Accurate Calculation of Weak Intermolecular Interactions

by

Jonathan Waldrop

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
August 3, 2019

Keywords: quantum chemistry, intermolecular interactions, symmetry-adapted perturbation theory

Approved by

Konrad Patkowski, Chair, Associate Professor of Chemistry and Biochemistry
Vincent Ortiz, Ruth W. Molette Professor of Chemistry and Biochemistry
German Mills, Professor of Chemistry and Biochemistry
Rik Blumenthal, Associate Professor of Chemistry and Biochemistry
Andrew Adamczyk, Assistant Professor of Department of Chemical Engineering

Abstract

A variety of weak intermolecular interactions is investigated using both high-accuracy and approximate methods. A high accuracy potential energy curve for the interaction of two krypton atoms was produced as a prerequisite for the computation of the most important thermophysical properties of the krypton gas. To obtain the desired spectroscopic accuracy around the potential miminum, corrections to the "gold standard" CCSD(T) were calculated, including higher-orders of coupled-cluster theory and the inclusion of relativistic effects. As an additional requirement for the computation of the second dielectric virial coefficient, similar steps were taken to produce an accurate analytic form for the interaction-induced isotropic pair polarizability for the krypton dimer. The resulting potential and polarizability functions were among the most accurate in the literature at time of publishing and the thermophysical properties were in marked agreement with experimental data. The interactions of CO_2 and models of metal-organic frameworks were explored using a variety of wavefunction and density functional methods as a way to evaluate the adequacy of these methods, and the potential multireference character of such models was considered. The interaction energy is decomposed using symmetry-adapted perturbation theory (SAPT) to provide insight into the character of the interactions. Lastly, recent work in the advancement and expansion of SAPT is detailed. This work includes the implementation of the second-order SAPT complete exchange terms using generalized Coulomb and exchange matrices and density-fitting, as well as the extension of the recent first-order spin-flip SAPT exchange energy beyond the single-exchange approximation.

Acknowledgments

I would first like to acknowledge the influence of my advisor, Dr. Konrad Patkowski. It is safe to say he taught me almost everything I know on the topics of computational and quantum chemistry. He has provided me with opportunities and challenges that have made me a better scientist and chemist, and has always been approachable and helpful.

I would also like to thank the Professors of my committee: Dr. Vincent Ortiz, Dr. German Mills, and Dr. Rik Blumenthal. Each has provided their own unique insight and broadened my view of the field of chemistry.

I would like to thank the members of the Patkowski Research Group, past and present, for their advice and assistance throughout my time here. Special mention goes to my contemporaries and office mates, Monika Kodrycka and Reza Hemmati, who served as helpful companions through our shared time at Auburn.

Lastly, I would like to thank my parents, John and Margaret, and my brother, Jacob. My parents always stressed the importance of education and supported my academic career. They also kept me on track at points where I would have gladly given up. My brother has always been a helpful sounding board for difficult topics.

Table of Contents

Abstract ii

Acknowledgments iii

List of Abbreviations ix

1 Introduction and Background 1

1.1 Electronic Structure Theory 2

1.1.1 Hartree-Fock Theory 3

1.1.2 Electron Correlation and Full Configuraton Interaction 6

1.1.3 Dynamic Correlation Methods 8

1.1.4 Nondynamical Correlation Methods 10

1.1.5 Density Functional Theory 12

1.2 Weak Intermolecular Interactions 13

1.2.1 Definitions of Interaction Energy 13

1.2.2 Basis Set Completeness 14

1.2.3 Method Selection 16

1.2.4 Symmetry-Adapted Perturbation Theory 18

1.2.5 Applications of Potential Energy Surfaces 22

1.3 Outline of Dissertation 22

2 Theoretical Calculations for the Kr Gas 24

2.1 Potential Energy Curve of the Kr Dimer 24

2.1.1 Investigation of Accuracy at Near-Minimum . . . 24
2.1.2 Fitting Analytic Potential . . . 27
2.2 Interaction-Induced Isotropic Pair Polarizability of the Kr Dimer . . . 30
2.2.1 Investigation of Accuracy at Near-Minimum . . . 30
2.2.2 Fitting an Analytic Function . . . 32
2.3 Evaluation of Computed Properties . . . 34

3 Investigation of Metal-Organic Framework Adsorption of CO_2 . . . 37
3.1 Model Structures of Metal-Organic Frameworks . . . 37
3.2 Single-Reference and Multireference Methods . . . 39
3.3 SAPT Decomposition . . . 41
3.4 Density Functional Methods . . . 43
3.5 Summary . . . 46

4 Developments in Symmetry-Adapted Perturbation Theory . . . 48
4.1 Second-Order Complete Exchange . . . 49
4.1.1 AO Formulation of the Complete $E^{(20)}_{exch-ind}$ . . . 52
4.1.2 AO/DF Formulation of the Complete $E^{(20)}_{exch-disp}$ . . . 55
4.2 First-Order SF-SAPT Complete Exchange . . . 59
4.2.1 Small Test Systems . . . 62
4.2.2 Pancake Bonded Systems . . . 63

5 Conclusions . . . 65
5.1 Future and Outlook . . . 66

References . . . 67

Appendices . . . 74

A Accurate *ab initio* potential for the krypton dimer and transport properties of the low-density krypton gas 75

B Accurate virial coefficients of gaseous krypton from state-of-the-art *ab initio* potential and polarizability of the krypton dimer 86

C Interactions of CO_2 with Metal-Organic Framework Models **[Undergoing Revision]** 97

D Spin splittings from first-order symmetry-adapted perturbation theory without single-exchange approximation 115

List of Figures

2.1 Contributions to the *ab initio* krypton dimer interaction energy and the total interaction energy (cm^{-1}) as functions of interatomic distance R (Å). 28

2.2 The *ab initio* values of E_{int} and the fitted Kr dimer potential (cm^{-1}) as functions of the interatomic distance R (Å). 29

2.3 Contributions to the interaction-induced isotropic polarizability $\Delta\alpha_{ave}$ for the krypton dimer and the total $\Delta\alpha_{ave}$ (a_0^3) as functions of the interatomic distance R (Å). 32

2.4 The *ab initio* values of $\Delta\alpha_{ave}$ and the fitted Kr dimer isotropic interaction-induced pair polarizability (a_0^3) as functions the interatomic distance R (Å). . . 33

3.1 SAPT0/aTZ interaction decomposition results (kcal/mol) for the various model systems with each metal. The Zn^{2+} results come from RHF-based SAPT, while the Mn^{2+} and Co^{2+} results are ROHF-based SAPT. 42

3.2 Comparison with benchmark of the DFT/aDZ E_{int}^{CP} (kcal/mol) for the diformate model-CO_2 systems. The bar colors indicate the method, while the hatchings indicate the type of empirical dispersion added. The lines bisecting the bars correspond to the result of that method in the aTZ basis set. 44

3.3 Comparison with benchmark of the DFT/aDZ E_{int}^{CP} (kcal/mol) for the paddlewheel model-CO_2 systems. The bar colors indicate the method, while the hatchings indicate the type of empirical dispersion added. The lines bisecting the bars correspond to the result of that method in the aTZ basis set. 44

3.4 Comparison with benchmark of the DFT/aDZ E_{int}^{CP} (kcal/mol) for the BDC model-CO_2 systems. The bar colors indicate the method, while the hatchings indicate the type of empirical dispersion added. The lines bisecting the bars correspond to the result of that method in the aTZ basis set. 45

3.5 Comparison with benchmark of the DFT/aDZ E_{int}^{CP} (kcal/mol) for the FDC model-CO_2 systems. The bar colors indicate the method, while the hatchings indicate the type of empirical dispersion added. The lines bisecting the bars correspond to the result of that method in the aTZ basis set. 45

4.1 The staggered conformation of the phenalenyl radical dimer, with carbon atoms colored gray and hydrogen atoms white. Sideview on left and topdown on right. 64

List of Tables

4.1 Average timings (s) for the calculation of the two second-order exchange corrections of the A24 dataset in either the MO or AO form. The column labeled AO_1 refers to the initial AO formulation for $E^{(20)}_{exch-ind}$, while AO_2 refers to the form based on the use of the modified HF coefficients. 57

4.2 Signed relative errors (%) of the single-exchange approximation for the two second-order exchange corrections for the A24 dataset. The values presented are for the minimum intermolecular separation R_0 and $0.8R_0$. Repeated pairs of molecules represent different geometric configurations.[97] 58

List of Abbreviations

AE All-Electron

AO Atomic Orbital

BDC Benzenedicarboxylate

BSSE Basis Set Superposition Error

CASPT2 Second-Order Complete Active Space Perturbation Theory

CASSCF Complete Active Space Self Consistent Field

CBS Complete Basis Set

CC Coupled-Cluster Theory

CI Configutation Interaction

CP Counterpoise Correction

DF Density-Fitting

DFT Denisty Functional Theory

DKH Douglas-Kroll-Hess

FC Frozen-Core

FCI Full Configuration Interaction

FDC Furandicarboxylate

GGA Generalized Gradient Approximation

HF Hartree-Fock Theory

KS Kohn-Sham Theory

MCSCF Multiconfigurational Self-Consistent Field Theory

MO Molecular Orbital

MOF Metal-Organic Framework

MP Møller-Plesset Perturbation Theory

MRCC Multireference Coupled-Cluster Theory

MRCI Multireference Configuration Interaction

MRPT Multireference Perturbation Theory

NEVPT2 Second-Order n-Electron Valence Perturbation Theory

NOONs Natural Orbital Occupation Numbers

PEC Potential Energy Curve

PES Potential Energy Surface

RS Rayleigh-Schrödinger Perturbation Theory

SAPT Symmetry-Adapted Perturbation Theory

SCF Self-Consistent Field Theory

SF-SAPT Spin-Flip Symmetry-Adapted Perturbation Theory

SRE Signed Relative Error

SRS Symmetrized Rayleigh-Schrödinger Perturbation Theory

X2C Exact Two-Component

Chapter 1

Introduction and Background

Electronic structure theory, and quantum mechanics in general, provide a unique approach to the elucidation of chemical problems. Through the application of these theories, a number of interesting chemical phenomena, which are difficult or impossible to investigate experimentally, can be evaluated with increasingly high levels of accuracy. Additionally, these *ab initio* methods can be used to calculate experimentally determinable values for comparison and validation. With the growth of computing power, these techniques have been adapted and implemented in a number of computer codes, simplifying the evaluation of otherwise complicated formulas. *Ab initio* methods and their computational implementations are not a silver bullet for all questions and have a number of general issues, with each specific theory or application providing its own potentially unique concerns. Below, Section 1.1 describes a number of electronic structure theory methods relevant to the remainder of this dissertation.

An area where the application of *ab initio* methods has been successful is the investigation of weak intermolecular interactions. This class of interactions, sometimes referred to as noncovalent interactions, covers a somewhat varied collection of interactions derived from both classical and quantum mechanical origins. The quantum character of some of these interactions necessitates an *ab initio* treatment to be correctly described, and a correlated method at that. Two specific approaches to the investigation of weak intermolecular interactions will be discussed in Section 1.2, along with the considerations required to evaluate and ensure their accuracy.

1.1 Electronic Structure Theory

The most common interest of electronic structure theory[1] is the solution of the time-independent Schrödinger Equation

$$H\Psi = E\Psi \tag{1.1}$$

where H is the non-relativistic Hamiltonian collecting all interactions between particles in the system, Ψ is a wavefunction of the system and an eigenfunction of the Hamiltonian, and E is the energy eigenvalue associated with the wavefunction. The full Hamiltonian is represented in atomic units as

$$H = -\sum_{A} \frac{1}{2M_A}\nabla_A^2 - \sum_{i} \frac{1}{2}\nabla_i^2 - \sum_{iA} \frac{Z_A}{r_{iA}} + \sum_{A\neq B} \frac{Z_A Z_B}{r_{AB}} + \sum_{i\neq j} \frac{1}{r_{ij}} \tag{1.2}$$

where the lowercase indices iterate over the electrons in the system, the uppercase indices iterate over the nuclei of the system, Z_A is the charge of a given nucleus, M_A is its mass, r_{xy} is the distance between two particles, and ∇_x^2 is the Laplacian of a particle. The first two terms are the kinetic energies of the nuclei and electrons, respectively, and the last three terms are the electron-nuclei, nuclei-nuclei, and electron-electron Coulombic interactions, respectively.

The Born-Oppenheimer Approximation proposes that since the nuclei of a system move at a much slower rate than the electrons, the nuclei can be considered to be fixed in place with the electrons moving around them.[2] Conversely, it can also be said that the electrons in the system move so quickly as to instantaneously adjust to any motion of the nuclei. This approximation leads to a separation of the wavefunction and Hamiltonian in terms of the electronic and nuclear parts. The electronic part is explicitly dependent on the electronic coordinates r and parametrically dependent on the nuclear coordinates R and has the electronic Hamiltionian

$$H_{elec} = -\sum_{i} \frac{1}{2}\nabla_i^2 - \sum_{iA} \frac{Z_A}{r_{iA}} + \sum_{i\neq j} \frac{1}{r_{ij}}. \tag{1.3}$$

The action of this Hamiltonian on the electronic part of the wavefunction produces an electronic energy, E_{elec}. It is helpful to define an operator h_i as the sum of the first terms in Eq. (1.3)

to simplify this equation into a one-electron and a two-electron part. The total energy, E_{tot}, is the sum of the electronic energy and the nuclear repulsion energy, $W_{AB}(R) = \sum_{A \neq B} \frac{Z_A Z_B}{r_{AB}}$, and serves as the potential for the nuclear Hamiltonian in the Born-Oppenheimer approximation. This potential is called the potential energy surface (PES), and its constituent one-dimensional slices are called potential energy curves (PECs). From here, we begin addressing ways of solving for the electronic energy. Much of the contents herein are drawn from a selection of texts on this topic, particularly Refs 1, 3, and 4.

1.1.1 Hartree-Fock Theory

Hartree-Fock (HF) theory, also called Self-Consistent Field (SCF) theory,[5–7] is the basis for a number of wavefunction methods and the first approximation to a solution for the electronic Schrödinger Equation

$$H_{elec}\Psi_{elec} = E_{elec}\Psi_{elec}, \tag{1.4}$$

where Ψ_{elec} is the electronic part of the total wavefunction. To be able to solve for E_{elec}, the form of Ψ_{elec} first needs to be defined.

The electronic wavefunction is a function of the spatial and spin coordinates of the electrons and, according to the Pauli principle, should be antisymmetric with the interchange of spatial and spin coordinates of any two electrons. These properties are enforced by representing the N-electron wavefunction as a Slater determinant built of one-electron spinorbitals

$$\Psi_{elec} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(r_1) & \chi_1(r_2) & \dots & \chi_1(r_N) \\ \chi_2(r_1) & \chi_2(r_2) & \dots & \chi_2(r_N) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_N(r_1) & \chi_N(r_2) & \dots & \chi_N(r_N) \end{vmatrix} \tag{1.5}$$

where $\{\chi_N\}$ are a set of orthonormal molecular spinorbitals.[8] In general, a spinorbital is the product of a spatial orbital ψ and one of the spin functions α or β. This set of spin functions is complete and orthonormal. In the case of a closed-shell system, the set of spinorbitals is

the product of some set of spatial orbitals $\{\psi_{\frac{N}{2}}\}$ and each of the spin functions to produce $\frac{N}{2}$ doubly-occupied molecular orbitals (MOs).

It is common to represent the spatial orbitals as linear combinations of a basis set of hydrogenic atomic orbitals (AOs)

$$\psi_i = \sum_{\mu} C_{i\mu}\phi_{\mu} \tag{1.6}$$

where $\{\phi_\mu\}$ are a set of μ atomic orbitals, and $C_{i\mu}$ is the coefficient of the μth atomic orbital in the ith molecular orbital. If this basis set were complete, the molecular orbitals could be represented exactly, but it is practically infeasible to work with a complete basis set. Additionally, the atomic orbitals are normalized, but are not necessarily orthogonal to those centered on different nuclei. The most preferable form of these basis functions would be a Slater function, which best replicates the expected qualities of the molecular orbitals. Unfortunately, Slater functions lead to difficulties during the computation of common terms needed for electronic structure calculations, specifically the two-electron integrals that will appear later in this section. To overcome this issue, the Slater orbitals are approximated by a linear combination of Gaussian orbitals, which better facilitate the calculation of the two-electron integrals.

HF theory takes advantage of the variational principle and the application of the linear variational method. The variational principle can be summarized as the fact that any normalized wavefunction ψ_A that is an approximation to the ground state wavefunction ψ_0 of a system will have an energy E_A, calculated as the expectation value of the Hamiltonian, that is an upper bound of the true ground state energy E_0,

$$E_A = \int \psi_A^*(r) H \psi_A(r) dr = \langle \psi_A | H | \psi_A \rangle \geq E_0. \tag{1.7}$$

Based on this principle, the best approximate function can be obtained by minimizing its corresponding energy. When the MOs of the Slater determinant are represented as a linear combination of AOs, the optimal coefficients can be determined by minimizing the Roothaan equations[9]

$$\mathbf{FC} = \mathbf{SC}e, \tag{1.8}$$

where $\mathbf{C}$ are the coefficients relating the MOs to the AO basis, e are the energies of the MOs, $\mathbf{S}$ is an overlap matrix with elements $\langle\mu|\nu\rangle$, and $\mathbf{F}$ is the Fock matrix. The Fock matrix is the HF approximation to the electronic Hamiltonian. The elements of the matrix are

$$\langle\phi_\mu|F|\phi_\nu\rangle = \langle\phi_\mu|h|\phi_\nu\rangle + \langle\phi_\mu|J|\phi_\nu\rangle - \langle\phi_\mu|K|\phi_\nu\rangle \tag{1.9}$$

where $\langle\phi_\mu|h|\phi_\nu\rangle$ is the expectation value of the one-electron operator h,

$$\begin{aligned}\langle\phi_\mu|J|\phi_\nu\rangle = J_{\mu\nu} &= \sum_{j}^{MOs}\sum_{\lambda\sigma}^{AOs} C_{j\lambda}C_{j\sigma}\int \phi_\mu^*(r_1)\phi_\nu(r_1)\frac{1}{r_{12}}\phi_\lambda^*(r_2)\phi_\sigma(r_2)dr_1dr_2 \\ &= \sum_{j}^{MOs}\sum_{\lambda\sigma}^{AOs} C_{j\lambda}C_{j\sigma}\,\langle\mu\lambda|\nu\sigma\rangle \\ &= \sum_{\lambda\sigma}^{AOs} P_{\lambda\sigma}\,\langle\mu\lambda|\nu\sigma\rangle\,,\end{aligned} \tag{1.10}$$

and

$$\langle\phi_\mu|K|\phi_\nu\rangle = K_{\mu\nu} = \sum_{\lambda\sigma} P_{\lambda\sigma}\,\langle\mu\lambda|\sigma\nu\rangle\,. \tag{1.11}$$

The term in Eq. (1.10) is called a Coulomb integral, and also defines the two-electron integrals $\langle\mu\lambda|\nu\sigma\rangle$ and the density matrix $P_{\lambda\sigma}$. Eq. (1.11) is the exchange integral that arises from the Slater determinant form of the wavefunction. The form of these integrals is the origin of HF's classification as a mean-field method, since each electron experiences an average interaction with all other electrons and is not correlated to their instantaneous positions.

At this point, two difficulties appear in the HF method. The first is the presence of $\mathbf{S}$ in Eq. (1.8), making this a pseudoeigenvalue problem. This issue is due to the lack of orthogonality in the basis functions. Diagonalizing $\mathbf{S}$ reduces it to identity, giving an orthonormal basis and simplifying the equations to an eigenvalue problem that is solvable. The greater issue is the presence of the coefficients $C_{i\mu}$ in the density matrix and the density matrix's presence in the Fock matrix. The HF method is attempting to find the optimal values of $C_{i\mu}$, which are in turn required in the equations being solved. A remedy is found in the iterative solving of the equations. With some initial values for $C_{i\mu}$, the equations can be solved for a new set of

coefficients. These coefficients are used to construct a new density matrix and calculate the HF electronic energy

$$E_{elec}^{HF} = \sum_{\mu\nu} \frac{1}{2} P_{\mu\nu}(h_{\mu\nu} + F_{\mu\nu}). \tag{1.12}$$

This new density matrix is used to solve for another new set of coefficients and the process continues until the results between consecutive iterations are determined to be the same within a predetermined criterion.

HF is commonly used in its restricted formulation RHF, where the equations are simplified based on the assumption that all orbitals are doubly occupied and that the spatial parts of orbitals with α and β are identical. The restricted formalism can be extended to cases with singly occupied orbitals as restricted open-shell HF (ROHF), or the system can be treated in an unrestricted way as UHF.

1.1.2 Electron Correlation and Full Configuraton Interaction

HF is a good first approximation to the electronic structure of the ground state, even with an incomplete basis set and its mean field treatment of electron correlation. But it has been found that the lack of electronic correlation often leaves HF results quantitatively inaccurate, and qualitatively inaccurate in some cases. A simple example of HF's failure is the interaction within any noble gas dimer. For these systems, HF predicts a repulsive potential at short range that decays to zero at infinite separation, while it is known that there is a shallow attractive well at some short distance for each dimer. The correlation energy that is absent in the HF result is defined as

$$E_{corr} = E_0 - E_{HF} \tag{1.13}$$

i.e. as the difference between the exact ground state energy and the HF approximation. As the E_{HF} is always an upper bound of E_0, the correlation energy E_{corr} will always be negative. There are a number of methods that attempt to recover some or all of E_{corr}, some of which will now be discussed.

The completion of an HF calculation provides a number of one-electron molecular spinorbitals equal to two times the number of basis functions used. While it is possible to use a basis

set with the minimum number of basis functions to accommodate the N electrons, most calculations are performed using a basis set with more functions than the number of electrons in the system. As the number of basis functions approaches infinity, the molecular spinorbitals are able to approach their exact form. The HF approximation to the ground state of the system is the wavefunction where the N electrons occupy the N molecular spinorbitals with the lowest energies. While this is the best approximation for a single determinant representation of the wavefunction, it is not necessarily the best approximation that can be constructed from these components.

Using the molecular spinorbitals from the HF calculation, it is possible to construct a number of alternate determinants where one or more molecular spinorbitals that are occupied in the HF ground state have been replaced with an equivalent number of unoccupied, or virtual, spinorbitals. These determinants can be used as a basis for the construction of a new trial function that is a linear combination

$$\Psi = c_0\psi_0 + c_1\psi_1 + c_2\psi_2 + ... \tag{1.14}$$

where ψ_0 is the HF wavefunction, ψ_1 is representative of all determinants where one spinorbital occupied in the HF wavefunction has been replaced with a virtual spinorbital, ψ_2 is representative of two such interchanges, and the expansion continues until all N electrons have been moved into virtual spinorbitals. The coefficients in Eq. (1.14) can then be optimized in a linear variational scheme as in HF to minimize the electronic energy.

This approach is called Configuration Interaction (CI), or Full CI (FCI) in the case where all possible determinants are included in the expansion. FCI is the exact result for a given basis set, containing all information that can be recovered. FCI is also one of the most computationally expensive methods, as the number of determinants that need to be included grows as $\binom{2M}{N}$ with the number of electrons N and the basis set size M. This scaling becomes unreasonable for conventional computers very quickly, and so approximations must be made to alleviate the cost growth while maintaining as much accuracy as possible.

1.1.3 Dynamic Correlation Methods

The FCI coefficients provide a useful framework to discuss electronic correlation. Within the large expansion of determinants, it is often seen that the HF ground state has the largest coefficient and that all others contribute some smaller part to the FCI wavefunction. These small contributions are associated with the instantaneous correlation of the electrons to each other's position. This kind of correlation is referred to as dynamic correlation. It is also generally the case that coefficients get smaller with more virtual excitations separating a determinant from the HF wavefunction. Based on this observation, a hierarchy of methods can be defined by limiting the expansion in Eq. (1.14) based on the number of virtual excitations. As an example, the most common of these methods is Configuration Interaction with Single and Double Excitations (CISD), where the expansion is truncated to only the terms explicitly shown in Eq. (1.14). The inclusion of more and more excitations (e.g. CISDT or CISDTQ) improves the calculated results toward FCI, but likewise increases the cost of the calculation.

A somewhat related, alternative approximation is the Coupled-Cluster Theory (CC).[10] In CC, the exact wavefunction is constructed as

$$\Psi = e^{\mathbf{T}}\psi_0 = \Big(1 + \mathbf{T} + \frac{\mathbf{T}^2}{2!} + \frac{\mathbf{T}^3}{3!} + ...\Big)\psi_0 \tag{1.15}$$

where $\mathbf{T}$ is the cluster operator

$$\mathbf{T} = \mathbf{T}_1 + \mathbf{T}_2 + \mathbf{T}_3 + ... \tag{1.16}$$

$\mathbf{T}_n$ is an operator that produces all possible n-tuple excited determinants from the HF reference state. If the complete cluster operator is used this method becomes equivalent to FCI, but the cluster operator is often truncated to a particular excitation level similar to the CI truncations. Similar to CISD, truncation to double excitations leads to coupled-cluster with singles and doubles (CCSD). When the cluster operator is truncated to a given number of excitations, the Taylor expansion in Eq. (1.15) leads to an approximation to higher excitations as products of

lower excitations (e.g. a quadruple excitation from the product of two double excitations). It is this feature of CC that provides its key improvement over CI, size consistency.

Size consistency is best explained through an example like the previously mentioned noble gas dimers. The interaction energy (discussed in greater detail in Section 1.2) between two noble gas atoms at a particular distance is the difference between the total energy of the dimer at that distance and the energies of the lone atoms. If this energy is calculated using CISD, each of the calculations contains up to double excitations. When the difference is calculated, the monomer values can contain cases where both lone atoms are doubly excited. These cases correspond to a quadruple excitation of the dimer and are therefore unreachable for the given level of theory, resulting in an uneven handling of the dimer and monomers and skewing the interaction energy. These same values calculated at the CCSD level would contain the same possible excitations, as the double excitations for both monomers would appear together in the Taylor expansion for the dimer.

Møller-Plesset perturbation theory (MP), a particular application of the more general Rayleigh-Schrödinger perturbation theory (RS), takes a slightly different approach to the calculation of the correlation energy.[11] The basis of perturbation theory is that the solution of a difficult problem can be found by starting from a similar problem that is easier to solve. This solvable part is then perturbed into the problem of interest. This process can be expressed as the partitioning

$$H = H_0 + \lambda V, \tag{1.17}$$

where H is the difficult operator of interest, H_0 is the similar, solvable operator, V is the perturbation operator that contains the difference between the two problems, and λ is an ordering parameter. For the calculation of the electronic energy, this can be done by defining the unperturbed (simpler) problem as the HF problem and the perturbation as the difference between HF and the exact result. By inserting Eq. (1.17) into Eq. (1.4), performing a Taylor expansion around λ, and grouping the resulting equations into equalities involving like powers of λ, one obtains up to an infinite number of corrections to the unperturbed problem. The unperturbed

energy and the first two corrections are acquired from the following terms

$$E_0^{(0)} = \langle \psi_0^{(0)} | H_0 | \psi_0^{(0)} \rangle \tag{1.18}$$

$$E_0^{(1)} = \langle \psi_0^{(0)} | V | \psi_0^{(0)} \rangle \tag{1.19}$$

$$E_0^{(2)} = \langle \psi_0^{(0)} | V | \psi_0^{(1)} \rangle \,, \tag{1.20}$$

where $E_0^{(n)}$ and $\psi_0^{(n)}$ are the nth order energy and wavefunction corrections, respectively. It can be shown that the first-order MP correction, $E_0^{(0)} + E_0^{(1)}$, is equivalent to the HF energy, and that improvements are seen at the second-order and onward (MPn with $n \geq 2$). The second-order correction to the electronic energy, a method called MP2, is known to have the form

$$E_0^{(2)} = \frac{1}{4} \sum_{ijab} \frac{| \langle ij||ab \rangle |^2}{e_i + e_j - e_a - e_b} \tag{1.21}$$

where i and j index the occupied orbitals, a and b index the virtual orbitals, and the integrals $\langle ij||ab \rangle$ are antisymmetric two-electron integrals equal to $\langle ij|ab \rangle - \langle ij|ba \rangle$.

As an aside before moving on, it should be noted that perturbation theory can be applied in a number of other scenarios. The application of this technique to the calculation of interaction energies will be discussed in section 1.2.4. Another application is as a further improvement of truncated CC methods. For instance, CCSDT can be approximated in a perturbative manner from a CCSD calculation. The resulting method is termed CCSD(T), where the parentheses imply that the triples have been calculated perturbatively.[12] CCSD(T) is a considerably faster calculation than CCSDT and is known to give high quality results, earning it the nickname as the "gold standard" of computational chemistry.[13–15]

1.1.4 Nondynamical Correlation Methods

Returning to the FCI coefficients, it is possible that the HF wavefunction is not the only determinant with a considerable contribution to the FCI wavefunction. For some systems, the ground state cannot be adequately approximated by a single determinant due to degenerate or

near degenerate states. An example of such a system is the singlet O_2 molecule, where one $\pi^\star$ orbital is doubly occupied and the other is empty. While both electrons occupy the same $\pi^\star$ orbital, they are equally likely to be found in either $\pi^\star$ orbital. Therefore, both configurations have equivalent contribution to the FCI wavefunction and ignoring either is a poor approximation to this result. This kind of correlation is referred to as strong, static, or nondynamical correlation.

Multiconfigurational Self-Consistent Field Theory (MCSCF) can be viewed as an improvement on SCF that goes beyond a single determinantal treatment. It is reasonable to describe the basic idea as a highly selective CI truncation. The most important electronic configurations, determined by chemical or physical intuition, are selected and the energy minimized with respect to their coefficients. One caveat solving for such a multireference problem is the adequacy of the MOs for the description of all important states. Returning to the O_2 example, assume that the pair of electrons was placed arbitrarily in one of the π orbitals and the HF calculation performed. Now, the occupied π orbital was optimized to minimize the energy, while the other π orbital was not necessarily as well optimized. The configuration made by moving the electrons from the optimized one to the unoptimized one is therefore not equivalent. The solution to this issue is the simultaneous optimization of the orbitals along with the CI coefficients, which is specifically the case in MCSCF.

There are various methods for selecting the configurations to include in the MCSCF calculation. The method called Complete Active Space Self Consistent Field theory (CASSCF) selects the configurations by identifying a number of electrons and a specific set of MOs, called the active space, and producing all possible configurations from the distribution of those electrons within the active space.[16] This selection can be represented as (*# of electrons*, *# of orbitals*), e.g. two electrons distributed in four orbitals would be (2, 4).

When performed with the correct configurations, MCSCF is able to capture the most important static correlation, but is still missing the dynamic correlation from other configurations. Multireference Configuration Interaction (MRCI) is the adaptation of the truncated CI methods to an MCSCF ground state, and provides a way to capture both dynamic and static correlation. This specifically denotes truncated methods, as a FCI calculation will reach the

same result from an HF or MCSCF reference as long as the same basis set is used. Additionally, the other methods mentioned for the calculation of dynamic correlation have formulations based on the usage of a multiconfigurational reference. Second-Order Complete Active Space Perturbation Theory (CASPT2)[17–19] and Second-Order n-Electron Valence Perturbation Theory (NEVPT2)[20–22] are two implementations of Multireference Perturbation Theory (MRPT). These formulations of MRPT and the Multireference Coupled-Cluster Theory (MRCC)[23] are similarly motivated to their single reference equivalents, though some additional complexity arises from dealing with multiple configurations and leads to increased computational cost. Some of the complexities will be discussed in later sections.

1.1.5 Density Functional Theory

The last of the general purpose electronic structure methods that will be mentioned here is also the most different from the others. Density Functional Theory (DFT) arose from a desire to reformulate the problem of finding the energy of a molecular system in terms of a physical observable, the electronic density $\rho(r)$, instead of the wavefunction. The electron density is the square of the wavefunction, which can be expressed as

$$\rho(r) = 2\sum_{i}^{N/2} |\psi_i(r)|^2. \tag{1.22}$$

The works of Hohenberg, Kohn, and Sham proved[24] and provided[25] a way to determine the energy as a functional of the electronic density

$$E[\rho(r)] = T[\rho(r)] + V_{ne}[\rho(r)] + V_{ee}[\rho(r)] + V_{xc}[\rho(r)] \tag{1.23}$$

where $T[\rho(r)]$ is the kinetic energy of the non-interacting electrons, $V_{ne}[\rho(r)]$ is the nuclear-electronic attraction, $V_{ee}[\rho(r)]$ is the classical electron-electron repulsion, and the exchange-correlation functional $V_{xc}[\rho(r)]$ collects a number of classical and quantum effects. The resulting method is called Kohn-Sham DFT, or Kohn-Sham theory (KS).

If the exact form of the exchange-correlation functional were known, DFT would be able to provide the exact solution for a given system.[26] That said, the exact form of this functional is not known, and in its place are a number of approximations in an ever growing hierarchy. The general process of calculating the electronic energy with a given functional is similar to HF in both methodology and cost. The results can be vastly superior due to the correlation effects in DFT. Care does need to be taken with DFT methods, as the approximate exchange-correlation functionals are difficult to improve in any systematic way and their results can vary in quality from system to system.

1.2 Weak Intermolecular Interactions

Weak intermolecular interactions are ubiquitous in the world at large. They have key influence on such properties as crystal structure, protein structure, and adsorption. Despite the widespread importance of this class of interactions, they can be distinctly difficult to handle in a computational fashion. The size of these values is often many orders of magnitude smaller than the energies of the systems for which they are being calculated. To accurately model weak interactions, a number of factors have to be taken into consideration and will be discussed here.

1.2.1 Definitions of Interaction Energy

The interaction energy E_{int} for the interaction of two monomers A and B can be defined as

$$E_{int} = E_{AB}(AB) - E_A(A) - E_B(B), \tag{1.24}$$

where $E_{AB}(AB)$, $E_A(A)$, and $E_B(B)$ are the energies of the dimer and monomers. In this context, "dimer" will refer to an interacting complex of "monomer" subsystems, regardless of the homogeneity or heterogeneity or the complex. The subscripts imply that the calculations were performed in the basis set of the dimer or one of the monomer subsystems. This definition is the mathematical representation of the statement that the interaction energy is the difference between the energy of the interacting monomers and the monomers at an infinite separation. Therefore, to calculate the interaction energy, one can perform three calculations to obtain some

approximation of these values and calculate this difference. This procedure is referred to as the supermolecular method.

Somewhat obviously, it is important that the energies of the dimer and monomers are calculated in an equal fashion, otherwise the resulting value of E_{int} will be skewed in some way. One potential imbalance inherent to this first definition is the selection of the basis sets for the dimer and monomers. Eq. (1.24) states that the energies of the dimer and subsystems are calculated in their own basis set, i.e. the dimer calculation is performed in a larger basis set than either of the monomers. This means that the orbitals of the monomer calculations are more restricted than their counterparts in the dimer calculations. This issue is termed basis set superposition error (BSSE), and a solution is found in the counterpoise correction of Boys and Bernardi,[27]

$$E_{int}^{CP} = E_{AB}(AB) - E_{AB}(A) - E_{AB}(B) \tag{1.25}$$

where the terms on the right hand side are equivalent to the terms in Eq. (1.24) but with the common subscript implying that they are all calculated in the dimer basis set. As the size of the basis set approaches infinity, the corrected and non-corrected definitions become equivalent.

Another way of defining the interaction energy was used in the work described in Chapter 3,

$$E_{int} = E_{AB}(AB) - E_{AB}^{\infty}(AB). \tag{1.26}$$

This definition replaces the infinite reference in Eqs. (1.24) and (1.25), $(E(A) + E(B))$, with the value from a single calculation of the dimer where the monomers are at some large distance that approximates infinity, $E_{AB}^{\infty}(AB)$. The benefit of this version of the procedure is that it circumvents issues of size consistency in the treatment of the dimer and monomers.

1.2.2 Basis Set Completeness

On the topic of basis sets, the incompleteness of the basis set is one of the leading limitations in the accuracy of electronic structure calculations. The closer the basis set comes to the complete basis set limit (CBS), the closer each method comes to its exact result. The balance to this is that increasing the basis set size also increases the cost of a given calculation. To put this fact

into perspective, HF scales as $\mathcal{O}(N^4)$ with the basis set size N, while MP2 scales as $\mathcal{O}(N^5)$ and CCSD(T) scales as $\mathcal{O}(N^7)$. These steep scalings put substantial limitations on how close to the CBS limit one can get for a system, limiting some calculations to only a minimal basis set and outright preventing others. While the HF energy converges to the CBS limit quickly, correlation energy does so at a much slower rate. This slow convergence necessitates either large basis sets or some alternative way of reaching the CBS limit.

Extrapolation provides a way of approximating the CBS limit from smaller basis sets. Basis set families have been designed so that they successively capture more and more of the correlation energy. By performing calculations in two or more of these basis sets, the CBS limit can be approximated by extrapolating from these results. One such family is the Dunning-type correlation consistent basis sets, denoted as cc-pVXZ (XZ where X =D, T, Q, 5, ...)[28,29] The cardinal number X primarily indicates the number of basis functions used to represent the valence orbitals, but the Dunning-type sets follow a specific design pattern that is also conveyed in the cardinal number. If the energy of a system is calculated in consecutive basis sets, the CBS value of E_{corr} can be approximated as

$$E_{corr}^{CBS} \approx \frac{E_{corr}^{mZ} m^3 - E_{corr}^{nZ} n^3}{m^3 - n^3} \tag{1.27}$$

where m and n are the cardinal numbers of the basis sets and $n = m - 1$.[30] This correlation energy is added to the HF energy from the higher basis set, as it is assumed to be sufficiently converged. An extrapolation to CBS from two given sets will be denoted as ($set1$, $set2$), where $set2$ has the larger cardinal number.

In the particular case of weak intermolecular interactions, there are some additional considerations. Because the interactions of interest are relatively long-ranged (compared to covalent bonds), the basis set needs to allow the electron density between monomers to interact at a distance. One way of accomplishing this is the use of diffuse basis functions that cover a wide spatial range. For Dunning-type sets, the prefix aug- implies the set has had diffuse, or augmentation, functions added to it, e.g. aug-cc-pVXZ (aXZ).

Another way of ensuring that the basis set is able to properly represent the electron density between the interacting systems is the addition of midbond functions.[31] These are additional basis functions that are added somewhere between the interacting systems, usually along the line between the centers of mass. The addition of midbond functions increases the coverage of the basis set in the space where the interaction is most likely to occur. The selection of what kind of functions and how many to use is not rigidly specific; however, one simple method is to use the basis functions for the hydrogen atom from whichever set is being used for the rest of the system. When used with Dunning sets, this prescription with be signified by the notation cc-pVXZM (XZM).

1.2.3 Method Selection

Returning to the discussion of potential imbalances in the methods used to calculate the energies, the first and most important fact is that the energies of the dimer and monomers must be calculated using the same level of theory. If E_{int} is calculated as the difference between $E(AB)$ calculated with HF, $E(A)$ calculated with MP2, and $E(B)$ calculated with FCI, the resulting value is effectively useless due to the wildly varying treatment of correlation in these systems. Stated more succinctly, the usefulness of the supermolecular method is dependent on a cancellation of errors between its component values. The usage of multireference methods introduces another way of potentially unbalancing the treatment of the components. Care must be taken to ensure that the active space of the dimer contains all orbitals corresponding to those present in the monomer active spaces.

Understanding that the method selection needs to be balanced for the component calculations, the next consideration is the quality of the results being acquired. Because E_{int} is much smaller in magnitude than $E(AB)$, $E(A)$, and $E(B)$, it is critical that these values be calculated with a high degree of accuracy to ensure that E_{int} is not drowned out by error. As stated above, FCI is the most accurate method, but is too computationally costly to perform for all but the smallest systems. CCSD(T) is known to provide high accuracy results when calculated in a large enough basis set, and at a significantly lower cost than FCI. Even still, CCSD(T) in a reasonably sized basis set is also too costly for many systems of interest.

One way of reaching a desired level of accuracy is by starting from the best method and basis set available, and then building up a number of corrections from superior methods in smaller basis sets. These composite methods are predicated on the idea that the improvements of one method over another are similar between different basis sets. An example build up of such a composite method could begin with a calculation at the MP2/(aQZ, a5Z) level. This result can then be improved by the addition of a CCSD correction from the aQZ set

$$\delta_{CCSD}^{aQZ} = E_{corr}^{CCSD/aQZ} - E_{corr}^{MP2/aQZ}, \tag{1.28}$$

providing an approximation to CCSD/CBS. This result could be further improved by a correction to CCSD(T),

$$\delta_{CCSD(T)}^{aTZ} = E_{corr}^{CCSD(T)/aTZ} - E_{corr}^{CCSD/aTZ}, \tag{1.29}$$

and a correction to CCSDT(Q),

$$\delta_{CCSDT(Q)}^{aDZ} = E_{corr}^{CCSDT(Q)/aDZ} - E_{corr}^{CCSD(T)/aDZ}, \tag{1.30}$$

and so on. The final result would be described as the sum of these corrections, MP2/(aQZ, a5Z) + δ_{CCSD}^{aQZ} + $\delta_{CCSD(T)}^{aTZ}$ + $\delta_{CCSDT(Q)}^{aDZ}$. This kind of composite scheme can be used to account for a number of different corrections, including core-core and core-valence correlation, relativistic effects, or improvements to the Born-Oppenheimer Approximation.

As mentioned before, DFT can also provide results with considerably high accuracy at a low cost, but the different approximate functionals need to be tested for their adequacy in dealing with a particular type of system. This has been accomplished by benchmarking databases of molecular compounds and complexes with high levels of theory and comparing the results of the different functionals.[32] This practice provides a way of statistically determining the appropriateness of a functional for certain interactions or compounds, and provides a kind of confidence value to the result. To obtain high quality interaction energies from DFT, it is usually necessary to include some correction for dispersion. One such correction is Grimme's empirical pairwise correction, the most popular version being D3.[33–36] Updated parameters for this

correction were published recently, with the intention of improving the results and providing a more even treatment of intermolecular distances other than the minimum.[37]

1.2.4 Symmetry-Adapted Perturbation Theory

Turning away from the supermolecular method now, we introduce an alternative method of calculating the interaction energy. Symmetry-adapted perturbation theory (SAPT) builds from the same basic starting point as Møller-Plesset perturbation theory, but is focused on the interaction energy between monomers.[38] In contrast to the supermolecular method, SAPT calculates the interaction energy directly instead of as a difference of total energies. Consequently, SAPT is not burdened by several of the considerations outlined in the preceding sections. Additionally, SAPT can provide a decomposition of the interation energy into physically meaningful components. These include electrostatics (the Coulombic interactions between the charged particles of the monomers), induction (the effect that the electric field of one monomer has on the other), and dispersion (the instantaneous development of multipole interactions arising from the correlated motion of electrons between monomers). These components are further accompanied by their exchange counterparts, arising from the Pauli exclusion principle.

SAPT partitions the Hamiltonian as

$$H = H_A + H_B + V, \tag{1.31}$$

where H_A is the Hamiltonian for monomer A, H_B is the Hamiltonian for monomer B, and V is the pertubation operator that collects all interactions between particles from A and B. The perturbation operator has the form

$$V = \sum_{AB} \frac{Z_A Z_B}{r_{AB}} - \sum_{j,A} \frac{Z_A}{r_{jA}} - \sum_{i,B} \frac{Z_B}{r_{iB}} + \sum_{i,j} \frac{1}{r_{ij}}, \tag{1.32}$$

where A are the nuclei of monomer A, B are the nuclei of monomer B, i are the electrons of monomer A, and j are the electrons of monomer B. In SAPT, the unperturbed dimer wavefunction is defined as the product of the wavefunctions of the monomers,

$$\Psi^{(0)} = \Psi_0^A \Psi_0^B, \tag{1.33}$$

making the zeroth-order energy the sum of the monomer ground state energies,

$$E^{(0)} = E_0^A + E_0^B. \tag{1.34}$$

From this point, the same RS corrections (Eqs. (1.19) and (1.20)) can be applied to SAPT to produce the first-order correction

$$\begin{aligned} E_{RS}^{(10)} &= \langle \Psi^{(0)} | V | \Psi^{(0)} \rangle \\ &= \sum_{AB} \frac{Z_A Z_B}{r_{AB}} - \sum_j \langle \psi_j | A | \psi_j \rangle - \sum_i \langle \psi_i | B | \psi_i \rangle + \sum_{ij} \langle ij | ij \rangle \\ &= V_0 + \sum_j A_{jj} + \sum_i B_{ii} + \sum_{ij} \langle ij | ij \rangle \,, \end{aligned} \tag{1.35}$$

where i indexes the spinorbitals of monomer A, j indexes the spinorbitals of monomer B, V_0 is the nuclear repulsion between monomers, A_{jj} is a one-electron integral over the nuclear potential of monomer A, and B_{ii} is a one-electron integral over the nuclear potential of monomer B. The negative signs in the second line of Eq. (1.35) are folded into the definitions of A_{jj} and B_{ii}. The second index in the superscript (10) denotes that the reference is at the HF level and no intramolecular correlation has been included. This first-order correction is the electrostatic energy $E_{elst}^{(10)}$ between the two monomers, as can be seen from the terms in Eq. (1.35). Furthermore, the second-order RS correction can be broken down into induction (both B$\rightarrow$A and A$\rightarrow$B) and dispersion. The RS corrections derived in this way are sometimes called the polarization terms.

To obtain the exchange components, the dimer wavefunction first needs to be inspected. The dimer wavefunction should be antisymmetric with respect to the interchange of coordinates for any two electrons, but the SAPT zeroth-order wavefunction is not antisymmetric with respect to the interchange of coordinates for electrons on different monomers. The dimer wavefunction has to be properly antisymmetrized to acquire the exchange energy. This symmetry adaptation can be accomplished in various ways based on where the symmetry enforcement enters the theory.[38] The most common method of symmetry adaptation is the Symmetrized Rayleigh-Schrödinger (SRS) theory,[39] where the wavefunction corrections are acquired from RS and an antisymmetrizer is introduced in the formulas for the energy corrections. The first-order SRS correction is

$$E_{SRS}^{(10)} = E_{elst}^{(10)} + E_{exch}^{(10)} = \frac{\langle \Psi^{(0)} | V\mathcal{A} | \Psi^{(0)} \rangle}{\langle \Psi^{(0)} | \mathcal{A} | \Psi^{(0)} \rangle} \tag{1.36}$$

where $\mathcal{A}$ is an antisymmetrizer. This correction is further represented as

$$\frac{\langle \Psi^{(0)} | V\mathcal{A} | \Psi^{(0)} \rangle}{\langle \Psi^{(0)} | \mathcal{A} | \Psi^{(0)} \rangle} = \frac{V_0 \mathcal{S} + \sum_{ir} B_{ir} \mathcal{S}^{ir} + \sum_{jr} A_{jr} \mathcal{S}^{jr} + \sum_{ijrs} \langle ij||rs \rangle \, \mathcal{S}^{ij,rs}}{\mathcal{S}}, \tag{1.37}$$

where r and s index occupied spinorbitals on both monomers, $\mathcal{S}$ is the determinant of the overlap matrix $\mathbf{S}$ of both monomers

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_{ii} & \mathbf{S}_{ij} \\ \mathbf{S}_{ji} & \mathbf{S}_{jj} \end{bmatrix} = \begin{bmatrix} \mathbf{1} & \mathbf{S}_{ij} \\ (\mathbf{S}_{ij})^T & \mathbf{1} \end{bmatrix}, \tag{1.38}$$

$\mathcal{S}^{ir}$ is a first cofactor of the overlap matrix, and $\mathcal{S}^{ij,rs}$ is a second cofactor of the overlap matrix. $\mathcal{S}^{ir}$ is constructed by deleting the row i and column r from the overlap matrix and multiplying the determinant of this reduced matrix by $(-1)^{i+r}$. The second cofactor extends this process to an additional row and column, and the new determinant is multiplied by $(-1)^{i+j+r+s}$. A second cofactor has the following relationship with the first cofactors,

$$\mathcal{S}^{ij,rs} = \frac{1}{\mathcal{S}} (\mathcal{S}^{ir} \mathcal{S}^{js} - \mathcal{S}^{is} \mathcal{S}^{jr}). \tag{1.39}$$

Cramer's Rule can be applied to relate the overlap matrix and its cofactors in the following way,

$$D_{ri} = \frac{\mathcal{S}^{ir}}{\mathcal{S}} \tag{1.40}$$

where D_{ri} are elements of the inverse of the overlap matrix, $\mathbf{D} = \mathbf{S}^{-1}$. Taking advantage of these relations, the terms in Eq. (1.37) can be rewritten as

$$E_{SRS}^{(10)} = V_0 + \sum_{ir} B_{ir} D_{ri} + \sum_{jr} A_{jr} D_{rj} + \sum_{ijrs} \langle ij||rs\rangle \left(D_{ri} D_{sj} - D_{si} D_{rj}\right) \tag{1.41}$$

Higher order SRS corrections provide the exchange counterparts to the second and higher order RS corrections.

The above method provides the complete $E_{exch}^{(10)}$, because the complete antisymmetrizer is used.[40] It is commonplace, however, to use an approximation of the complete antisymmetrizer

$$\mathcal{A} \approx \mathcal{A}_A \mathcal{A}_B (1 + \mathcal{P}), \;\; \mathcal{P} = -\sum_{ij} \mathcal{P}_{ij} \tag{1.42}$$

where $\mathcal{P}_{ij}$ are the single-exchange operators that exchange pairs of electrons between monomers. This is called the single-exchange, or $\mathcal{S}^2$, approximation because of the limitations it puts on the exchange between monomers.[38] While widely used in SAPT, the $\mathcal{S}^2$ approximation is known to break down at short intermolecular distances.

A hierarchy of SAPT results are derived from the number of terms used to calculate the interaction energy. The simplest of these is referred to as SAPT0,[41] defined as

$$E_{int}^{SAPT0} = E_{elst}^{(10)} + E_{exch}^{(10)} + (E_{ind}^{(20)} + E_{exch-ind}^{(20)} + \delta E_{HF}^{(2)}) + (E_{disp}^{(20)} + E_{exch-disp}^{(20)}). \tag{1.43}$$

The sum of the terms collected in the first set of parentheses is considered the induction energy, and the sum in the second set is considered the dispersion energy. The terms $E_{ind}^{(20)}$ and $E_{disp}^{(20)}$ are the induction and dispersion parts of the second-order RS correction, while $E_{exch-ind}^{(20)}$ and $E_{exch-disp}^{(20)}$ are their exchange counterparts from the second-order SRS correction. The term $\delta E_{HF}^{(2)}$ is the difference between the supermolecular HF E_{int} and the sum of the first four terms

on the R.H.S. of Eq. (1.43). HF intrinsically contains some higher orders of induction and this term is relatively cheap to include. As implied above, the reference for practical SAPT calculations tends to be from the HF level. As is typical, the simplest case is a closed-shell RHF reference. For open-shell systems, ROHF and UHF based SAPT0 have been derived.[42–44]

1.2.5 Applications of Potential Energy Surfaces

The effective calculation of E_{int} as a function of the degrees of freedom between the interacting molecules allows for the development of an approximate PES (or PEC). Though these potentials neglect potential many-body effects, a number of quantities of interest can be extracted from them. The PES, as defined above, is the potential associated with atomic movement, which can be used to solve the nuclear part of the Schrödinger equation. The results of this part of the equation provide information about the rotational and vibrational behavior of the system, which can be used to predict rovibrational spectra.[45,46] Other quantities that can be derived from the PES include improvements to ideal gas approximations, such as the virial coefficients for gas properties. The PES also provides the information necessary to calculate various cross sections of collisions. Such cross sections play a role in the calculation of transport properties, such as viscosity, which are based on the interactions within the system.[47] Additionally, scattering cross sections can be calculated to predict the results of various experimental methods. Characteristic points of the PES can be used to identify transition states of reactions. Furthermore, the lowest energy pathway connecting the transition state and related minima provides an approximation of the reaction pathway. The potential energy surface can also be used in simulations of condensed phase to predict properties such as radial distributions.[48]

1.3 Outline of Dissertation

The remaining chapters of this dissertation will be organized as follows. Each chapter will begin with an overview of the project and topics particular to it. Chapter 2 discusses the prerequisite *ab initio* calculations for the computation of selected thermophysical properties and second virial coefficients of the krypton gas. Next, Chapter 3 summarizes the investigation of the interactions between CO_2 and representative models of metal-organic frameworks. Lastly,

Chapter 4 describes recent work in expanding the capabilities of SAPT calculations. These chapters recompile and expand on the works submitted in partial fulfillment of this dissertation, which can be found in Appendices A-D.

Chapter 2

Theoretical Calculations for the Kr Gas

With the growth of computational chemistry, it has become possible to produce highly accurate PESs for many systems of interest. A particular category of systems that have attracted repeated attention is the collection of noble gas dimers.[45–47,49–52] With the accuracy of such potentials, theoretical calculations of thermophysical properties for these systems can have estimated uncertainties as good as, or smaller than, the most advanced experimental data.[47,51–57] To this end, a state-of-the-art PEC for the Kr dimer was developed, going beyond the CCSD(T) standard to acquire spectroscopic accuracy ($\leq$1 cm^{-1}) around the curve minimum. An analytic potential was fitted to the high-accuracy points and used in the calculation of several of the most important thermophysical properties of the Kr gas, specfically: viscosity, thermal conductivity, self-diffusion coefficient, thermal diffusion factor, and the second pressure, acoustic, and dielectric virial coefficients. As an additional requirement for the computation of the second dielectric virial coefficient, the interaction-induced isotropic pair polarizability was produced at a similar theoretical level as the PEC and fitted to an analytic form. This chapter summarizes the development of the new *ab initio* PEC[51] and pair polarizability[57] for the Kr dimer, originally described in Appendices A and B, respectively.

2.1 Potential Energy Curve of the Kr Dimer

2.1.1 Investigation of Accuracy at Near-Minimum

To develop the analytic potential for the PEC, the interaction energy for the krypton atom pair needed to be calculated at several interatomic distances. Furthermore, the level of theory

needed to acquire the desired spectroscopic accuracy around the minimum had to be determined. An initial investigation into the uncertainty of the calculated interaction energy was performed at a near-minimum interatomic distance, $R = 4.0$ Å. Throughout the chapter, the interaction energy is the counterpoise corrected value, defined in Eq. (1.25), and the basis sets are Dunning-type sets.[58]

The first consideration was how well the frozen-core (FC) CCSD(T) interaction energy was converged to the CBS limit. Frozen-core implies that only the valence electrons are allowed to correlate during the underlying CCSD calculation, while the core orbitals are held in their HF form. Table I in Appx. A presents the pertinent results. The top row of this table shows the convergence of the FC E_{int} with respect to the increasing cardinal number of the basis set, i.e. its size. The uncertainty for each basis set is assumed to be the difference between itself and the basis set one cardinal number lower, e.g. the uncertainty for the a5Z value is its difference with aQZ, $\pm$10.45 cm^{-1}. It is clear that this term is not completely converged, even in the large a5Z basis. The next row provides the results based on extrapolating from successive basis sets using the previously described X^{-3} scheme of Eq. (1.27). For the extrapolated values, the uncertainty is based on the difference between the extrapolated value and the value of the larger set used in the extrapolation. Based on the trend with the (aTZ, aQZ) result, the (aQZ, a5Z) result appears better converged but still has an uncertainty of $\pm$10.84 cm^{-1}. Next, the addition of midbond functions was tested. It is obvious that the aXZM results converge faster to CBS, but still have not reached saturation. Applying extrapolation alongside the midbond functions improves convergence, but the uncertainty of the (aQZM, a5ZM) result is still more than the desired 1 cm^{-1} threshold.

In an attempt to reach the target uncertainty, the F12 method of explicit correlation was applied to accelerate the basis set convergence.[59,60] As implied, this method introduces explicit electron correlation terms into the CC calculation. Two variants of this method, F12a and F12b,[61,62] were used in conjunction with the midbond functions, and extrapolation was again performed. The common implementation of CCSD(T)-F12 is a bit of a misnomer. The full CCSD(T)-F12 method is quite complicated, so it is common to perform the standard perturbative triple calculation (i.e. without explicit correlation) on top of a CCSD-F12 calculation. A

method of approximating the explicitly correlated perturbative triples is to scale the standard triples by the ratio of the corresponding MP2-F12 and MP2 values,[63,64]

$$E_{int}^{(T**)} = E_{int}^{(T)} \frac{E_{corr}^{MP2-F12}}{E_{corr}^{MP2}}. \tag{2.1}$$

These MP2 values are side products of the CCSD(T)-F12 calculations and require no additional expensive computations. The double star notation in Eq. (2.1) is to specify that the MP2 ratio for the dimer calculation is used to scale the dimer and monomer triples, as opposed to using the monomer ratios for the monomer triples. Applying explicit correlation and midbond functions shows a substantial increase in the rate of convergence, even before extrapolation. The addition of extrapolation shows that each of these methods appears to converge to approximately the same CBS value, around -133.5 cm^{-1}. Two of these methods, CCSD(T)-F12a/(aQZ, a5Z) and CCSD(T**)-F12b/(aQZ, a5Z), also have uncertainties lower than the desired threshold. CCSD(T**)-F12b/(aQZ, a5Z) has the lower uncertainty and was selected as the method for calculating the leading FC term of the interaction energy for all points.

With the selection of the leading term, the next consideration is determining what corrections to the leading term contribute more than 1 cm^{-1} at the near minimum separation. The results for the investigation of these corrections can be found in Table II of Appx. A. The first term inspected was the removal of the frozen-core approximation and allowing the core electrons to be excited. This correction is referred to as the core-core and core-valence correlation correction, but will be termed the all-electron (AE-FC) correction here. Due to the considerable number of core electrons in the Kr atom, it is conceivable that this correction could be of some importance. The AE-FC corrections were calculated in the aug-cc-pCVXZ (aCXZ)[65] and aug-cc-pwCVXZ (awCXZ)[66] basis sets, which are Dunning-type sets designed for the appropriate treatment of core electrons. The key observation is that even the least saturated basis sets provide corrections greater than 1 cm^{-1}, confirming the importance of this correction for an accurate interaction energy. The addition of midbonds to this correction had only moderate benefit over an increase in cardinal number, and the highest available set with midbonds was limited to aQZM. The (aCQZ, aC5Z) and (awCQZ, awC5Z) extrapolated values are similarly

converged and have similar uncertainties. The (awCQZ, awC5Z) result was picked as the level for the AE-FC correction because it is slightly better converged.

The second correction of interest was the improvement beyond the CCSD(T) level to higher levels of CC, specifically to CCSDT(Q). At first glance, the (Q)-(T) correction appears insignificant within the required uncertainty. Closer inspection reveals that the size of this correction is due to a cancellation between its components, the T-(T) and (Q)-T corrections. Individually, these corrections are large enough to warrant inclusion in the interaction energy, but they are of similar magnitude and opposite direction. Based on the individual size of the component corrections (as well is its increased importance at short intermolecular distances seen in the next section), it was decided to include the correction in the aQZ set.

The last considered correction to the interaction energy is the correction due to relativistic effects. Electrons in the core of the Kr atom can reach speeds near enough to the speed of light that relativity can have considerable effect. The relativistic correction was investigated at the second-order Douglas-Kroll-Hess (DKH) level.[67,68] The basis sets used for this correction were the aXZ sets, but the contraction of the Gaussian orbitals was removed to provide a significantly larger basis space. The relativistic correction is of similar magnitude to the AE-FC correction, which necessitates its inclusion in the interaction energy. The (decontracted aQZ, decontracted a5Z) result appears sufficiently converged and was selected as the level of the correction to include in the final points.

2.1.2 Fitting Analytic Potential

The total Kr-Kr interaction energy for this work is defined as

$$E_{int} = E_{int}^{CCSD(T)/FC} + \Delta E_{int}^{CCSD(T)/AE-FC} + \Delta E_{int}^{(Q)-(T)} + \Delta E_{int}^{rel}, \tag{2.2}$$

where the leading FC term and the corrections were calculated at the previously determined levels. The resulting interaction energy was calculated for selected interatomic distances along the PEC, ranging from 2.60 Å to 12.00 Å and with an increased density of points around the

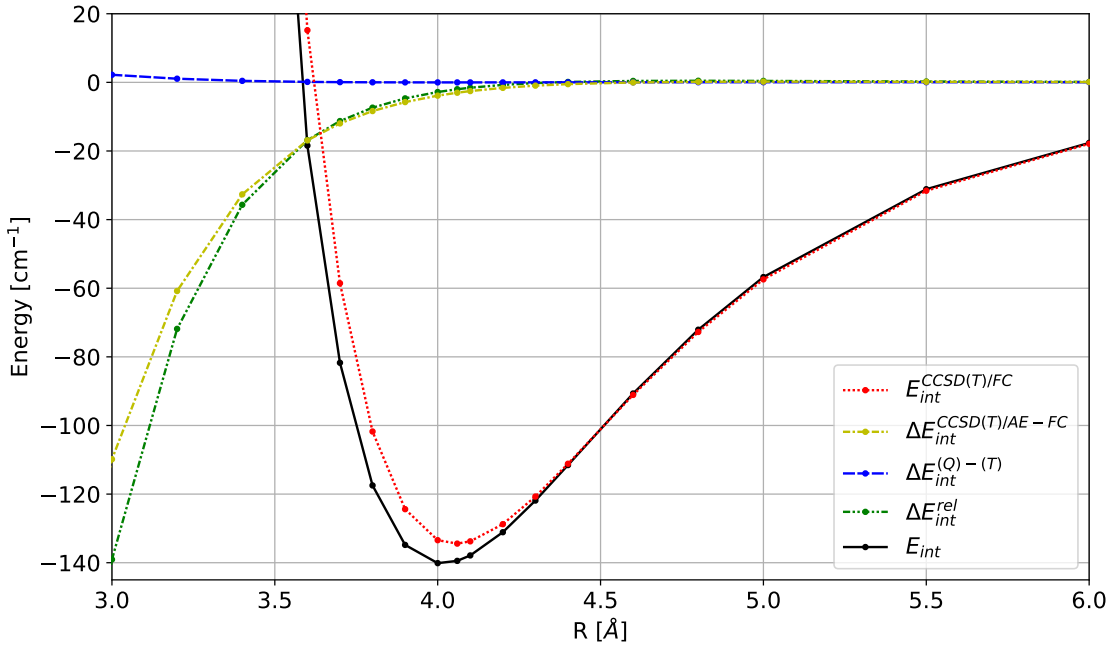


Figure 2.1: Contributions to the *ab initio* krypton dimer interaction energy and the total interaction energy (cm^{-1}) as functions of interatomic distance R (Å).

expected minimum. The selected distances can be seen in Table III of Appx. A, along with the numerical values for the terms of the interaction energy at each point.

The role of the leading term and corrections at different interatomic distances can be observed graphically in Fig. 2.1. At long and medium ranges, the FC term is shown to be a good approximation to the complete E_{int}, while the corrections all tend to zero for distances greater than 4.5 Å. In the well region, the $\Delta E_{int}^{CCSD(T)/AE-FC}$ and ΔE_{int}^{rel} corrections are equally important. These corrections deepen the well compared to the FC value and shift the apparent minimum to a shorter separation. Past the well and into the repulsive wall, ΔE_{int}^{rel} becomes more negative than $\Delta E_{int}^{CCSD(T)/AE-FC}$, while both corrections counteract some of the repulsion from $E_{int}^{CCSD(T)/FC}$. $\Delta E_{int}^{(Q)-(T)}$ has minimal effect throughout much of the PEC, but does play a small role in the wall.

To approximate the potential at arbitrary interatomic distances, an analytic formula was fitted to the *ab initio* data points. The form of the function is

$$V(R) = (A + BR + CR^{-1})e^{-\alpha R} - \sum_{n=3}^{4} f_{2n}(\beta R)\frac{C_{2n}}{R^{2n}} \tag{2.3}$$

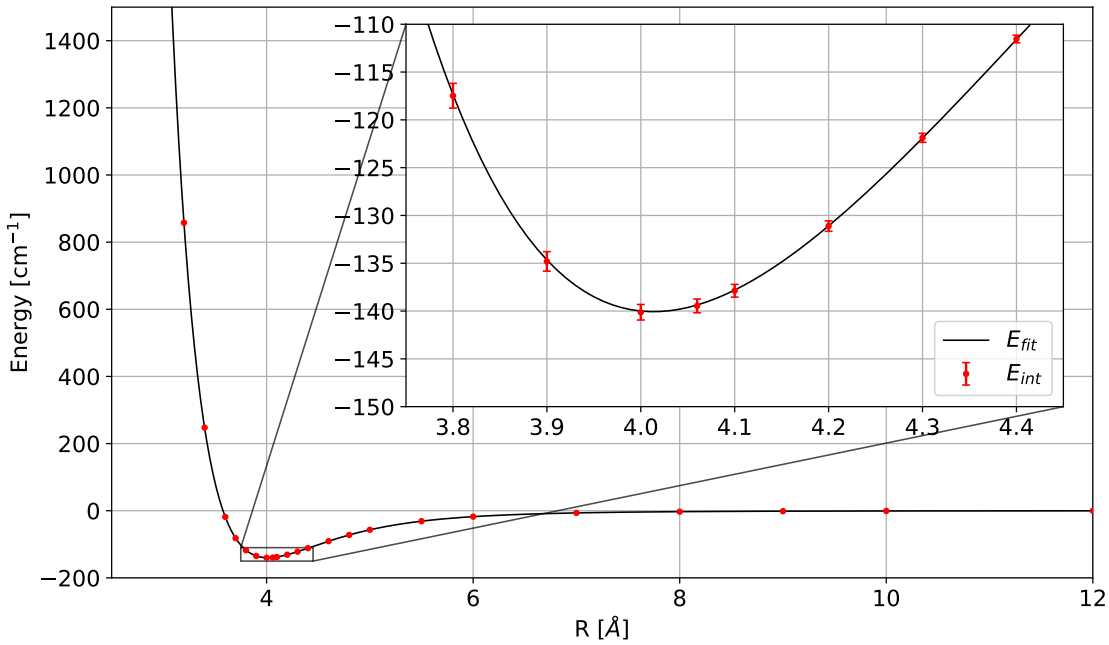


Figure 2.2: The *ab initio* values of E_{int} and the fitted Kr dimer potential (cm^{-1}) as functions of the interatomic distance R (Å).

where A,B,C,α,β,C_6, and C_8 are fitting parameters. The functions f_{2n} are the Tang-Toennies damping functions,[69]

$$f_{2n}(x) = 1 - e^{-x} \sum_{k=0}^{2n} \frac{x^k}{k!}, \tag{2.4}$$

which serve to damp the dispersion interactions of the C_6 and C_8 terms in Eq. (2.3). These two parameters were fitted first to the long range points ($R \geq 8.0$Å), then held constant during the fitting of the other parameters. The fitting was accomplished by the method of least-squares, with each data point weighted by scaling the squared difference of the fitted value by the inverse of the squared uncertainty associated with the *ab initio* result at that point. This weighting scheme ensures that the least uncertain points are prioritized over the others, e.g. points around the minimum versus points high on the repulsive wall. The optimized values of the parameters are provided in Table IV of Appx. A, while the interaction energy values of the fitted potential at the selected data points are in Table III of the appendix.

Figure 2.2 shows the calculated values of E_{int}, with their uncertainties, along with the fitted potential. The number of parameters in the parentheses in Eq. (2.3) was selected to ensure that the fitted potential passed within the range $[E - U(E), E + U(E)]$ for all points,

while also keeping the number of fitted parameters to a minimum. The zoomed-in region of Fig. 2.2 shows this behavior clearly around the potential well. This potential predicts a minimum distance of 4.01 Å with an interaction energy value of -140.06 cm^{-1}, which agrees well with previously reported values[70] and values from detailed works more recent than that presented here.[52]

The reasonability of the fitted potential is assured from the shortest *ab initio* data point to infinite separation, but not so for particularly short separations. Specifically, the wall region of the potential discussed here actually flips over and proceeds to negative infinity starting at some short distance. To overcome this unphysical behavior, the values of the fitted potential are replaced with a well-behaved function starting at a distance shorter than 2.6 Å (the shortest fitting point) but before the function becomes ill-behaved. The form of this short range function is simple,

$$V_{sh}(R) = (A_{sh}/R)e^{-\alpha_{sh}R+\beta_{sh}R^2} \tag{2.5}$$

with the fitting parameters A_{sh}, α_{sh}, and β_{sh} fitted to ensure physical behavior and that the potential and its first derivative are continuous at the splicing point ($R = 1.8$ Å). The values of these parameters are listed alongside the other potential parameters. In addition, functions of the same form as Eq. (2.3) (and with the same short range form as Eq. (2.5)) were fitted to $E - U(E)$ and $E + U(E)$ for all points to facilitate the determination of uncertainty for the properties computed from the new dimer potential. The parameter values of these limit functions can be found in Table IV of Appx. B.

2.2 Interaction-Induced Isotropic Pair Polarizability of the Kr Dimer

2.2.1 Investigation of Accuracy at Near-Minimum

The pair potential discussed above is sufficient for the calculation of the second pressure and acoustic virial coefficients. For the second dielectric virial coefficient, the interaction-induced isotropic pair polarizability $\Delta\alpha_{ave}$ is also required. This quantity is the average change in polarizability from the lone monomers to the dimer due to their interaction, and is defined in

the following equations

$$\Delta\alpha_{\parallel}(R) = \alpha_{\parallel}^{Dimer}(R) - 2\alpha_{\parallel}^{Monomer}(R) \tag{2.6}$$

$$\Delta\alpha_{\perp}(R) = \alpha_{\perp}^{Dimer}(R) - 2\alpha_{\perp}^{Monomer}(R) \tag{2.7}$$

$$\Delta\alpha_{ave}(R) = \frac{1}{3}[\Delta\alpha_{\parallel}(R) + 2\Delta\alpha_{\perp}(R)] \tag{2.8}$$

where α is the polarizability of either the dimer or monomer (superscripts), either parallel or perpendicular to the axis of interaction (subscripts). A related value is the anisotropy $\Delta\alpha_{aniso}$, defined as

$$\Delta\alpha_{aniso}(R) = [\Delta\alpha_{\parallel}(R) - \Delta\alpha_{\perp}(R)]. \tag{2.9}$$

Though not necessary for the property computations, this value is a side product of the calculation of $\Delta\alpha_{ave}$.

As with the pair potential, the pair polarizability was first investigated at the same near-minimum distance (which is effectively the minimum determined from the potential). The CCSD(T)/FC value of $\Delta\alpha_{ave}(R)$ was first checked for convergence. Then, the same corrections beyond the FC term (AE-FC, relativistic, and higher order CC) were considered for their importance. Unlike for the interaction energy, the desired level of accuracy was not predetermined and the corrections were evaluated by their size relative to the FC term. The results for all terms inspected can be found in Table I of Appx. B. The parallel and perpendicular components of the CCSD(T)/FC term seem to converge at similar rates, and have values with relatively small magnitudes. The aCXZ and awCXZ basis sets were checked for this term, but provided no improvement over the more appropriate aXZ sets. Furthermore, the aXZ sets extended one cardinal number further than the other sets. Midbond functions were determined to be unnecessary given the effectiveness of the conventional basis sets. CCSD(T)/FC/a6Z was determined to be a suitable level of theory for the leading term.

The AE-FC correction was checked for the same basis sets as the FC term, where the awCXZ set was found to converge the fastest. The size of this correction was significant enough to warrant including it, but the calculation would be limited to the awCQZ basis since

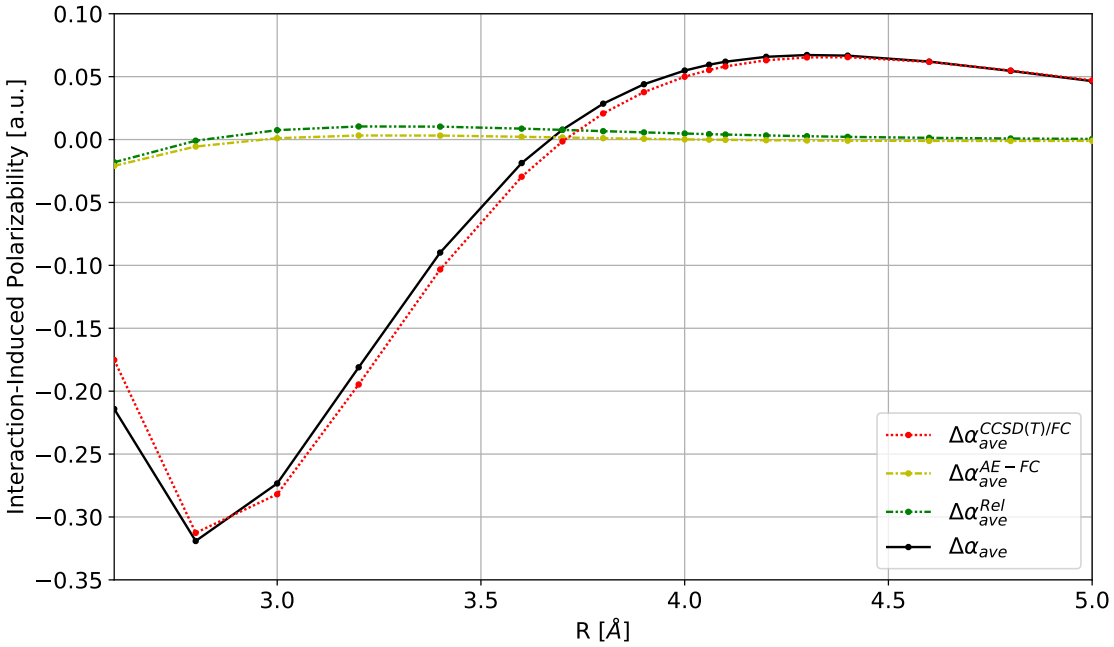


Figure 2.3: Contributions to the interaction-induced isotropic polarizability $\Delta\alpha_{ave}$ for the krypton dimer and the total $\Delta\alpha_{ave}$ (a_0^3) as functions of the interatomic distance R (Å).

the correction was already well enough converged at that point. The correction for higher orders of CC, specifically CCSDT, was found small enough to be neglected ($\Delta\alpha_{\parallel}^{T-(T)}$/aDZ = 0.0023 a_0^3 and $\Delta\alpha_{\perp}^{T-(T)}$/aDZ = -0.0003 a_0^3). The relativistic correction was calculated analytically using the exact two-component method (X2C),[71] as opposed to DKH used in the pair potential. The available implementation of this method necessitated validating due to its experimental nature. This process is described in Appx. B and the method was found to be sound. The decontracted basis sets were again found to be the most suitable for this correction, and the decontracted aQZ set was selected for use at all data points.

2.2.2 Fitting an Analytic Function

With the necessary components and basis sets selected, the interaction-induced pair polarizability was calculated for the same interatomic distances as used for the potential. The total $\Delta\alpha_{ave}$ is defined as

$$\Delta\alpha_{ave} = \Delta\alpha_{ave}^{CCSD(T)/FC} + \Delta\alpha_{ave}^{AE-FC} + \Delta\alpha_{ave}^{rel}. \tag{2.10}$$

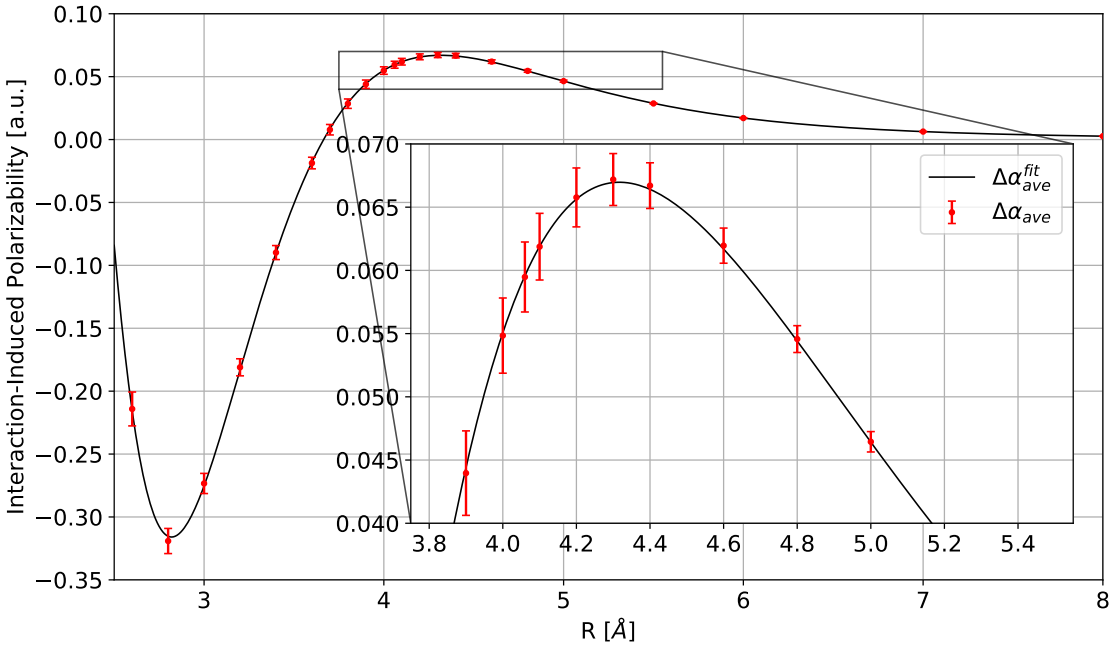


Figure 2.4: The *ab initio* values of $\Delta\alpha_{ave}$ and the fitted Kr dimer isotropic interaction-induced pair polarizability (a_0^3) as functions the interatomic distance R (Å).

The numerical results for all points can be found in Table II of Appx. B, and the various contributing terms are shown in Fig 2.3. Through most of the displayed range, the FC term remains a good approximation. Only at the shortest ranges do the corrections lead to any considerable difference in the total and FC values. Both the AE-FC and relativistic corrections are small and positive until these short ranges, where they become negative and about equal in magnitude. At $R \geq 3.0$ Å, the relativistic correction contributes slightly more than the AE-FC correction.

The fitting scheme used for the potential was also used for the pair polarizability, utilizing least-squares fitting and weighting the points as the inverse square of their uncertainty. The form of the function used to fit the isotropic pair polarizability is

$$\begin{aligned}\Delta\alpha_{ave}(R) =& (A^{(ave)}R^{-1} + B^{(ave)} + C^{(ave)}R + D^{(ave)}R^2)e^{-\alpha^{(ave)}R} \\ &+ \sum_{n=6,8} f_n(\beta^{(ave)}R)\frac{C_n^{(ave)}}{R^n}\end{aligned} \tag{2.11}$$

where $A^{(ave)}$, $B^{(ave)}$, $C^{(ave)}$, $D^{(ave)}$, $\alpha^{(ave)}$, $\beta^{(ave)}$, $C_6^{(ave)}$, and $C_8^{(ave)}$ are the fitting parameters and f_n are the Tang-Toennies damping functions. Again, the $C_6^{(ave)}$ and $C_8^{(ave)}$ terms were fitted to only the long range points and held constant as the other parameters were fitted to all data points. The upper- and lower-limit functions were fitted to $\Delta\alpha_{ave} + U(\Delta\alpha_{ave})$ and $\Delta\alpha_{ave} - U(\Delta\alpha_{ave})$, respectively, using the same functional form and procedure. The parameters for all three functions can be found in Table III of Appx B. As illustrated for most points in Fig. 2.4, the fitted function passes within the uncertainty range of all *ab initio* datapoints.

2.3 Evaluation of Computed Properties

The new potential was first used to calculate the vibrational transition frequencies of $^{84}Kr_2$ isotopomer, the results for which are found alongside previous literature values in Table V of Appx. A. The transitions are in good agreement with the empirical[70] and experimental results,[72] and are likely more accurate than the previous *ab initio* results.[73] Most importantly, the highly accurate experimental values for the $^{86}Kr_2$ isotopomer from LaRocque et al[74] are within the uncertainty of the theoretical results based on the new potential.

With the analytic potential fitted to the high accuracy *ab initio* results, the new values for the viscosity, thermal conductivity, self-diffusion coefficient, and thermal diffusion factor could be computed. Furthermore, the second virial coefficients can also be computed, with the dielectic one requiring the pair polarizability function. The method of computing these properties is described in Appxs. A and B, and will not be addressed in any detail here other than to say that they are calculated classically and that the virial coefficients utilize quantum corrections. The properties were calculated for temperatures ranging from 116 to 5000 K, and the results of these computations are found in Table VI of Appx. A and Table V of Appx. B. The accuracy of the new viscosity and thermal conductivity values will now be evaluated, while the uncertainties associated with the comparisons for the self-diffusion coefficient and thermal diffusion factor provide little insight.

The newly calculated viscosity values are initially compared with reported literature values in Fig. 2 of Appx. A. This initial comparison shows moderate agreement with the new values, with most literature data points having relative deviations with magnitudes greater than 0.5%.

It is useful to note that the experimental viscosity values are determined relative to some other calibrated value. The instruments used to gather the literature values were calibrated with older data with uncertainties larger than those available more recently, specifically those produced via *ab initio* calculations (other than this work). To improve the literature data, the points were worked backwards to the relative ratio of the viscosity of Kr with the reported reference gas and then the updated value was determined using a corresponding *ab initio* result for the reference gas. The comparison of the newly calculated results and the updated literature results is displayed in Appx. A, Fig. 3. This figure also contains previously calculated theoretical results and the uncertainty values of the newly calculated viscosity.

The re-evaluation described above moves most of the literature points within $\pm 0.5\%$ of the computed values, with several points falling within the uncertainty of the new results. The agreement of the updated literature results is also much better with the new viscosity results over the previous theoretical ones, with the empirical potential[70] leading to an overestimation across the range and the *ab initio* potential[73] underestimating most of the experimental values. Similar trends are seen for the thermal conductivity in Fig. 4 from Appx. A, where there are less experimental data points for comparison. While it is harder to differentiate between the thermal conductivity results based on the newer and older *ab initio* potentials, the new values have a more consistent agreement with the experimental literature.

For the second virial coefficients, the amount of literature data for comparison was limited and consisted mainly of previous theoretical results. The pressure and acoustic coefficients, seen in Figs. 3 and 4 in Appx. B, chiefly highlight the considerable agreement between the values of these works and those from the more recent potential of Jäger et al.[52] The values corresponding to their newer potential uniformly fall within the upper uncertainty bound of the coefficients produced in this work. For the second dielectric virial coefficient, a lack of experimental data limits the characterization of the newly computed results.

The good agreement with experimental results leads to the conclusion that this new potential is more accurate then those previously calculated, and that the properties computed from it are suitable for use as reference values. The recent development of a more accurate potential[52] did produce somewhat smaller uncertainties, but was not useful to draw conclusions about the

relative accuracy of the resulting properties and those presented here due to relatively high uncertainties in the experimental values.

Chapter 3

Investigation of Metal-Organic Framework Adsorption of CO_2

Metal-organic frameworks (MOFs) are composed of metal atoms or clusters coordinated to organic linker molecules that produce a crystalline structure with high porosity. This class of compounds has garnered interest for a variety of chemical purposes, including selective gas separation.[75,76] The structure and properties of an MOF are tunable via altering the metal(s) or linker(s), providing a pathway for optimizing interactions with a particular adsorbate.[77,78] Carbon dioxide is a particularly relevant adsorbate of interest.[79] The large scale of these MOFs makes them difficult cases for the application of expensive *ab initio* methods. For this reason, model systems are used in an attempt to replicate the interactions of the bulk system at a more tractable size.[77,80,81] High-accuracy results for such models can then be used to benchmark more approximate methods that can be applied to larger sized systems. Additionally, the interactions of these models can be decomposed (with methods like SAPT) to provide increased insight into the rational design of MOFs for specific purposes.[81–83] This chapter summarizes the work in Appendix C, wherein the interaction of CO_2 with MOF model systems is investigated with a number of *ab initio* methods.

3.1 Model Structures of Metal-Organic Frameworks

Each of the mimic systems in this work contains one of three metals species: Mn^{2+}, Co^{2+}, or Zn^{2+}. These metals were chosen as representatives of high spin open-shell, low spin open-shell, and closed-shell systems, respectively. The selection of these metals allows for testing the relative importance of a multireference treatment for MOF mimic systems. The mimics are

constructed from one or two atoms of one metal species and a varying number of organic ligands to represent the MOF linkers. Two of the models contain one metal and one ligand, either benzenedicarboxylate (BDC, also called terephthalate) or furandicarboxylate (FDC). Another model contains a single metal with two formate ligands, and the last contains two metals and four formates. These last two models are referred to as the diformate and paddlewheel models, respectively. The approximate structure of each mimic, with an arbitrary metal, interacting with CO_2 can be seen in Fig. 1 of Appx. C. The formate ligands all have a -1 charge and BDC and FDC have -2 charges, resulting in each of the model systems being electronically neutral.

The BDC and FDC systems were optimized at the B3LYP/aDZ[84,85] level, the paddlewheel systems at the MP2/aDZ level, and the diformate ones at MP2/aTZ. The systems containing Mn^{2+} and Zn^{2+} were optimized in a generally unconstrained fashion. In the diformate models, the mimic systems were required to have all atoms coplanar, but all other degrees of freedom were optimized freely. For the paddlewheel models, two minimum structures were obtained: one with the CO_2 aligned with the axis of the metal centers, and another where the CO_2 appears to be interacting with the formate ligands as well as one of the metal centers. A comparison of these geometries can be seen in Fig. 2 of Appx. C. While this second structure is indicative of real effects in MOFs (i.e. adsorbate interaction with the organic linkers), it reduces the symmetry of the dimer from C_{4v} to C_s. The axially aligned geometry was selected to take advantage of the higher symmetry and the computational advantage it grants. The optimized structures of the other systems were each nearly symmetric and were symmetrized for the same computational benefits. The BDC and diformate models have C_{2v} symmetry, while the FDC models have C_s symmetry. Additionally, a second optimization of each dimer was performed with the monomers held in their individually optimized geometries. This second dimer geometry was used exclusively for the calculation of simple one-dimensional PEC scans, where only the intermolecular separation was altered.

The systems containing Co^{2+} were treated more cautiously, due to their potentially multireference nature. For each model system, the corresponding Mn^{2+} structure was used as a starting point for CASSCF/aDZ scans of the most important parameters of each model. The metal-CO_2 distance ($R_{M^{2+}-CO_2}$) was scanned for all mimic dimers, along with the following

parameters that were checked for both the isolated mimics and the complexes with CO_2. For the paddlewheel systems, the important parameters are the distances from the model center of mass for the metals ($R_{X-M^{2+}}$) and the formates ($R_{X-Ligand}$). For the diformate, BDC, and FDC systems, the ligand-metal distances ($R_{Ligand-M^{2+}}$) were scanned. Additionally, the angle between the ligand, metal, and CO_2 ($\Theta_{Ligand-M^{2+}-CO_2}$) was scanned for FDC since it was not fixed by symmetry. In the BDC and FDC monomers, it was also necessary to scan the Θ_{O-C-O} angle of the carboxylate coordinated to the metal. This caution proved unnecessary, as latter results will show that the Co^{2+} systems can be well represented by a single reference calculation. The Co^{2+} containing structures were optimized in the same way as the other metals, and the cautious results were deemed similar enough to keep. Table 1 in Appx C. provides the values for the important parameters of each system. For all systems, Zn^{2+} is bound more closely to the ligands, followed first by Co^{2+} and then Mn^{2+}. The binding of CO_2 is less universal, being dependent on the mimic system and metal species.

3.2 Single-Reference and Multireference Methods

The first consideration after geometry optimization is determining an appropriate benchmark level of theory with which to compare various approximate methods. Each system was investigated with common single-reference (RHF, MP2, and RCCSD(T)) and multireference (CASSCF, CASPT2, NEVPT2, MRCISD) methods in the aDZ basis. The active space in all multireference calculations included the 3d and 4s orbitals of the metals. For the three smallest mimic systems, the importance of active space size was tested for the multireference methods and a second active space added the metal 4p orbitals. For the paddlewheel systems, the active space is ($2n$, 12), where n is the number of d-electrons on one metal center. The other systems are (n, 6) in the smaller space and (n, 9) in the larger one.

Regarding single-reference or multireference treatment and active space sizes, the natural orbital occupation numbers (NOONs) for each model system were calculated at the CASSCF/aDZ level and are found in Tables 6-9 of Appx. C. These NOONs show how the occupations of the active orbitals differ from the restricted occupation numbers (2, 1, or 0) and provide an indication of the multireference character of the system. The results here show that the systems

are not considerably multireference, with even the largest deviations being still quite close to the restricted expectations. The Mn^{2+} results are to be expected, with no change in the smaller active space and a small change in the larger one. Co^{2+} and Zn^{2+} deviate similar to each other in both active spaces, with Co^{2+} generally displaying more significant changes. Based on these results, the systems should be sufficiently single-reference for RCCSD(T) to be considered the benchmark (RCCSD for the Zn^{2+}-BDC system due to a suspect (T) value). The multireference methods will also be compared to these results for evaluation.

The results of these wavefunction methods are found in Tables 2-5 in Appx. C. The features of these tables that stand out most immediately are the MRCISD values in the E_{int}^{CP} columns. These values are completely incongruous with the other values and arise from the size inconsistency of the method. These values prompted the use of Eq. (1.26) to produce more reasonable MRCI results. The E_{int}^{CP*} values are defined according to Eq. (1.26), but with the addition of ghost atoms near the infinitely separated monomers to provide the desired counterpoise correction. This reference is visualized in Fig. 3 of Appx. C. Quick inspection of these tables show that E_{int}^{CP} and E_{int}^{CP*} are usually in near perfect agreement for all methods besides MRCISD, where the E_{int}^{CP*} values are more reasonable. In the cases where the two references differ considerably, the E_{int}^{CP*} is deemed the more trustworthy result.

Comparing with the CC results, CASPT2 provides satisfactory results and outperforms NEVPT2 for the diformate systems, but did not provide meaningful results for the other systems and was omitted. Additionally, NEVPT2 produced wildly inconsistent results for many of the systems and is not very useful in a predictive capacity. Of the multireference methods, only MRCISD proved to be a reliably predictive method (under the E_{int}^{CP*} definition) and was insensitive to the active space selection, as well. With that said, MP2 is consistently in similar or better agreement with the RCCSD(T) results, and is considerably cheaper to calculate than MRCISD.

Th relative strength of interaction for the different mimics is reasonable. The BDC and FDC models have much stronger interactions than the diformate and paddlewheel due to their obvious dipole moments. The character of these interactions will be discussed more in the next section. In terms of metal preference for each model, the interaction energies for each of the

models follow one of two trends. The diformate and paddlewheel systems show a trend of $Mn^{2+} > Zn^{2+} > Co^{2+}$, implying CO_2 binds most strongly with the Mn^{2+} version and least strongly with the Co^{2+} version. The FDC and BDC systems have a trend of $Zn^{2+} > Mn^{2+} > Co^{2+}$.

3.3 SAPT Decomposition

SAPT analysis was applied to each model to provide insight into the way each mimic interacts with CO_2. The SAPT0 level of theory was applied to all mimic-metal combinations, in both the aDZ and aTZ basis to test for basis set completeness. The Zn^{2+} results utilized the RHF-based SAPT,[38] while the Mn^{2+} and Co^{2+} results were obtained from ROHF-based SAPT.[42] The Mn^{2+} calculations were also performed with UHF-based SAPT for comparison.[43,44] SAPT2+3/aDZ[41] results were calculated for the Zn^{2+} systems to observe the changes to the interaction energy at higher levels of SAPT. For the Co^{2+}-FDC model, only the SAPT0/aTZ was obtained due to convergence issues in the aDZ set. For similar reasons, neither result was obtained for the Co^{2+}-BDC model.

The results for the Zn^{2+} containing systems are found in Table 11 of Appx C. Comparison of the SAPT0/aDZ and SAPT2+3/aDZ results shows that the most consistent change between these levels of theory is in the exchange term. Exchange increases significantly for each model while the induction and dispersion terms decrease more slightly. Electrostatics either increases slightly or remains roughly the same. The net result is that the total interaction energy decreases with the increase in the SAPT level. For comparison of UHF- and ROHF-based SAPT, Tables 12 and 13 of Appx. C present the Mn^{2+} results for these two methods. The differences between these two sets of data are minimal and mostly isolated to the dispersion terms, which become more negative for the UHF-based results. Lastly, Table 14 of Appx. C, containing the available Co^{2+} results, is introduced so that the role of basis set size can be discussed. As with the UHF vs. ROHF comparison, the SAPT0/aDZ and SAPT0/aTZ results differ only a small amount for all cases. The changes in the total energies are almost entirely due to the changes in the dispersion terms. This occurrence is very reasonable as dispersion is the only correlated term

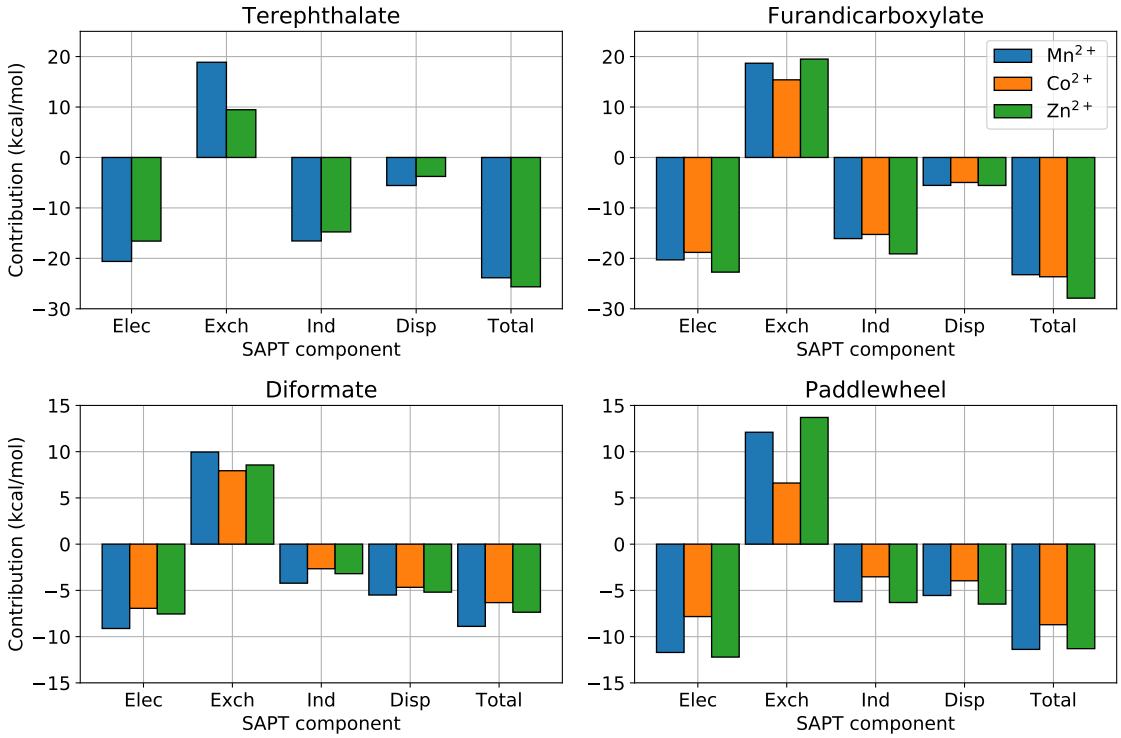


Figure 3.1: SAPT0/aTZ interaction decomposition results (kcal/mol) for the various model systems with each metal. The Zn^{2+} results come from RHF-based SAPT, while the Mn^{2+} and Co^{2+} results are ROHF-based SAPT.

and therefore will be slower to converge with the basis set, as opposed to the other terms which are all derived from a HF level treatment of the systems.

The most complete set of SAPT data (the SAPT0/aTZ results) is gathered in Figure 3.1 for comparison between the metals and models. The SAPT results qualitatively show the correct trend of interaction strength between the different models, with BDC being the most strongly interacting followed closely by FDC and then the paddlewheel and diformate. Quantitatively, the SAPT0 results persistently overestimate the interaction strength and produce more negative interaction energies than the CC results. SAPT gets the metal trend of the diformate and paddlewheel models correct, though for diformate it produces a much smaller difference between the Mn^{2+} and Zn^{2+} interactions than RCCSD(T). While lacking a Co^{2+} result for BDC, SAPT does get the order for the Mn^{2+} and Zn^{2+} interactions correct. The FDC model is the only case where SAPT shows a different metal trend than RCCSD(T), showing CO_2 interacting more strongly with Co^{2+} than Mn^{2+}. This discrepancy is of little consequence since the CC results for these two systems are quite close to begin with.

Comparing the SAPT components, it is obvious that the BDC and FDC are strongly induction bound. Again, this result is logical due to the dipole moments of these systems, the results for which can be found in Table 15 of Appx. C. The diformate and paddlewheel models have a more mixed character. This is especially true of the paddlewheel, where the induction and dispersion contributions are almost identical for each metal. For diformate, dispersion is a greater contribution than induction. Overall, dispersion is required for an adequate description of the interaction of each model system with CO_2. Regarding metal selection in each model, the components of the Co^{2+} systems generally have smaller magnitude than the other two metals, resulting in a net lower total interaction. The exception to this is the FDC models, where the induction and dispersion terms of the Co^{2+} model are close in value to the Mn^{2+} and lead to the change in metal preference mentioned above.

3.4 Density Functional Methods

The last set of methods tested were a selection of common density functionals. These included the previously mentioned B3LYP functional, along with the BLYP,[86] PBE,[87] PBE0,[88,89] M05-2X,[90] and M06-2X[91] functionals. PBE and BLYP are generalized gradient approximations (GGAs) that utilize the electron density and its gradient as parameters for the calculation of the KS energy. B3LYP and PBE0 are hybrid-GGAs and use a fraction of the SCF exchange energy to improve on the previous methods. The M05-2X and M06-2X functionals are referred to as interaction-optimized functionals and are heavily parameterized against various datasets. Given the importance of dispersion for these systems pointed out by the SAPT results, Grimme's empirical dispersion corrections for these functionals were checked with the original and updated parameters. These functionals were paired with the aDZ, aTZ, and aQZ basis sets, though basis set completeness proved to have little effect on these methods. The aTZ results are included only to illustrate this point and the aQZ results have been omitted entirely.

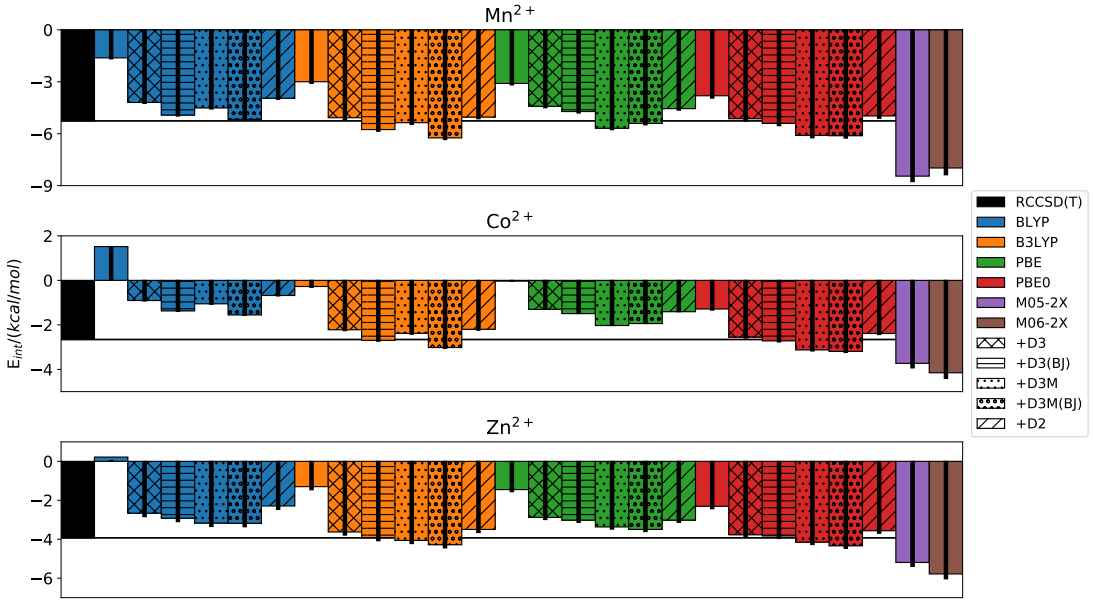


Figure 3.2: Comparison with benchmark of the DFT/aDZ E_{int}^{CP} (kcal/mol) for the diformate model-CO_2 systems. The bar colors indicate the method, while the hatchings indicate the type of empirical dispersion added. The lines bisecting the bars correspond to the result of that method in the aTZ basis set.

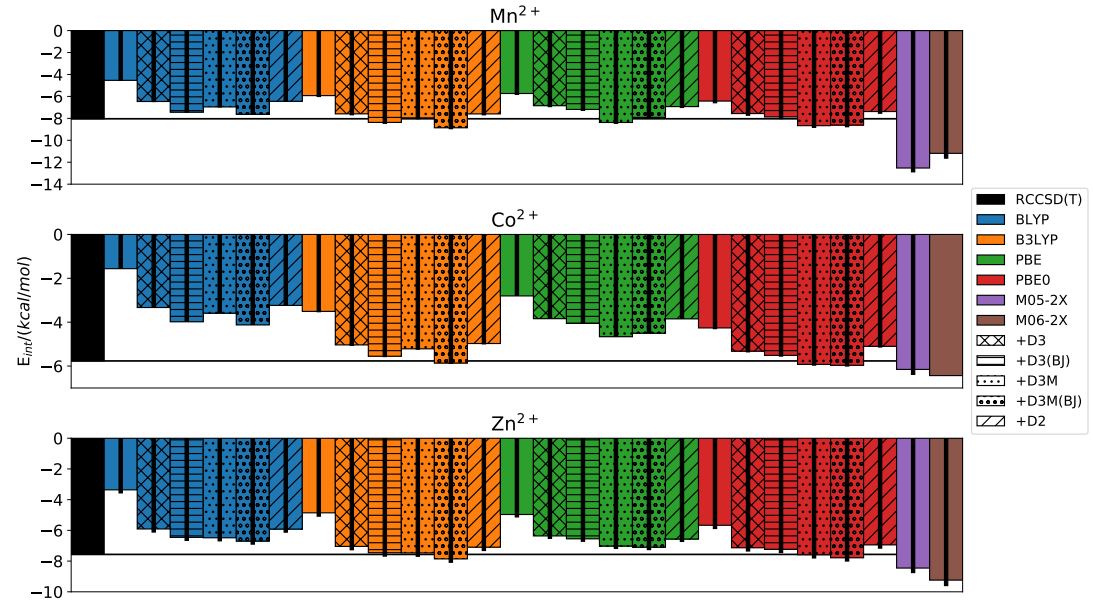


Figure 3.3: Comparison with benchmark of the DFT/aDZ E_{int}^{CP} (kcal/mol) for the paddlewheel model-CO_2 systems. The bar colors indicate the method, while the hatchings indicate the type of empirical dispersion added. The lines bisecting the bars correspond to the result of that method in the aTZ basis set.

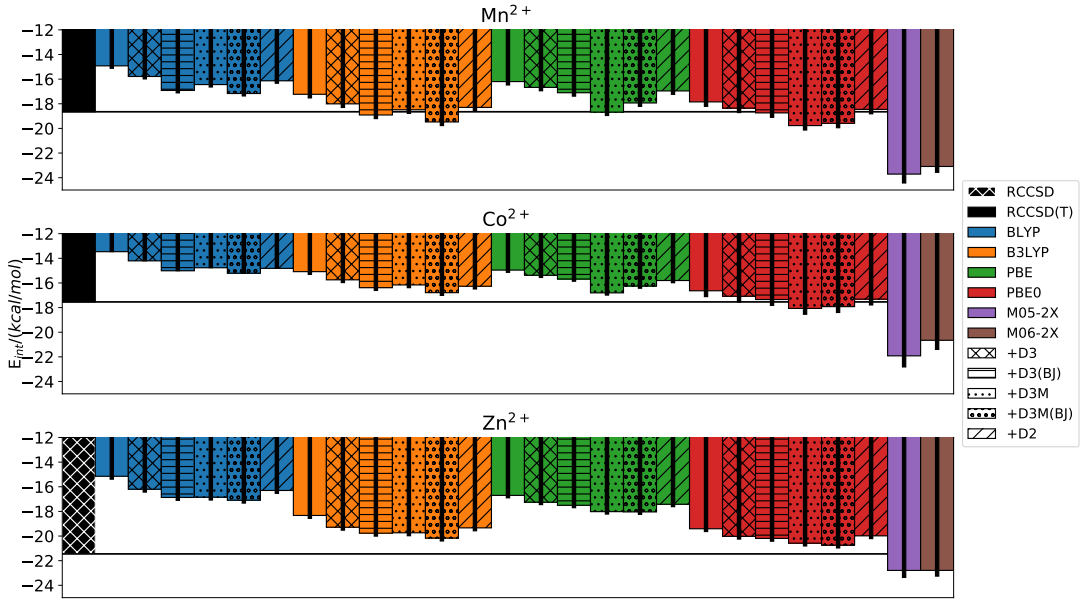


Figure 3.4: Comparison with benchmark of the DFT/aDZ E_{int}^{CP} (kcal/mol) for the BDC model-CO_2 systems. The bar colors indicate the method, while the hatchings indicate the type of empirical dispersion added. The lines bisecting the bars correspond to the result of that method in the aTZ basis set.

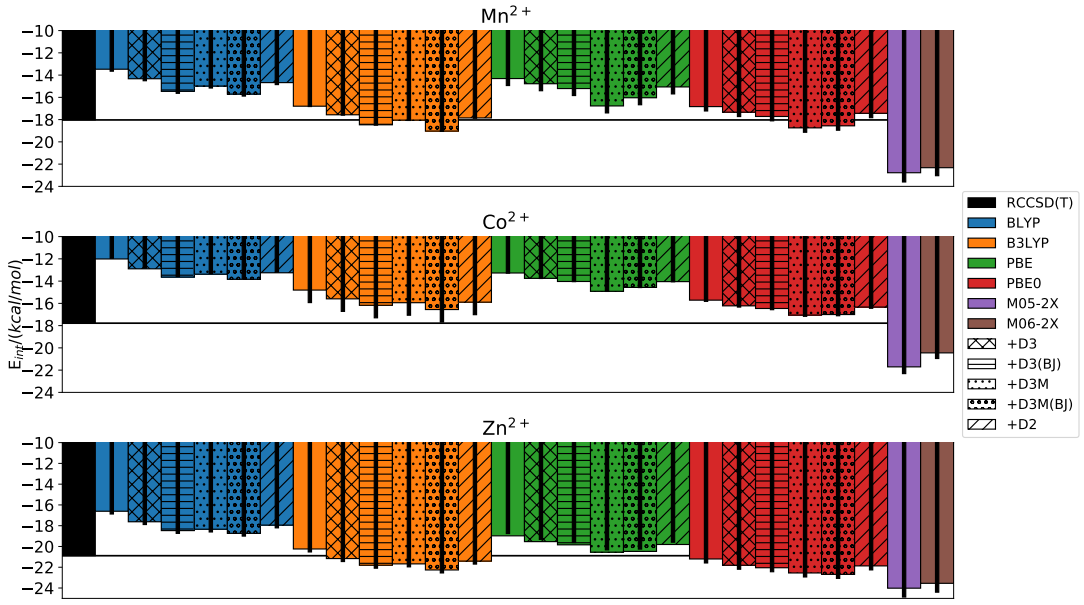


Figure 3.5: Comparison with benchmark of the DFT/aDZ E_{int}^{CP} (kcal/mol) for the FDC model-CO_2 systems. The bar colors indicate the method, while the hatchings indicate the type of empirical dispersion added. The lines bisecting the bars correspond to the result of that method in the aTZ basis set.

Figures 3.2-3.5 compare the DFT results, with and without the empirical dispersion corrections, to the RCCSD(T) benchmarks for each set of model systems. The black lines bisecting the DFT bars indicate the value of that method in the aTZ basis set. As mentioned above, these results differ minimally from the aDZ results. The only notable exception is for B3LYP and the Co^{2+}-FDC model, where the aTZ results improve over the aDZ results by about 2 kcal/mol. Another general trend is the unsurprising improvement seen with the addition of the dispersion corrections. The effects of these corrections correlate well with the SAPT analysis, with the diformate and paddlewheel systems being most dependent on dispersion to produce results near RCCSD(T). Though inclusion of dispersion is important for all of the mimics, the specific version of the dispersion correction is less important. The D3 corrections are better than the older D2, but no particular parameterization of D3 stands out as the best.

Regarding functional selection, BLYP is consistently in poor agreement with the CC results. This functional actually predicts a repulsive interaction between CO_2 and the Co^{2+} and Zn^{2+} diformate models. In some cases the addition of dispersion brings it quite close to the benchmark, but it underperforms most of the other functionals. PBE and B3LYP perform very similarly for many systems, though the B3LYP results are better for the FDC systems. With the addition of the dispersion corrections, B3LYP outperforms PBE slightly for those cases where they are close. The dispersion corrections were not necessary for the Minnesota functionals, M05-2X and M06-2X, as these two overestimate the benchmark interaction energy in all cases, sometimes quite severely. PBE0, with any dispersion correction, is the best performing functional tested here.

3.5 Summary

The model systems studied here proved sufficiently single-reference for an accurate treatment with CCSD(T). Of the multireference methods tested, only MRCISD provides any kind of reliably accurate results, and only from the E_{int}^{CP*} reference. Additionally, the computationally less expensive MP2 provides comparable results. The SAPT decomposition of the interaction

energies emphasizes the importance of dispersion for all models, which is echoed by the improvement of the DFT results by the addition of the empirical dispersion corrections. Of the density functionals, the hybrids are the top performers, with PBE0 being better than B3LYP.

Chapter 4

Developments in Symmetry-Adapted Perturbation Theory

As mentioned in subsection 1.2.4, the exchange corrections for SAPT are commonly calculated within the single-exchange approximation to simplify the somewhat complicated process of antisymmetrizing the dimer wavefunction. From 1976 until very recently, the only SAPT exchange correction whose complete form was known was the first-order correction. Two recent works by Schäffer and Jansen[92,93] derived the complete forms of the two second-order exchange corrections, the second-order exchange-induction and exchange-dispersion corrections. An example implementation of the MO forms of these corrections was incorporated into the PSI4NUMPY framework by this dissertation's author.[94] Furthermore, an AO transformation of the second-order exchange-induction correction and a mixed AO/Density-fitting (DF) transformation of the second-order exchange-dispersion correction were contributed to the PSI4 quantum chemistry software package.[95]

On a different but convergent topic, conventional SAPT calculations involving two high-spin open-shell molecules are restricted to the highest spin state of the dimer complex. The splitting between different spin states derives exclusively from the exchange energy of the states, while the RS corrections are the same. The recently developed spin-flip SAPT (SF-SAPT) introduced a method of treating arbitrary spin states of the complex through coupled spin-flipping of the monomers,[96] but this new first-order SF-SAPT exchange energy was derived within the single-exchange approximation. The work presented in Appendix D describes the re-derivation of the first-order SF-SAPT exchange term that replaces the single-exchange approximation with the new single-spin-flip approximation (1-flip), that is based on the complete exchange techniques of Refs. 92 and 93.

This chapter summarizes the work in Appendix D, as well as the author's contributions to the PSI4NUMPY project and the PSI4 quantum chemistry software package.

4.1 Second-Order Complete Exchange

The second-order SRS energy correction is

$$E_{SRS}^{(20)} = \frac{\langle \Psi^A \Psi^B | (V - E_{SRS}^{(10)}) \mathcal{A} | \Psi^{(10)} \rangle}{\langle \Psi^A \Psi^B | \mathcal{A} | \Psi^A \Psi^B \rangle}, \tag{4.1}$$

where $\Psi^{(10)}$ is the first-order RS wavefunction correction. The first-order wavefunction correction splits into the following terms

$$\Psi^{(10)} = \Psi_{Ind,A}^{(10)} + \Psi_{Ind,B}^{(10)} + \Psi_{Disp}^{(10)}, \tag{4.2}$$

where $\Psi_{Ind,A}^{(10)}$ and $\Psi_{Ind,B}^{(10)}$ are the first-order induction wavefunctions and $\Psi_{Disp}^{(10)}$ is the first-order dispersion wavefunction. These corrections have the following forms:

$$\Psi_{Ind,A}^{(10)} = \sum_{ia} s_{ia} \Psi_{i \to a}^{A} \Psi^{B}, \tag{4.3}$$

$$\Psi_{Ind,B}^{(10)} = \sum_{jb} s_{jb} \Psi^{A} \Psi_{j \to b}^{B}, \tag{4.4}$$

$$\Psi_{Disp}^{(10)} = \sum_{ijab} t_{ij}^{ab} \Psi_{i \to a}^{A} \Psi_{j \to b}^{B}, \tag{4.5}$$

where i and a are the occupied and virtual spinorbitals of monomer A, j and b are the occupied and virtual spinorbitals of monomer B, s_{ia} and s_{jb} are the first-order induction amplitudes, t_{ij}^{ab} are the dispersion amplitudes, and the subscripts on Ψ^A and Ψ^B imply an excitation of an electron from an occupied spinorbital to a virtual spinorbital. Based on the last four equations, the second-order SRS energy correction can be rewritten as

$$E_{SRS}^{(20)} = E_{Ind}^{(20)} + E_{Disp}^{(20)}, \tag{4.6}$$

where

$$E_{Ind}^{(20)} = \sum_{ia} s_{ia} \frac{\langle \Psi^A \Psi^B | (V - E_{SRS}^{(10)}) \mathcal{A} | \Psi_{i\to a}^A \Psi^B \rangle}{\langle \Psi^A \Psi^B | \mathcal{A} | \Psi^A \Psi^B \rangle} + \sum_{jb} s_{jb} \frac{\langle \Psi^A \Psi^B | (V - E_{SRS}^{(10)}) \mathcal{A} | \Psi^A \Psi_{j\to b}^B \rangle}{\langle \Psi^A \Psi^B | \mathcal{A} | \Psi^A \Psi^B \rangle}, \tag{4.7}$$

and

$$E_{Disp}^{(20)} = \sum_{ijab} t_{ij}^{ab} \frac{\langle \Psi^A \Psi^B | (V - E_{SRS}^{(10)}) \mathcal{A} | \Psi_{i\to a}^A \Psi_{j\to b}^B \rangle}{\langle \Psi^A \Psi^B | \mathcal{A} | \Psi^A \Psi^B \rangle}. \tag{4.8}$$

The denominator in these terms is defined in Eq. (1.37) and the numerators were determined by Schäffer and Jansen.[92,93] Taking advantage of Cramer's rule and the relationships between various determinants (outlined explicitly in Eqs. (13)-(18) of Appx D.), the complete second-order induction energy $E_{Ind}^{(20)}$ is[92]

$$E_{Ind}^{(20)} = E_{ind}^{(20)} + E_{exch-ind}^{(20)} = \sum_{ia} s_{ia} \Omega_{ia} + \sum_{jb} s_{jb} \Omega_{jb}, \tag{4.9}$$

where

$$\Omega_{ia} = \sum_{i'} (B_{i'a} - \sum_{rj} B_{i'r} D_{rj} S_{ja}) D_{ii'} + \sum_{j} (A_{ja} - \sum_{rj'} A_{jr} D_{rj'} S_{j'a}) D_{ij} + \sum_{i'js} (\langle i'j||as\rangle - \sum_{rj'} \langle i'j||rs\rangle \, D_{rj'} S_{j'a}) (D_{ii'} D_{sj} - D_{ij} D_{si'}) \tag{4.10}$$

and Ω_{jb} is similarly defined by selective replacement of some of the indices corresponding to monomers A and B. Furthermore, the complete dispersion energy $E_{Disp}^{(20)}$ is[93]

$$E_{Disp}^{(20)} = E_{disp}^{(20)} + E_{exch-disp}^{(20)} = \sum_{ijab} t_{ij}^{ab} \Gamma_{ia,jb}, \tag{4.11}$$

where

$$\begin{aligned}\Gamma_{ia,jb} = \Omega_{jb}(\sum_{j'} D_{ij'}S_{j'a}) - \Omega_{ja}(\sum_{i'} D_{ii'}S_{i'b}) + \Omega_{ia}(\sum_{i'} D_{ji'}S_{i'b}) \\ - \Omega_{ib}(\sum_{j'} D_{jj'}S_{j'a}) + \sum_{i'j'}\big(\langle i'j'||ab\rangle + \sum_{rsi''j''}\langle i'j'||rs\rangle\, D_{rj''}S_{j''a}D_{si''}S_{i''b} \\ - \sum_{ri''}\langle i'j'||ar\rangle\, D_{ri''}S_{i''b} - \sum_{rj''}\langle i'j'||rb\rangle\, D_{rj''}S_{j''a}\big)(D_{ii'}D_{jj'} - D_{ij'}D_{ji'}),\end{aligned} \tag{4.12}$$

and the forms of Ω_{ib} and Ω_{ja} can be inferred from Eq. (4.10). Eqs (4.10) and (4.12) can be made more compact by introducing the following T terms,

$$\begin{aligned}T_{ra} &= \sum_{j'} D_{rj'}S_{j'a} \\ T_{rb} &= \sum_{i'} D_{ri'}S_{i'b}\end{aligned} \tag{4.13}$$

and the modified virtual orbitals,

$$\begin{aligned}\chi_{\tilde{a}} &= \chi_a - \sum_r \chi_r T_{ra}, \\ \chi_{\tilde{b}} &= \chi_b - \sum_r \chi_r T_{rb}.\end{aligned} \tag{4.14}$$

Using the above equations, the compacted definitions for Ω_{ia} and $\Gamma_{ia,jb}$ are

$$\Omega_{ia} = \sum_{i'} B_{i'\tilde{a}}D_{ii'} + \sum_j A_{j\tilde{a}}D_{ij} + \sum_{i'js}\langle i'j||\tilde{a}s\rangle\,(D_{ii'}D_{sj} - D_{ij}D_{si'}) \tag{4.15}$$

and

$$\Gamma_{ia,jb} = \Omega_{jb}T_{ia} - \Omega_{ja}T_{ib} + \Omega_{ia}T_{jb} - \Omega_{ib}T_{ja} + \sum_{i'j'}\langle i'j'||\tilde{a}\tilde{b}\rangle\,(D_{ii'}D_{jj'} - D_{ij'}D_{ji'}). \tag{4.16}$$

The complete second-order exchange-induction energy $E^{(20)}_{exch-ind}$ and exchange-dispersion energy $E^{(20)}_{exch-disp}$ are obtained be the subtraction of the corresponding RS corrections from the appropriate complete SRS correction.

The last step necessary to arrive at implementable versions of these terms is to simplify them from spinorbital to orbital form via spin-integration. Taking into account that $\mathbf{D}$ and $\mathbf{T}$ are spin-diagonal, the Ω terms integrate to

$$\begin{aligned}\Omega_{ia} &= 2\sum_{i'} B_{i'\tilde{a}} D_{ii'} + 2\sum_{j} A_{j\tilde{a}} D_{ij} \\ &+ 4\sum_{i'js} \langle i'j|\tilde{a}s\rangle\, D_{ii'} D_{sj} - 2\sum_{i'js} \langle i'j|\tilde{a}s\rangle\, D_{ij} D_{si'} \\ &- 2\sum_{i'js} \langle i'j|s\tilde{a}\rangle\, D_{ii'} D_{sj} + 4\sum_{i'js} \langle i'j|s\tilde{a}\rangle\, D_{ij} D_{si'},\end{aligned} \tag{4.17}$$

where the indices will now index the orbitals instead of the spinorbitals (and will do so throughout the remainder of this section). Likewise, Γ integrates to the following form,

$$\begin{aligned}\Gamma_{ia,jb} &= 2\Omega_{jb} T_{ia} - \Omega_{ja} T_{ib} + 2\Omega_{ia} T_{jb} - \Omega_{ib} T_{ja} \\ &+ 4\sum_{i'j'} \langle i'j'|\tilde{a}\tilde{b}\rangle\, D_{ii'} D_{jj'} - 2\sum_{i'j'} \langle i'j'|\tilde{a}\tilde{b}\rangle\, D_{ij'} D_{ji'} \\ &- 2\sum_{i'j'} \langle i'j'|\tilde{b}\tilde{a}\rangle\, D_{ii'} D_{jj'} + 4\sum_{i'j'} \langle i'j'|\tilde{b}\tilde{a}\rangle\, D_{ij'} D_{ji'}.\end{aligned} \tag{4.18}$$

These orbital forms are presented in a reference implementation prepared by this author as part of the PSI4NUMPY project.

4.1.1 AO Formulation of the Complete $E^{(20)}_{exch-ind}$

The MO formulation of the terms in Eqs. (4.17) and (4.18) is restrictive for their practical application to large systems, since the transformation of the two-electron integrals scales like $\mathcal{O}(N^5)$. One way to relieve some of the cost is to circumvent the integral transformations by recasting the equations into their AO forms. This recast is accomplished by taking advantage of the fact that the one and two-electron integrals can be represented as

$$A_{ij} = \sum_{KL} C_{iK} C_{jL} \mathbf{A}_{KL}, \tag{4.19}$$

$$B_{ij} = \sum_{KL} C_{iK} C_{jL} \mathbf{B}_{KL}, \tag{4.20}$$

$$\langle ij|ab\rangle = \sum_{KLMN} C_{iK} C_{jM} C_{aL} C_{bN} \langle KM|LN\rangle , \tag{4.21}$$

where the capital Roman letters index the atomic orbitals and $\mathbf{A}_{KL}$, $\mathbf{B}_{KL}$, and $\langle KM|LN\rangle$ are the one and two-electron integrals in the AO basis. In doing so, the two-electron integrals can be handled as generalized Coulomb and exchange matrices, defined as

$$\mathbf{J}[\mathbf{X}]_{KL} = \sum_{MN} \langle KM|LN\rangle \, \mathbf{X}_{MN} \qquad \mathbf{K}[\mathbf{X}]_{KL} = \sum_{MN} \langle KN|ML\rangle \, \mathbf{X}_{MN}, \tag{4.22}$$

where $\mathbf{X}_{MN}$ is some suitable matrix in the AO basis. The overlap matrix is similarly decomposable into the HF coefficients and $\mathbf{S}_{KL}$, the AO overlap matrix. Another way of cutting down the cost of transforming the two-electron integrals is by the process of density-fitting.[44] This process approximates the AO two-electron integrals as

$$\langle KM|LN\rangle \approx \sum_{X} b_X^{KL} b_X^{MN}, \tag{4.23}$$

where

$$b_X^{KL} = \sum_{Y} \int\int \phi_K(r_1)\phi_L(r_1) \frac{1}{r_{12}} \eta_Y(r_2) dr_1 dr_2 (\mathbf{J}^{-\frac{1}{2}})_{XY}, \tag{4.24}$$

$$\mathbf{J}_{XY} = \int\int \eta_X(r_1) \frac{1}{r_{12}} \eta_Y(r_2) dr_1 dr_2, \tag{4.25}$$

and $\{\eta\}$ is an auxiliary basis set indexed by X and Y. Density-fitting can be paired with the computation of the generalized Coulomb and exchange matrices, and both methods have efficient implementations in PSI4.

The second-order exchange-induction energy is amenable to a purely AO reformulation. Starting from Eq. (4.17), expanding the modified virtual orbitals and T's, and breaking r into i

and j, the orbital formula for the term in Eq. (4.9) containing s_{ia} is rewritten as

$$
\begin{aligned}
s_{ia}\Omega_{ia} = 2s_{ia}\Big[& \sum_{i'} \big(B_{i'a} - \sum_{i'i''j} B_{i'i''}D_{i''j}S_{ja} - \sum_{i'jj'} B_{i'j}D_{jj'}S_{j'a}\big)D_{ii'} \\
& + \sum_{j} \big(A_{ja} - \sum_{i'jj'} A_{ji'}D_{i'j'}S_{j'a} - \sum_{jj'j''} A_{jj'}D_{j'j''}S_{j''a}\big)D_{ij} \\
& + 2\sum_{i'i''j} \big(\langle i'j|ai''\rangle - \sum_{i'''j'} \langle i'j|i'''i''\rangle\, D_{i'''j'}S_{j'a} - \sum_{j'j''} \langle i'j|j'i''\rangle\, D_{j'j''}S_{j''a}\big)D_{ii'}D_{i''j} \\
& + 2\sum_{i'jj'} \big(\langle i'j|aj'\rangle - \sum_{i''j''} \langle i'j|i''j'\rangle\, D_{i''j''}S_{j''a} - \sum_{j''j'''} \langle i'j|j''j'\rangle\, D_{j''j'''}S_{j'''a}\big)D_{ii'}D_{j'j} \\
& + 2\sum_{i'i''j} \big(\langle i'j|i''a\rangle - \sum_{i'''j'} \langle i'j|i''i'''\rangle\, D_{i'''j'}S_{j'a} - \sum_{j'j''} \langle i'j|i''j'\rangle\, D_{j'j''}S_{j''a}\big)D_{ij}D_{i''i'} \\
& + 2\sum_{i'jj'} \big(\langle i'j|j'a\rangle - \sum_{i''j''} \langle i'j|j'i''\rangle\, D_{i''j''}S_{j''a} - \sum_{j''j'''} \langle i'j|j'j''\rangle\, D_{j''j'''}S_{j'''a}\big)D_{ij}D_{j'i'} \\
& - \sum_{i'i''j} \big(\langle i'j|ai''\rangle + \sum_{i'''j'} \langle i'j|i'''i''\rangle\, D_{i'''j'}S_{j'a} + \sum_{j'j''} \langle i'j|j'i''\rangle\, D_{j'j''}S_{j''a}\big)D_{ij}D_{i''i'} \\
& - \sum_{i'jj'} \big(\langle i'j|aj'\rangle + \sum_{i''j''} \langle i'j|i''j'\rangle\, D_{i''j''}S_{j''a} + \sum_{j''j'''} \langle i'j|j''j'\rangle\, D_{j''j'''}S_{j'''a}\big)D_{ij}D_{j'i'} \\
& - \sum_{i'i''j} \big(\langle i'j|i''a\rangle + \sum_{i'''j'} \langle i'j|i''i'''\rangle\, D_{i'''j'}S_{j'a} + \sum_{j'j''} \langle i'j|i''j'\rangle\, D_{j'j''}S_{j''a}\big)D_{ii'}D_{i''j} \\
& - \sum_{i'jj'} \big(\langle i'j|j'a\rangle + \sum_{i''j''} \langle i'j|j'i''\rangle\, D_{i''j''}S_{j''a} + \sum_{j''j'''} \langle i'j|j'j''\rangle\, D_{j''j'''}S_{j'''a}\big)D_{ii'}D_{j'j}\Big].
\end{aligned}
\tag{4.26}
$$

The systematic decomposition of the integrals and overlap terms into their AO forms provides the coefficients necessary to "back-transform" the induction amplitudes and D terms in the following ways,

$$(\mathbf{D}_{ij})_{KL} = \sum_{ij} C_{iK}C_{jL}D_{ij}, \tag{4.27}$$

and

$$(s\mathbf{D}_{ij})_{KL} = \sum_{ija} C_{aK}C_{jL}s_{ia}D_{ij}, \tag{4.28}$$

where the subscripts inside the parentheses specify which block of the $\mathbf{D}$ matrix was included in the transformation (e.g. $D_{ii'}$, D_{ij}, or $D_{jj'}$). With these definitions, the AO form of Eq. (4.26)

is

$$
\begin{aligned}
s_{ia}\Omega_{ia} = 2(s\mathbf{D}_{ii'}) \cdot \Big[& \mathbf{B} - \mathbf{S}\mathbf{D}_{ij}^{T}\mathbf{B} - \mathbf{S}\mathbf{D}_{jj'}\mathbf{B} \\
& + 2\mathbf{J}[\mathbf{D}_{ij}] - 2\mathbf{S}\mathbf{D}_{ij}^{T}\mathbf{J}[\mathbf{D}_{ij}] - 2\mathbf{S}\mathbf{D}_{jj'}\mathbf{J}[\mathbf{D}_{ij}] \\
& + 2\mathbf{J}[\mathbf{D}_{jj'}] - 2\mathbf{S}\mathbf{D}_{ij}^{T}\mathbf{J}[\mathbf{D}_{jj'}] - 2\mathbf{S}\mathbf{D}_{jj'}\mathbf{J}[\mathbf{D}_{jj'}] \\
& - \mathbf{K}^{T}[\mathbf{D}_{ij}] + \mathbf{S}\mathbf{D}_{ij}^{T}\mathbf{K}^{T}[\mathbf{D}_{ij}] + \mathbf{S}\mathbf{D}_{jj'}\mathbf{K}^{T}[\mathbf{D}_{ij}] \\
& - \mathbf{K}[\mathbf{D}_{jj'}] + \mathbf{S}\mathbf{D}_{ij}^{T}\mathbf{K}[\mathbf{D}_{jj'}] + \mathbf{S}\mathbf{D}_{jj'}\mathbf{K}[\mathbf{D}_{jj'}] \Big] \\
+ 2(s\mathbf{D}_{ij}) \cdot \Big[& \mathbf{A} - \mathbf{S}\mathbf{D}_{ij}^{T}\mathbf{A} - \mathbf{S}\mathbf{D}_{jj'}\mathbf{A} \\
& + 2\mathbf{J}[\mathbf{D}_{ii'}] - 2\mathbf{S}\mathbf{D}_{ij}^{T}\mathbf{J}[\mathbf{D}_{ii'}] - 2\mathbf{S}\mathbf{D}_{jj'}\mathbf{J}[\mathbf{D}_{ii'}] \\
& + 2\mathbf{J}[\mathbf{D}_{ij}] - 2\mathbf{S}\mathbf{D}_{ij}^{T}\mathbf{J}[\mathbf{D}_{ij}] - 2\mathbf{S}\mathbf{D}_{jj'}\mathbf{J}[\mathbf{D}_{ij}] \\
& - \mathbf{K}[\mathbf{D}_{ii'}] + \mathbf{S}\mathbf{D}_{ij}^{T}\mathbf{K}[\mathbf{D}_{ii'}] + \mathbf{S}\mathbf{D}_{jj'}\mathbf{K}[\mathbf{D}_{ii'}] \\
& - \mathbf{K}[\mathbf{D}_{ij}] + \mathbf{S}\mathbf{D}_{ij}^{T}\mathbf{K}[\mathbf{D}_{ij}] + \mathbf{S}\mathbf{D}_{jj'}\mathbf{K}[\mathbf{D}_{ij}] \Big],
\end{aligned}
\tag{4.29}
$$

where all multiplication operations inside the square brackets are matrix multiplication, and the dots outside of the brackets imply the scalar product of $(s\mathbf{D}_{ii'})$ or $(s\mathbf{D}_{ij})$ and the sum of the bracketed terms. As before, an equivalent series of steps can be performed to arrive at the AO form for $s_{jb}\Omega_{jb}$.

4.1.2 AO/DF Formulation of the Complete $E^{(20)}_{exch-disp}$

As seen in the above reformulation, the Ω terms appearing in Eq. (4.18) can be calculated entirely in an AO fashion. However, the two-electron integrals in that equation can not be expressed as generalized two-electron matrices and must be treated in a DF fashion to gain any amount of speed-up. As an initial step toward the AO/DF form of Eq. (4.18), the modified virtual orbitals $\tilde{\chi}_a$ and $\tilde{\chi}_b$ will be used to define the modified virtual coefficients $\tilde{C}_{aK}$ and $\tilde{C}_{bK}$ that satisfy

$$
\begin{aligned}
\tilde{\chi}_a &= \sum_K \tilde{C}_{aK}\phi_K, \\
\tilde{\chi}_b &= \sum_K \tilde{C}_{bK}\phi_K.
\end{aligned}
\tag{4.30}
$$

Expanding the T's in Eq. (4.14) and decomposing the MOs and overlap terms provides

$$\begin{aligned}\sum_K \tilde{C}_{aK}\phi_K &= \sum_K C_{aK}\phi_K - \sum_{KL}(\mathbf{D}_{rj}\mathbf{S})_{KL}C_{aL}\phi_K,\\ \sum_K \tilde{C}_{bK}\phi_K &= \sum_K C_{bK}\phi_K - \sum_{KL}(\mathbf{D}_{ri}\mathbf{S})_{KL}C_{bL}\phi_K,\end{aligned} \tag{4.31}$$

from which the definitions of $\tilde{C}_{aK}$ and $\tilde{C}_{bK}$ can be inferred. Utilizing these modified coefficients, the two-electron terms in Eq. (4.18) decompose to

$$\begin{aligned}&4\sum_{i'j'KLMN} C_{i'K}C_{j'M}\tilde{C}_{aL}\tilde{C}_{bN}\langle KM|LN\rangle D_{ii'}D_{jj'}\\ &-2\sum_{i'j'KLMN} C_{i'K}C_{j'M}\tilde{C}_{aL}\tilde{C}_{bN}\langle KM|LN\rangle D_{ij'}D_{ji'}\\ &-2\sum_{i'j'KLMN} C_{i'K}C_{j'M}\tilde{C}_{aN}\tilde{C}_{bL}\langle KM|LN\rangle D_{ii'}D_{jj'}\\ &+4\sum_{i'j'KLMN} C_{i'K}C_{j'M}\tilde{C}_{aN}\tilde{C}_{bL}\langle KM|LN\rangle D_{ij'}D_{ji'}.\end{aligned} \tag{4.32}$$

Introducing the DF approximation of Eq. (4.23), the terms in Eq. (4.32) can be rewritten as

$$\begin{aligned}&4\sum_{i'j'X} b_X^{i'\tilde{a}}b_X^{j'\tilde{b}}D_{ii'}D_{jj'} - 2\sum_{i'j'X} b_X^{i'\tilde{a}}b_X^{j'\tilde{b}}D_{ij'}D_{ji'}\\ &-2\sum_{i'j'X} b_X^{i'\tilde{b}}b_X^{j'\tilde{a}}D_{ii'}D_{jj'} + 4\sum_{i'j'X} b_X^{i'\tilde{b}}b_X^{j'\tilde{a}}D_{ij'}D_{ji'},\end{aligned} \tag{4.33}$$

where the dual-basis b tensors are defined as

$$\begin{aligned}b_X^{i'\tilde{a}} &= \sum_{KL} b_X^{KL}C_{i'K}\tilde{C}_{aL}, \quad b_X^{i'\tilde{b}} = \sum_{KL} b_X^{KL}C_{i'K}\tilde{C}_{bL},\\ b_X^{j'\tilde{a}} &= \sum_{KL} b_X^{KL}C_{j'K}\tilde{C}_{aL}, \quad b_X^{j'\tilde{b}} = \sum_{KL} b_X^{KL}C_{j'K}\tilde{C}_{bL}.\end{aligned} \tag{4.34}$$

Furthermore, it is possible to construct an alternative AO form for the Ω terms using these modified coefficients. Beginning with Eq. (4.17) and decomposing the integrals into their AO

Table 4.1: Average timings (s) for the calculation of the two second-order exchange corrections of the A24 dataset in either the MO or AO form. The column labeled AO_1 refers to the initial AO formulation for $E^{(20)}_{exch-ind}$, while AO_2 refers to the form based on the use of the modified HF coefficients.

Term	MO	AO_1	AO_2
$E^{(20)}_{exch-ind}$	12.648	0.064	0.057
$E^{(20)}_{exch-disp}$	48.333		9.553

forms leads to

$$\begin{aligned}\Omega_{ia} =& 2\Big(\sum_{i'KL} D_{ii'} C_{i'K} (\mathbf{B} + 2\mathbf{J}[\mathbf{D}_{rj'}] - \mathbf{K}[\mathbf{D}_{rj'}])_{KL} \tilde{C}_{aL} \\ &+ \sum_{j'KL} D_{ij'} C_{j'K} (\mathbf{A} + 2\mathbf{J}[\mathbf{D}_{ri'}] - \mathbf{K}[\mathbf{D}_{ri'}])_{KL} \tilde{C}_{aL} \Big),\end{aligned} \tag{4.35}$$

and similar results for the other Ω terms. This version of the Ω terms is equivalent to the previous one, but is more preferable when the exchange-dispersion is also expected to be calculated as it utilizes many of the same intermediate terms as the two-electron parts of $\Gamma_{ia,jb}$.

The first AO algorithm for the second-order exchange-induction term was implemented into the public version of PSI4 in Python, while the AO/DF algorithm for the second-order exchange-dispersion term was implemented in C++. Both the MO and AO algorithms were tested in the PSI4NUMPY framework to provide a fair comparison of the timings. All pertinent algorithms were tested on a subset of the A24 dataset[97] with the two Argon containing systems removed (due to basis set availability). Additionally, the equilibrium separations for the Ne and He dimers were added. Both $E^{(20)}_{exch-ind}$ and $E^{(20)}_{exch-disp}$ were calculated for the resulting 24 systems at both the minimum geometries and at 0.8 times their minimum intermolecular separation. Table 4.1 shows the average time to calculate these quantities for the 48 geometries. The two AO algorithms for the exchange-induction term are both around 200x faster than the MO version, with the algorithm based on the modified virtual coefficients being slightly faster on average. The speed-up of the exchange-dispersion term is more limited at around 5x.

Table 4.2: Signed relative errors (%) of the single-exchange approximation for the two second-order exchange corrections for the A24 dataset. The values presented are for the minimum intermolecular separation R_0 and $0.8R_0$. Repeated pairs of molecules represent different geometric configurations.[97]

System	$E^{(20)}_{exch-ind}$		$E^{(20)}_{exch-disp}$	
R/R_0	0.8	1.0	0.8	1.0
$H_2O \cdots NH_3$	15.15	2.95	-6.54	-0.44
$H_2O \cdots H_2O$	10.55	1.81	-4.86	-0.40
$HCN \cdots HCN$	24.27	1.75	-8.10	0.24
$HF \cdots HF$	6.99	1.15	-4.88	-0.48
$NH_3 \cdots NH_3$	8.27	1.12	-2.30	-0.17
$CH_4 \cdots HF$	5.14	0.70	-1.33	-0.02
$NH_3 \cdots CH_4$	5.64	0.59	-0.22	0.08
$CH_4 \cdots H_2O$	3.98	0.35	-0.15	0.05
$CH_2O \cdots CH_2O$	7.45	1.34	-4.08	-0.61
$C_2H_4 \cdots H_2O$	7.28	1.32	-2.19	-0.17
$C_2H_4 \cdots CH_2O$	4.46	0.71	-1.39	-0.07
$C_2H_2 \cdots C_2H_2$	10.82	1.01	0.01	0.18
$C_2H_4 \cdots NH_3$	6.24	0.93	-0.56	0.05
$C_2H_4 \cdots C_2H_4$	5.28	0.83	-0.19	0.12
$C_2H_4 \cdots CH_4$	4.44	0.50	0.28	0.09
$CH_4 \cdots BH_3$	6.17	1.11	-2.77	-0.23
$CH_4 \cdots C_2H_6$	2.90	0.33	0.00	0.00
$C_2H_6 \cdots CH_4$	3.76	0.24	0.00	0.00
$CH_4 \cdots CH_4$	2.29	0.23	-0.20	0.02
$C_2H_4 \cdots C_2H_2$	7.15	1.35	-0.23	0.08
$C_2H_4 \cdots C_2H_4$	8.08	1.64	-0.18	0.13
$C_2H_2 \cdots C_2H_2$	6.28	1.10	-0.59	0.01
$He \cdots He$	0.63	0.00	0.00	0.00
$Ne \cdots Ne$	0.16	0.00	-0.46	-0.03

For the 48 geometries described above, the signed relative errors (SREs) of the single-exchange approximation relative to the complete exchange values for both second-order corrections were calculated as

$$SRE(E^{(20)}_{exch-X}) = \frac{E^{(20)}_{exch-X}(S^\infty) - E^{(20)}_{exch-X}(S^2)}{E^{(20)}_{exch-X}(S^\infty)} \cdot 100\% \tag{4.36}$$

and are presented in Table 4.2. These relative errors clearly show the breakdown of the S^2 approximation at shorter separations, as well as the variable severity of the breakdown. The

SREs for the exchange-induction values are generally larger in magnitude than the exchange-dispersion values. Furthermore, the error in the exchange-induction S^2 results can be quite considerable, e.g. the HCN dimer where $E^{(20)}_{exch-ind}(S^2)$ recovers only 75% of $E^{(20)}_{exch-ind}(S^\infty)$ at $0.8R_0$. The exchange-induction SREs are also always positive, showing that the single-exchange approximation always underestimates $E^{(20)}_{exch-ind}$. The exchange-dispersion results are seen to be positive or negative, though they trend toward negative for the shorter separations.

4.2 First-Order SF-SAPT Complete Exchange

While SAPT is a useful tool for the examination of interactions between molecules, it does have limitations to its applications. Until recently, one such limitation was the inability to treat arbitrary spin states of complexes containing multiple open-shell molecules. Even when the interacting molecules are specifically in their individual high-spin states, the complex can have a range of different spin states. Given that molecules A and B have corresponding spin quantum numbers S_A and S_B, the complex can have a spin quantum number ranging from $|S_A - S_B|$ to $S_A + S_B$. The interaction energies of these various spin states differ only in the exchange terms derived from the SRS corrections, while the RS corrections are the same for all spin states. To calculate the SRS correction for a given spin state of the complex, the dimer wavefunction needs to either be a pure spin function or to be projected onto the subspace that corresponds to a particular spin.[98,99] The conventional formalism takes advantage of the fact that the dimer function $\Psi^{(0)}$ is a pure spin function when all unpaired electrons in Ψ_A and Ψ_B have the same spin. This formulation means that the spin projection is not necessary, but also limits the complex to its highest spin state.

If the SRS corrections for a lower spin state are desired, the antisymmetrizer $\mathcal{A}$ is accompanied by the spin projection operator P_{SM_S} that projects the dimer wavefunction onto a particular spin state. In the case of the first-order SRS correction, this leads to

$$E^{(10)}_{SRS} = \frac{\langle \Psi^{(0)} | V \mathcal{A} P_{SM_S} | \Psi^{(0)} \rangle}{\langle \Psi^{(0)} | \mathcal{A} P_{SM_S} | \Psi^{(0)} \rangle}. \tag{4.37}$$

If $\Psi^{(0)}$ is the product of Ψ_A and Ψ_B, which are ROHF wavefunctions where all unpaired electrons on A are spin up and all unpaired electrons on B are spin down, the action of P_{SM_S} on $\Psi^{(0)}$ is represented as

$$P_{SM_S}\Psi_A\Psi_B = c_0\Psi_A\Psi_B + c_1\Psi_A^{\downarrow}\Psi_B^{\uparrow} + c_2\Psi_A^{\downarrow\downarrow}\Psi_B^{\uparrow\uparrow} + \ldots, \tag{4.38}$$

where c_n are the Clebsch-Gordan coefficients

$$c_0 = \langle S(S_A - S_B)|S_A S_A S_B(-S_B)\rangle\,, \tag{4.39}$$

$$c_1 = \langle S(S_A - S_B)|S_A(S_A - 1)S_B(-S_B + 1)\rangle\,, \tag{4.40}$$

$$c_2 = \langle S(S_A - S_B)|S_A(S_A - 2)S_B(-S_B + 2)\rangle\,, \tag{4.41}$$

and the wavefunctions with the arrows are spin-flipped wavefunctions.[96] The spin-flipped wavefunctions are normalized linear combinations of all wavefunctions where a number of unpaired electrons, equal to the number of arrows, has had their spins flipped. As the unpaired electrons in A are all spin up and the unpaired electrons in B are all spin down, the electrons in Ψ_A can only be flipped down and the electrons in Ψ_B can only be flipped up. The terms arising from the action of P_{SM_S} on the dimer wavefunction are therefore the result of a particular number of intermolecular spin-flips, giving rise to the name spin-flip SAPT (SF-SAPT).

From Eqs. (4.37) and (4.38), Patkowski et al. derived the first-order ROHF-based SF-SAPT exchange correction within the S^2 approximation.[96] The resulting formulas express the first-order exchange energy of the different spin states of the complex as the sum of a diagonal term and a spin-flip term, which is scaled using the appropriate Clebsch-Gordan coefficients based on the desired spin state. A result of the S^2 approximation is that all terms containing wavefunctions with more than a single spin-flip reduce to zero, limiting the formalism to only one spin-flip. While the resulting formalism provides the opportunity for the calculation of spin splittings between various complex spin states, the use of the S^2 approximation engenders the same issues as in the conventional SAPT formalism. As such, the first-order SF-SAPT exchange correction was re-derived without the S^2 approximation.

The new derivation of the SF-SAPT exchange correction also begins with Eqs. (4.37) and (4.38), but obviously forgoes the introduction of the approximate antisymmetrizer. Therefore, Eqs. (4.37) is expanded as

$$E_{SRS}^{(10)} = \frac{c_0 \langle \Psi_A \Psi_B | V\mathcal{A} | \Psi_A \Psi_B \rangle + c_1 \langle \Psi_A \Psi_B | V\mathcal{A} | \Psi_A^{\downarrow} \Psi_B^{\uparrow} \rangle + \ldots}{c_0 \langle \Psi_A \Psi_B | \mathcal{A} | \Psi_A \Psi_B \rangle + c_1 \langle \Psi_A \Psi_B | \mathcal{A} | \Psi_A^{\downarrow} \Psi_B^{\uparrow} \rangle + \ldots}. \tag{4.42}$$

While the S^2 approximation reduced all terms with more than a single spin-flip to zero, these terms do not necessarily vanish when the complete antisymmetrizer is used. Instead, the decision is made to ignore all terms with more than a single spin-flip (i.e. all terms not explicitly present in Eq. (4.42)), giving rise to the single-spin-flip approximation (1-flip). The leading terms in the numerator and denominator are effectively the same ones required for the calculation of the conventional first-order exchange correction. The second terms can be considered as a particular subset of the double excitations appearing in Eq. (4.8), where only the unpaired electrons are promoted to the virtual orbitals with the same spatial part but opposite spin.[100] The application of the techniques used by Schäffer and Jansen for the complete $E_{exch-disp}^{(20)}$ to the derivation of these new spin-flip terms is detailed in Appx. D, as well as the reformulation of the newly derived MO terms into their AO forms.

The resulting equations are similar to the S^2 variant, in that the numerator and denominator are both sums of a diagonal part, that is common to all spin states, and a spin-flipped part that is scaled by the function in Eq. (37) of Appx D to select for a specific spin state. This function is the same as the one that appears in the S^2 terms and is derived from the Clebsch-Gordan coefficients and the normalization of the spin-flipped wavefunctions. The 1-flip approximation is formally exact for any complex where one of the interacting molecules is a doublet, as this limits the number of possible spin-flips to only one. The effect of the 1-flip approximation on systems where both molecules have spin states higher than a doublet was investigated on several smaller systems, along with a comparison of the new approximation versus the previous S^2 approximation.

4.2.1 Small Test Systems

The first test system considered was the interaction of Li and H. As both Li and H are doublets, the 1-flip approximation is not actually an approximation for this complex, as stated above. Additionally, the singular electron on H means that the single-exchange approximation appears to be complete for this test case. Both approximate values of $E_{exch}^{(10)}$ for the singlet and triplet states at various intermolecular separations are presented in Table I of Appx. D, alongside the FCI-based SAPT results[101] (i.e. the HF wavefunctions are replaced with FCI wavefunctions in the calculation of the SAPT corrections). Though both approximations are expected to be the same, they begin to deviate from each other around 10.0 bohr. An investigation into the difference between the two approximations led to the realization that, even for systems where one molecule has only one electron, the common implementation of the S^2 approximation is never formally exact. The S^2 approximation can be described as two parts. The first is the replacement of the antisymmetrizer with its approximate form containing the single-exchange operator. The second part is the elimination of all terms that contain overlap terms higher than S^2. It is this second part that renders the S^2 approximation inexact for all cases (for more detail, see the text around Eq. (58) in Appx. D).

For Li$\cdots$Li and Li$\cdots$N, the 1-flip approximation is still exact due to Li's doublet state. The comparisons of the S^2 and 1-flip $E_{exch}^{(10)}$ values for these two systems are shown in Figs. 1 and 2 of Appx. D, respectively. In the Li dimer case, it can be seen that the S^2 results for the singlet and triplet states cross at short distance, while the behavior at the shortest distance for Li$\cdots$N suggests that the same thing will happen at shorter separations. These results contrast the 1-flip ones, where the correct ordering of the spin states is observed throughout the entire presented range.

The N dimer is a case where neither approximation is formally exact, and provides a chance to juxtapose the S^2 and 1-flip results. The first-order exchange correction for the highest spin state, the septet, can be calculated exactly in the conventional formalism,[44] and is presented together with both SF-SAPT results in Fig. 3 of Appx. D. The 1-flip approximation is observed to be considerably milder than the S^2 one, with its deviation from the exact result being smaller

and beginning at a shorter distance. Similar results are seen for the Mn dimer in Table II of that appendix, though it is important to note the similarity of the energy differences between the highest and lowest spin states produced by both approximations. The adequacy of the S^2 spin state splittings is also seen for the different orientations of the O_2 dimer in Fig. 4 of Appx D. In the presented range, only the shortest separations of the linear configuration of this dimer show any noticeable difference between the approximations.

4.2.2 Pancake Bonded Systems

As a more interesting application of first-order SF-SAPT, the spin splittings between the highest and lowest spin states of a selection of pancake bonded systems were calculated.[102–105] Pancake bonded systems consist of interacting doublet radicals, where the singly-occupied orbitals are highly delocalized over the molecule. The resulting interactions are different from typical π-stacking interactions and have strengths intermediate between covalent and van der Waals complexes. Additionally, the orientation of these systems prefers a direct alignment of the atoms of the molecules, as opposed to the typically "slipped" orientations of noncovalently interacting systems. This direct alignment can be seen for one of the test systems, the phenalenyl (PLY) dimer, in Fig. 4.1. Along with the PLY dimer, four dimers of PLY derivative and the trioxotriangulene (TOT) dimer were studied. The structures of these systems can be seen in Fig. 5 of Appx. D. For PLY and its derivatives, only the staggered conformations was studied. Both the staggered and eclipsed conformations of the TOT dimer were considered.

The singlet-triplet splittings for PLY_2 calculated with both SF-SAPT variations are presented in Fig. 6 of Appx. D, alongside CASSCF, M05-2X, and multireference averaged quadratic coupled cluster (MR-AQCC) results for comparison.[106,107] SF-SAPT has the worst agreement with the CC based reference, which is reasonable given that the method is still only first-order. The more interesting feature of this figure is the lack of difference between the S^2 and 1-flip splitting results. As is to be expected, the S^2 curve does show some deviation as the intermolecular separation decreases, but overall the inclusion of complete exchange does little to improve the splitting estimates. Fig. 7 of Appx. D shows a similar trend for the PLY

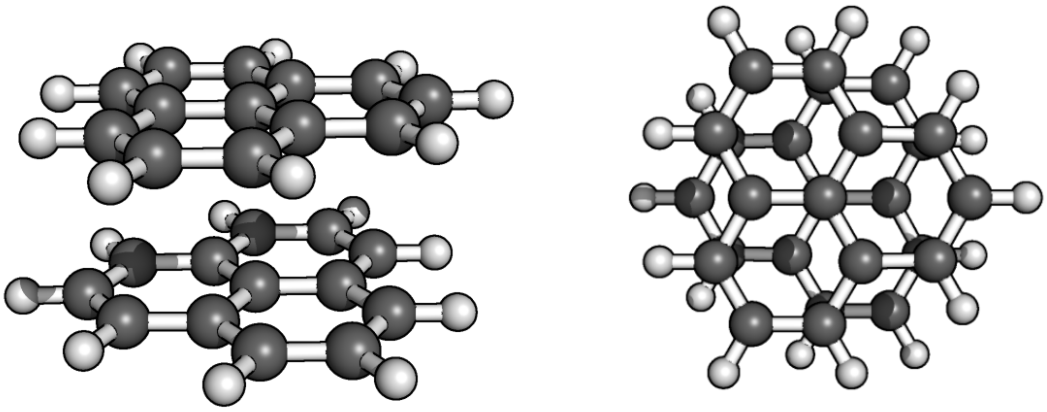


Figure 4.1: The staggered conformation of the phenalenyl radical dimer, with carbon atoms colored gray and hydrogen atoms white. Sideview on left and topdown on right.

derivative dimers, though the S^2 approximation actually overestimates the splitting values at the shorter separations presented.

Appendix D's Table III presents the splittings for both conformations of the TOT dimer and provides some explanation for the surprisingly good S^2 based splitting values. In the smaller systems, it can be observed that the S^2 approximation deteriorates earlier for the highest spin state than for the lowest. However, in the Mn and TOT dimers, it is observed that the S^2 approximation begins to break down at around the same distance for the highest and lowest spin states. The net effect is that the difference between the $E_{exch}^{(10)}$ values remains relatively unchanged even as the individual results begin to deviate from the 1-flip ones.

Chapter 5

Conclusions

Toward the calculation of new, high-accuracy estimates of the most important thermophysical properties for the krypton gas, both a new potential energy curve and the interaction-induced isotropic pair polarizability for the Kr dimer were calculated. To ensure the desired level of accuracy in the properties calculated from these quantities, considerable effort was taken to go beyond the state-of-the-art CCSD(T)/CBS level of theory. At the frozen core level, the basis space was fully saturated using a combination of midbond functions, basis set extrapolation, and explicit correlation. The importance of higher orders of coupled cluster theory, relativistics, and core-core and core-valence correlation were all considered to insure spectroscopic accuracy around the minimum separation. Analytic functions were fitted to the *ab initio* results for use in the calculation of the viscosity, thermal conductivity, self-diffusion coefficient, thermal diffusion factor, and the second pressure, acoustic, and dielectric virial coefficients of dilute krypton gas. The resulting property values are among the most accurate available.

The interactions of CO_2 with representative models of MOFs were studied with high-accuracy *ab initio* methods, DFT, and SAPT. The models included metal species with a varying number of d-electrons to investigate the importance of a multireference treatment of such systems. The tested systems proved to be sufficiently single-reference for evaluation with CCSD(T), with MP2 generally outperforming MRCISD. SAPT decomposition of the interactions showed the persistent importance of dispersion, even for those systems that were predominately induction bound. The CC results were used to test a selection of density functionals, where the necessity of empirical dispersion echoed the SAPT results. PBE0 proved the most accurate functional of those considered.

The second-order SAPT exchange corrections without the single exchange approximation, $E^{(20)}_{exch-ind}$ and $E^{(20)}_{exch-disp}$, were implemented in the public version of PSI4, after being reformulated into AO and combined AO/DF forms, respectively. In addition, the first-order SF-SAPT exchange correction was derived without the S^2 approximation, instead introducing the 1-flip approximation. The new approximation is shown to be considerably milder than the S^2 one, and is formally exact for any complex containing a doublet. Both approximations are applied to the calculation of the spin state splittings for a number of pancake bonded systems, where the 1-flip approximation shows little improvement over the previous S^2 results. The splittings resulting from the S^2 approximation are similar to the 1-flip results due to an even deviation of the individual $E^{(10)}_{exch}$ values of the highest and lowest spin states of the complexes.

5.1 Future and Outlook

In the case of the MOF models, the work presented here is reasonably only a start. The models studied are a minuscule subset of such structures, with a wide variety of organics and metal species in different configurations to consider. Additionally, while the models studied were not especially multireference, this does not lead to the conclusion that a multireference treatment is never required. The application of SAPT to systems like these is still non-trivial, but the benefits provided by SAPT analysis outweigh the difficulty. Furthermore, a wider selection of density functionals could be considered.

There remain a number of directions to continue the development of the SAPT formalism. The third order exchange correction has been derived within the single-exchange approximation,[108] but the corresponding complete exchange forms have not yet been derived. Beyond that, ever higher orders of SAPT corrections are also feasible. The same can be said for SF-SAPT, where the obvious next step is to move on to the second-order exchange corrections.

References

[1] Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: An Introduction to Advanced Electronic Structure Theory*; McGraw-Hill, 1989.

[2] Born, M.; Oppenheimer, J. R. *Ann. d. Phys.* **1927**, *84*, 457.

[3] Cramer, C. J. *Essentials of Computational Chemistry*; John Wiley Sons, 2002.

[4] Leach, A. R. *Molecular Modelling: Principles and Applications*; Prentice-Hall, 2001.

[5] Hartree, D. R. *Proc. Cambridge Phil. Soc.* **1928**, *24*, 111.

[6] Fock, V. *Z. f. Phys.* **1930**, *61*, 126.

[7] Hartree, D. R.; Hartree, W. *Proc. Roy. Soc.* **1935**, *150*, 9.

[8] Slater, J. C. *Phys. Rev.* **1930**, *35*, 210.

[9] Roothaan, C. C. J. *Rev. Mod. Phys.* **1951**, *23*, 69.

[10] Cizek, J. *J. Chem. Phys.* **1966**, *45*, 4256.

[11] Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618.

[12] Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, *157*, 479.

[13] Woon, D. E. *J. Chem. Phys.* **1994**, *100*, 2838.

[14] Woon, D. E.; Dunning, T. H. *J. Chem. Phys.* **1994**, *101*, 8877.

[15] Peterson, K. A.; Dunning, T. H. *J. Phys. Chem.* **1995**, *99*, 3898.

[16] Roos, B. O.; Taylor, P. R.; Siegbahn, P. E. M. *Chem. Phys.* **1980**, *48*, 157.

[17] Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. *J. Chem. Phys.* **1992**, *96*, 1218.

[18] Werner, H.-J. *Mol. Phys.* **1996**, *89*, 645.

[19] Celani, P.; Werner, H.-J. *J. Chem. Phys.* **2000**, *112*, 5546.

[20] Angeli, C.; Cimiraglia, R.; Evangelisti, S.; Leininger, T.; Malrieu, J.-P. *J. Chem. Phys.* **2001**, *114*, 10252.

[21] Angeli, C.; Cimiraglia, R.; Malrieu, J.-P. *J. Chem. Phys.* **2002**, *117*, 9138.

[22] Angeli, C.; Pastore, M.; Cimiraglia, R. *Theor. Chem. Acc.* **2006**, *117*, 743.

[23] Evangelista, F. A. *J. Chem. Phys.* **2018**, *149*, 030901.

[24] Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864.

[25] Kohn, W.; Sham, L. J. *Phys. rev.* **1965**, *140*, A1133.

[26] Becke, A. D. *J. Chem. Phys.* **2014**, *140*, 18A301.

[27] Boys, S.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.

[28] Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007.

[29] Kendall, R. A.; Dunning, T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796.

[30] Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. *Chem. Phys. Lett.* **1998**, *286*, 243.

[31] Tao, F.-M.; Pan, Y.-K. *J. Chem. Phys.* **1992**, *97*, 4989.

[32] Patkowski, K.; Dixon, D. A., Ed.; Annual Reports in Computational Chemistry, Vol. 13; Elsevier, 2017; p 3.

[33] Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787.

[34] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104.

[35] Grimme, S.; Ehrlich, S.; Goerigk, L. *J. Comput. Chem.* **2011**, *32*, 1456.

[36] Caldeweyher, E.; Bannwarth, C.; Grimme, S. *J. Chem. Phys.* **2017**, *147*, 034112.

[37] Smith, D. G. A.; Burns, L. A.; Patkowski, K.; Sherrill, C. D. *J. Phys. Chem. Lett.* **2016**, *7*, 2197.

[38] Jeziorski, B.; Moszynski, R.; Szalewicz, K. *Chem. Rev.* **1994**, *94*, 1887.

[39] Jeziorski, B.; Chałsiński, G.; Szalewicz, K. *Int. J. Quantum Chem.* **1978**, *14*, 271.

[40] Jeziorski, B.; Bulski, M.; Piela, L. *Int. J. Quantum Chem.* **1976**, *10*, 281.

[41] Parker, T. M.; Burns, L. A.; Parrish, R. M.; Ryno, A. G.; Sherrill, C. D. *J. Chem. Phys.* **2014**, *140*, 094106.

[42] Żuchowski, P. S.; Podeszwa, R.; Moszyński, R.; Jeziorski, B.; Szalewicz, K. *J. Chem. Phys.* **2008**, *129*, 084101.

[43] Hapka, M.; Żuchowski, P. S.; Szczęśniak, M. M.; Chałasiński, G. *J. Chem. Phys.* **2012**, *137*, 164104.

[44] Gonthier, J. F.; Sherrill, C. D. *J. Chem. Phys.* **2016**, *145*, 134106.

[45] Hellmann, R.; Bich, E.; Vogel, E. *Mol. Phys.* **2008**, *106*, 133.

[46] Jäger, B.; Hellmann, R.; Bich, E.; Vogel, E. *Mol. Phys.* **2009**, *107*, 2181.

[47] Hellmann, R.; Jäger, B.; Bich, E. *J. Chem. Phys.* **2017**, *147*, 034304.

[48] Bukowski, R.; Szalewicz, K.; Groenenboom, G. C.; van der Avoird, A. *Science* **2007**, *315*, 1249.

[49] Przybytek, M.; Cencek, W.; Komasa, J.; Łach, G.; Jeziorski, B.; Szalewicz, K. *Phys. Rev. Lett.* **2010**, *104*, 183003.

[50] Patkowski, K.; Szalewicz, K. *J. Chem. Phys.* **2010**, *133*, 094304.

[51] Waldrop, J. M.; Song, B.; Patkowski, K.; Wang, X. *J. Chem. Phys.* **2015**, *142*, 204307.

[52] Jäger, B.; Hellmann, R.; Bich, E.; Vogel, E. *J. Chem. Phys.* **2016**, *144*, 114301.

[53] Vogel, E.; Jäger, B.; Hellmann, R.; Bich, E. *Mol. Phys.* **2010**, *108*, 3335.

[54] Cencek, W.; Przybytek, M.; Komasa, J.; Mehl, J. B.; Jeziorski, B.; Szalewicz, K. *J. Chem. Phys.* **2012**, *136*, 224303.

[55] Bich, E.; Hellmann, R.; Vogel, E. *Mol. Phys.* **2008**, *106*, 813.

[56] Song, B.; Wang, X.; Liu, Z. *Mol. Simul.* **2016**, *42*, 9.

[57] Song, B.; Waldrop, J. M.; Wang, X.; Patkowski, K. *J. Chem. Phys.* **2018**, *148*, 024306.

[58] Wilson, A. K.; Woon, D. E.; Peterson, K. A.; Dunning, T. H. *J. Chem. Phys.* **1999**, *110*, 7667.

[59] Hättig, C.; Klopper, W.; Köhn, A.; Tew, D. P. *Chem. Rev.* **2012**, *112*, 4–74.

[60] Kong, L.; Bischoff, F. A.; Valeev, E. F. *Chem. Rev.* **2012**, *112*, 75.

[61] Adler, T. B.; Knizia, G.; Werner, H.-J. *J. Chem. Phys.* **2007**, *127*, 221106.

[62] Knizia, G.; Adler, T. B.; Werner, H.-J. *J. Chem. Phys.* **2009**, *130*, 054104.

[63] Marchetti, O.; Werner, H.-J. *J. Phys. Chem. A* **2009**, *113*, 11580.

[64] Marshall, M. S.; Sherrill, C. D. *J. Chem. Theory Comput.* **2011**, *7*, 3978.

[65] Woon, D. E.; Dunning, T. H. *J. Chem. Phys.* **1995**, *103*, 4572.

[66] Peterson, K. A.; Dunning, T. H. *J. Chem. Phys.* **2002**, *117*, 10548.

[67] Douglas, M.; Kroll, N. M. *Annals of Physics* **1974**, *82*, 89.

[68] Hess, B. A. *Phys. Rev. A* **1986**, *33*, 3742.

[69] Tang, K. T.; Toennies, J. P. *J. Chem. Phys.* **1984**, *80*, 3726.

[70] Dham, A. K.; Allnatt, A. R.; Meath, W. J.; Aziz, R. A. *Mol. Phys.* **1989**, *67*, 1291.

[71] Dyall, K. G. *J. Chem. Phys.* **2001**, *115*, 9136.

[72] Tanaka, Y.; Yoshino, K.; Freeman, D. E. *J. Chem. Phys.* **1973**, *59*, 5160.

[73] Slavíček, P.; Kalus, R.; Paška, P.; Odvárková, I.; Hobza, P.; Malijevský, A. *J. Chem. Phys.* **2003**, *119*, 2102.

[74] LaRocque, P. E.; Lipson, R. H.; Herman, P. R.; Stoicheff, B. P. *J. Chem. Phys.* **1986**, *84*, 6627.

[75] Li, J.-R.; Sculley, J.; Zhou, H.-C. *Chem. Rev.* **2012**, *112*, 869.

[76] Odoh, S. O.; Cramer, C. J.; Truhlar, D. G.; Gagliardi, L. *Chem. Rev.* **2015**, *115*, 6051.

[77] Howe, J. D.; Liu, Y.; Flores, L.; Dixon, D. A.; Sholl, D. S. *J. Chem. Theory Comput.* **2017**, *13*, 1341.

[78] Yazaydin, A. O.; Snurr, R. Q.; Park, T.-H.; Koh, K.; Liu, J.; LeVan, M. D.; Benin, A. I.; Jakubczak, P.; Lanuza, M.; Galloway, D. B.; Low, J. J.; Willis, R. R. *J. Am. Chem. Soc.* **2009**, *131*, 18198.

[79] Yu, J.; Xie, L.-H.; Li, J.-R.; Ma, Y.; Seminario, J. M.; Balbuena, P. B. *Chem. Rev.* **2017**, *117*, 9674.

[80] Grajciar, L.; Bludský, O.; Nachtigall, P. *J. Phys. Chem. Lett.* **2010**, *1*, 3354.

[81] Goings, J. J.; Ohlsen, S. M.; Blaisdell, K. M.; Schofield, D. P. *J. Phys. Chem. A* **2014**, *118*, 7411.

[82] McDaniel, J. G.; Schmidt, J. *J. Phys. Chem. A* **2013**, *117*, 2053.

[83] Lin, L.-C.; Berger, A. H.; Martin, R. L.; Kim, J.; Swisher, J. A.; Jariwala, K.; Rycroft, C. H.; Bhown, A. S.; Deem, M. W.; Haranczyk, M.; Smit, B. *Nat. Mater.* **2012**, *11*, 633.

[84] Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

[85] Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623.

[86] Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.

[87] Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865.

[88] Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158.

[89] Ernzerhof, M.; Scuseria, G. E. *J. Chem. Phys.* **1999**, *110*, 5029.

[90] Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2006**, *2*, 364.

[91] Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215.

[92] Schäffer, R.; Jansen, G. *Theor. Chem. Acc.* **2012**, *131*, 1235.

[93] Schäffer, R.; Jansen, G. *Mol. Phys.* **2013**, *111*, 2570.

[94] Smith, D. G. A.; et al. *J. Chem. Theory Comput.* **2018**, *14*, 3504.

[95] Parrish, R. M.; et al. *J. Chem. Theory Comput.* **2017**, *13*, 3185.

[96] Patkowski, K.; Żuchowski, P. S.; Smith, D. G. A. *J. Chem. Phys.* **2018**, *148*, 164110.

[97] Řezáč, J.; Hobza, P. *J. Chem. Theory Comput.* **2013**, *9*, 2151.

[98] Jeziorski, B. *Mol. Phys.* **2010**, *108*, 3043.

[99] Patkowski, K.; Jeziorski, B.; Korona, T.; Szalewicz, K. *J. Chem. Phys.* **2002**, *117*, 5124.

[100] Waldrop, J. M.; Patkowski, K. *J. Chem. Phys.* **2019**, *150*, 074109.

[101] Patkowski, K.; Korona, T.; Jeziorski, B. *J. Chem. Phys.* **2001**, *115*, 1137.

[102] Kertesz, M. *Chem. Eur. J.* **2019**, *25*, 400.

[103] Cui, Z.-h.; Lischka, H.; Beneberu, H. Z.; Kertesz, M. *J. Am. Chem. Soc.* **2014**, *136*, 5539.

[104] Mou, Z.; Kubo, T.; Kertesz, M. *Chem. Eur. J.* **2015**, *21*, 18230.

[105] Mou, Z.; Kertesz, M. *Angew. Chem. Int. Ed.* **2017**, *56*, 10188.

[106] Mou, Z.; Tian, Y.-H.; Kertesz, M. *Phys. Chem. Chem. Phys.* **2017**, *19*, 24761.

[107] Szalay, P. G.; Bartlett, R. J. *Chem. Phys. Lett.* **1993**, *214*, 481.

[108] Patkowski, K.; Szalewicz, K.; Jeziorski, B. *J. Chem. Phys.* **2006**, *125*, 154107.

Appendices

Appendix A

Accurate *ab initio* potential for the krypton dimer and transport properties of the low-density krypton gas

THE JOURNAL OF CHEMICAL PHYSICS **142**, 204307 (2015)

Accurate *ab initio* potential for the krypton dimer and transport properties of the low-density krypton gas

Jonathan M. Waldrop,[1,a)] Bo Song,[2,a)] Konrad Patkowski,[1,b)] and Xiaopo Wang[2,c)]
[1]*Department of Chemistry and Biochemistry, Auburn University, Auburn, Alabama 36849, USA*
[2]*Key Laboratory of Thermo-Fluid Science and Engineering, Ministry of Education, School of Energy and Power Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China*

(Received 14 April 2015; accepted 13 May 2015; published online 28 May 2015)

A new highly accurate potential energy curve for the krypton dimer was constructed using coupled-cluster calculations up to the singles, doubles, triples, and perturbative quadruples level, including corrections for core-core and core-valence correlation and for relativistic effects. The *ab initio* data points were fitted to an analytic potential which was used to compute the most important transport properties of the krypton gas. The viscosity, thermal conductivity, self-diffusion coefficient, and thermal diffusion factor were calculated by the kinetic theory at low density and temperatures from 116 to 5000 K. The comparisons with literature experimental data as well as with values from other pair potentials indicate that our new potential is superior to all previous ones. The transport property values computed in this work are recommended as standard values over the complete temperature range. © *2015 AIP Publishing LLC.* [http://dx.doi.org/10.1063/1.4921623]

I. INTRODUCTION

Thanks to algorithmic improvements and a steady increase in computing power, it has recently become possible to compute noncovalent interaction energies between atoms and/or small molecules to unprecedented accuracy. The newly developed, highly accurate *ab initio* interaction potentials have had significant impact on both theory and experiment; their recent successes include the determination[1,2] of helium pair potential to a millikelvin accuracy necessary (and sufficient) to devise improved, helium-based pressure and temperature standards,[3] the resolution of a long-standing disagreement between theory and experiment for the binding energy of the beryllium dimer,[4–10] and the assignment of the extremely complicated infrared spectrum of the *ortho*-H_2–CO van der Waals complex.[11,12] These successes would not have been possible without both an accurate determination of the complete basis set (CBS) limit of the leading interaction energy contribution, computed by the coupled-cluster approach with singles, doubles, and perturbative triples (CCSD(T)), and a reliable account of all important interaction energy corrections past the frozen-core (FC) CCSD(T) level, including a correction for core-core and core-valence correlation, for coupled-cluster excitations beyond CCSD(T), and for relativistic effects.

In the field of thermophysical properties, the progress on *ab initio* potentials has also improved the accuracy of transport property data for monatomic gases. Cencek *et al.*[1] evaluated the zero-density viscosity and thermal conductivity of helium from the state-of-the-art potential of Przybytek *et al.*[2] The relative uncertainties of the *ab initio* values are only 2×10^{-5} for temperatures above 50 K and increase to 5×10^{-4} at low temperatures. This level of uncertainty is significantly smaller than that of the corresponding measurements of the viscosity and thermal conductivity. Bich *et al.*[13,14] determined a pair potential for neon by first-principles calculations and further derived the viscosity and thermal conductivity of neon from their high-quality potential. The estimated uncertainties of the two properties are about ±0.1% for the complete temperature range of the calculations except at the lowest temperatures. Vogel *et al.*[15] and Song *et al.*[16] independently computed the transport properties of argon from two different *ab initio* potentials, that is, the ones of Jäger *et al.*[17,18] and Patkowski and Szalewicz,[19] respectively. The calculated transport property values are as accurate as the best experimental measurements at room temperature. For temperatures above and below ambient temperature, the theoretical values are expected to be more reliable than the available experimental data.

The *ab initio* values of viscosity are becoming increasingly important for the calibration of high-precision viscometers. Vogel[20] based their measurements on the theoretical viscosity of argon at 298.15 K[15] and reduced the measurement uncertainties to 0.2% at temperatures up to 700 K. Berg and Moldover[21] determined recommended viscosities of ten common gases at 25 °C by re-evaluating the literature viscosity data with a new helium result calculated *ab initio*.[1] Their work established a scale for gas viscosities that is more accurate than most of the reported values. Another potential application of *ab initio* results is their use in the development of wide-ranging correlations for transport properties. Roder *et al.*[22] showed in their measurements on argon that the relative uncertainty of the thermal conductivity data increased from 0.3% at higher densities to 1% or even 2% at lower densities. In a historical review on the thermal conductivity,[23] it was suggested that the experimental data seemed to deviate systematically from expectations in the low-pressure region. As a result, Assael

[a)]J. M. Waldrop and B. Song contributed equally to this work.
[b)]E-mail: patkowsk@auburn.edu
[c)]E-mail: wangxp@mail.xjtu.edu.cn

et al.[24] utilized theoretical values instead of critically assessed experimental data to develop the correlation for the thermal conductivity of hydrogen in the dilute-gas limit. The thermal conductivity is well represented by the new correlation over the extended temperature range between the triple point and 1000 K.

Since *ab initio* values of transport properties are of interest in many fields, the study here is aimed at their accurate determination for the krypton gas. In this work, we develop a new potential energy curve for the krypton dimer by high-level *ab initio* calculations. The interatomic interaction energies from high-level coupled-cluster calculations (including core and relativistic corrections) are fitted to an analytical representation which is applied in this work to investigate transport properties of krypton at low density. The viscosity, thermal conductivity, self-diffusion coefficient, and thermal diffusion factor are computed by the classical kinetic theory of dilute gases over a wide range of temperatures from 116 K (the triple-point temperature of krypton) to 5000 K.

II. THE *AB INITIO* DATA POINTS

The potential energy curve for the krypton dimer was calculated using the coupled-cluster method up to perturbative quadruple excitations (CCSDT(Q)).[25,26] All CCSD(T), explicitly correlated CCSD(T) (CCSD(T)–F12), and relativistic results were calculated using the MOLPRO code.[27] The CCSDT and CCSDT(Q) corrections[28–30] were calculated using the MRCC code.[31] Basis sets used in these calculations include aug-cc-pVXZ (aVXZ), aug-cc-pCVXZ (aCVXZ), and aug-cc-pwCVXZ (awCVXZ).[32–34] Some calculations used midbond functions (denoted by, e.g., aVXZM), which were hydrogenic functions from the same basis as for the krypton atoms. These functions were placed halfway between the krypton atoms. The counterpoise correction was utilized in all calculations to counter the basis set superposition error.[35]

An extensive examination of a near van der Waals minimum point ($R = 4.0$ Å, 1 Å $= 10^{-10}$ m) was performed to determine the basis sets and levels of theory required to produce the desired level of accuracy. The interaction energy was calculated as

$$E_{int} = E_{int}^{CCSD(T)/FC} + \Delta E_{int}^{CCSD(T)/AE-FC} + \Delta E_{int}^{(Q)-(T)/FC} + \Delta E_{int}^{rel}. \quad (1)$$

The subscript *int* implies the counterpoise corrected interaction energy. The first term represents the FC approximation at the CCSD(T) level. The second term is the correction to the FC calculation for the core-core and core-valence correlation. The next term is the correction from the CCSD(T) level to the CCSDT(Q) level and the last term accounts for relativistic effects. The correlation energies were extrapolated to the CBS limit using the X^{-3} scheme. The notation *method*/(*set*1,*set*2) indicates the theory level and basis sets used in the extrapolation. The SCF energy, which converges faster, was not extrapolated, but taken from the larger of the two basis sets.

The interaction energies at the near van der Waals minimum distance computed at the CCSD(T)/FC level can be found in Table I. The use of midbond functions obviously increases the rate of convergence, as does the extrapolation. For further comparison, explicitly correlated CCSD(T)-F12 calculations were performed using midbond functions.[36,37] The CCSD(T)-F12a and CCSD(T)-F12b triples energies were scaled proportionally to the ratio of the MP2-F12 and MP2 correlation energies,[38,39]

$$E_{int}^{(T**)-F12} = E_{int}^{(T)} \cdot \frac{E_{corr}^{MP2-F12}}{E_{corr}^{MP2}}. \quad (2)$$

The use of a double asterisk denotes that the scaling factor for the dimer was also employed for the counterpoise-corrected monomer calculations.[40,41] It should be noted that the CCSD(T)-F12 results were extrapolated using the same X^{-3} scheme despite their faster formal convergence. The comparison of the extrapolated scaled and unscaled CCSD(T)-F12a and CCSD(T)-F12b values shows high consistency for the largest basis set. The uncertainty for the extrapolated energies is taken as the difference between the extrapolated value

TABLE I. Frozen-core CCSD(T) interaction energies (cm^{-1}) at the near van der Waals minimum separation ($R = 4.0$ Å). The rows marked "ext." show the CBS-extrapolated results where the values for "X" are obtained from the $(X-1, X)$ extrapolation. The letter M in the basis set symbol stands for the hydrogenic set of midbond functions from the same aVXZ basis.

	$X =$			
Method	D	T	Q	5
$E_{int}^{CCSD(T)/FC}$/aVXZ	−17.459	−82.369	−113.474	−123.924
ext.		−112.000	−136.369	−134.764
$E_{int}^{CCSD(T)/FC}$/aVXZM	−69.585	−122.362	−129.610	−131.625
ext.		−144.339	−134.868	−133.726
$E_{int}^{CCSD(T)-F12a/FC}$/aV$X$ZM	−104.767	−130.423	−132.275	−132.899
ext.		−141.938	−133.697	−133.549
$E_{int}^{CCSD(T**)-F12a/FC}$/aV$X$ZM	−117.101	−136.819	−135.130	−134.365
ext.		−145.346	−133.909	−133.558
$E_{int}^{CCSD(T)-F12b/FC}$/aV$X$ZM	−95.504	−126.021	−129.366	−131.346
ext.		−139.095	−131.817	−133.421
$E_{int}^{CCSD(T**)-F12b/FC}$/aV$X$ZM	−107.839	−132.417	−132.221	−132.812
ext.		−141.450	−132.088	−133.429

TABLE II. Interaction energy corrections (cm^{-1}) at the near van der Waals minimum separation ($R = 4.0$ Å). The rows marked "ext." show the CBS-extrapolated results where the values for "X" are obtained from the $(X-1, X)$ extrapolation. The letter M in the basis set symbol stands for the hydrogenic set of midbond functions from the same aCVXZ/awCVXZ basis.

	X =			
Method	D	T	Q	5
$\Delta E_{int}^{CCSD(T)/AE-FC}$/aCV$XZ$	−2.288	−4.695	−4.613	−4.275
ext.		−5.708	−4.554	−3.921
$\Delta E_{int}^{CCSD(T)/AE-FC}$/aCV$XZ$M	−2.330	−4.254	−4.300	
ext.		−5.064	−4.334	
$\Delta E_{int}^{CCSD(T)/AE-FC}$/awCV$XZ$	−4.119	−5.057	−4.545	−4.216
ext.		−5.452	−4.171	−3.872
$\Delta E_{int}^{CCSD(T)/AE-FC}$/awCV$XZ$M	−3.814	−4.826	−4.201	
ext.		−5.252	−3.745	
$\Delta E_{int}^{T-(T)/FC}$/aV$XZ$	0.155	1.530	1.853	
ext.		2.109	2.088	
$\Delta E_{int}^{(Q)-T/FC}$/aV$XZ$	−0.814	−1.259	−1.745	
ext.		−1.447	−2.099	
$\Delta E_{int}^{(Q)-(T)/FC}$/aV$XZ$	−0.659	0.271	0.108	
ext.		0.662	−0.011	
ΔE_{int}^{rel}/decontracted aVXZ	−5.582	−4.507	−3.566	−3.196
ext.		−3.955	−2.894	−2.812

and the result in the larger basis set used for extrapolation. The CCSD(T**)-F12b/(aVQZM,aV5ZM) result exhibits the lowest uncertainty, yielding a value of -133.43 ± 0.62 cm^{-1} ($1\ cm^{-1} \approx 1.986\,45 \times 10^{-23}$ J).

Table II provides the values for the post-CCSD(T)/FC corrections at $R = 4.0$ Å. The core and relativistic terms are shown to be significant at the CBS limit, with values of $-3.87\ cm^{-1}$ and $-2.81\ cm^{-1}$, respectively. The coupled-cluster contributions beyond CCSD(T) are accidentally very small for this system (due to a nearly perfect cancellation between the CCSDT–CCSD(T) and CCSDT(Q)–CCSDT contributions); however, we still included them as they become somewhat more important at short range. The second-order Douglas-Kroll-Hess Hamiltonian[42,43] was used with the CCSD(T)/AE method and decontracted aVXZ basis sets to calculate the relativistic effects. Additionally, the contributions to the relativistic correction from the two-electron Breit-Pauli terms[44] were investigated (at $R = 4.06$ Å). Their contributions to the total interaction energy were found to be insignificant and we did not consider these terms any further. The relativistic correction to the SCF energy is not fully converged at large values of R, which is taken into consideration in the uncertainty estimation. The uncertainty of the relativistic terms is determined in the same manner as for the other terms, but with an addition of the difference between the SCF energies in the basis sets used in the extrapolation. The best correction estimates were added to the CCSD(T**)-F12b/(aVQZM,aV5ZM) value to provide the best estimate of the near-minimum interaction energy, -140.12 ± 0.82 cm^{-1}. The uncertainty of this estimate was determined by the quadratic addition of the uncertainties of the individual terms.

Using the best level of theory determined above, the *ab initio* calculations were performed at 25 distances R: 2.6, 2.8, 3.0, 3.2, 3.4, 3.6, 3.7, 3.8, 3.9, 4.0, 4.06, 4.1, 4.2, 4.3, 4.4, 4.6, 4.8, 5.0, 5.5, 6.0, 7.0, 8.0, 9.0, 10.0, and 12.0 Å. We observed that at the largest R, the CCSD(T**)-F12b interaction energy is not as well converged as the conventional CCSD(T) one. This behavior has been observed before and is likely due to both auxiliary basis set incompleteness and the residual inaccuracies of the approximate CCSD(T)-F12b approach.[45] As a result, we employed the conventional CCSD(T)/(aVQZM,aV5ZM) values for $R \geq 8.0$ Å. Table III contains the total interaction energy, total uncertainty, and interaction energy contributions at all values of R.

III. THE ANALYTIC KRYPTON-KRYPTON POTENTIAL

The *ab initio* data points obtained in Sec. II were fitted using a weighted least-squares routine to a function of the form

$$V(R) = \left(A + BR + \frac{C}{R}\right) e^{-\alpha R} - \sum_{n=3}^{4} f_{2n}(\beta R)\frac{C_{2n}}{R^{2n}}, \quad (3)$$

where $A, B, C, \alpha, \beta, C_6$, and C_8 are fit parameters and $f_{2n}(x)$ are the Tang-Toennies damping functions,[46]

$$f_{2n}(x) = 1 - e^{-x}\sum_{k=0}^{2n}\frac{x^k}{k!}. \quad (4)$$

For each R, the inverse of the squared uncertainty was used as the weight during fitting. The long-range terms C_6 and C_8 were fitted to points with $R \geq 8.0$ Å without any damping and then frozen for the rest of the fit. The parameter values obtained in the fitting process are shown in Table IV (one should note that these parameters are expressed in atomic units). Using the parameters provided, the potential passes through the $[E - U(E), E + U(E)]$ uncertainty range for each *ab initio* data point, deviating from the computed values by $0.16\,U(E)$ on the average. One should note that since only the two leading asymptotic terms are included in Eq. (3), the fitted

TABLE III. Contributions to the *ab initio* krypton dimer interaction energy, the total interaction energy, and the total uncertainty $U(E_{int})$ in cm^{-1}. The column marked "E_{fit}" shows the values for the analytical potential fitted to the *ab initio* data.

R (Å)	$E_{int}^{CCSD(T)/FC}$ [a]	$\Delta E_{int}^{CCSD(T)/AE-FC}$ [b]	$\Delta E_{int}^{(Q)-(T)/FC}$ [c]	ΔE_{int}^{rel} [d]	E_{int}	$U(E_{int})$ [e]	E_{fit}
2.60	10 593.159	−328.707	6.708	−468.592	9802.568	36.642	9798.985
2.80	5 199.108	−192.796	4.046	−259.852	4750.507	21.518	4755.736
3.00	2 396.447	−109.840	2.208	−139.063	2149.751	12.221	2149.689
3.20	989.473	−60.805	1.067	−71.849	857.886	6.869	856.318
3.40	315.487	−32.638	0.438	−35.682	247.605	3.861	246.681
3.60	15.182	−16.900	0.137	−16.846	−18.427	2.221	−18.612
3.70	−58.545	−11.965	0.063	−11.280	−81.727	1.675	−81.704
3.80	−101.762	−8.360	0.019	−7.375	−117.478	1.296	−117.349
3.90	−124.393	−5.747	−0.003	−4.666	−134.809	1.019	−134.654
4.00	−133.429	−3.872	−0.011	−2.812	−140.124	0.815	−139.999
4.06	−134.431	−3.016	−0.012	−2.002	−139.461	0.720	−139.366
4.10	−133.765	−2.538	−0.011	−1.562	−137.876	0.666	−137.802
4.20	−128.764	−1.603	−0.006	−0.737	−131.110	0.553	−131.093
4.30	−120.718	−0.954	0.001	−0.207	−121.878	0.467	−121.910
4.40	−111.157	−0.514	0.009	0.121	−111.542	0.398	−111.604
4.60	−91.093	−0.029	0.023	0.415	−90.684	0.298	−90.763
4.80	−72.746	0.161	0.033	0.464	−72.089	0.226	−72.153
5.00	−57.418	0.210	0.039	0.419	−56.751	0.174	−56.777
5.50	−31.587	0.159	0.041	0.245	−31.142	0.096	−31.112
6.00	−17.899	0.089	0.033	0.131	−17.646	0.059	−17.610
7.00	−6.547	0.027	0.018	0.040	−6.461	0.025	−6.458
8.00	−2.791	0.009	0.009	0.015	−2.757	0.018	−2.757
9.00	−1.329	0.004	0.005	0.007	−1.314	0.009	−1.315
10.00	−0.689	0.002	0.003	0.003	−0.682	0.004	−0.682
12.00	−0.224	0.000	0.001	0.001	−0.221	0.001	−0.221

[a]CCSD(T**)-F12b/(aVQZM,aV5ZM) for $R < 8.0$ Å. CCSD(T)/(aVQZM,aV5ZM) for $R \geq 8.0$Å.
[b]Extrapolated from the awCVQZ and awCV5Z basis sets.
[c]Extrapolated from the aVTZ and aVQZ basis sets.
[d]Extrapolated from CCSD(T)/AE (decontrated aVQZ, decontrated aV5Z) computed using second-order DKH Hamiltonian.
[e]Obtained by quadratically adding the uncertainties of the contributions.

parameters C_6 and C_8 should be viewed as effective constants that include some contribution from C_{10} and higher terms. In particular, our parameter C_6 = 126.79 a.u. is likely less accurate than the value 129.6 a.u. obtained from the experimental dipole oscillator strength distributions by Kumar and Meath.[47]

Without enforcing any particular short-range form, there is no guarantee that $V(R)$ will exhibit reasonable behavior at distances shorter than 2.6 Å (the lowest-R data point). Indeed, it turns out that the fitted $V(R)$ function behaves reasonably down to about 1.8 Å but becomes negative for still shorter R. To eliminate this unphysical behavior, we decided to replace $V(R)$ by a simpler, well-behaving function $V_{sh}(R)$ for $R < 1.8$ Å,

$$V_{sh}(R) = \left(\frac{A_{sh}}{R}\right) e^{-\alpha_{sh}R+\beta_{sh}R^2}, \tag{5}$$

where the parameter A_{sh} = (36 × 36) hartree×bohrs = 1.5052 × 10^8 cm^{-1} Å^6 enforces the correct $R \to 0$ asymptotics, and the parameters α_{sh} and β_{sh} are chosen such that the potential and its first derivative are continuous at the splicing point R = 1.8 Å. The values of the short-range parameters are also included in Table IV.

TABLE IV. Parameters of the analytical krypton dimer potential $V(R)$ in atomic units of energy and length (1 hartree ≈ 4.359 74 × 10^{-18} J, 1 bohr ≈ 5.291 772 × 10^{-11} m).

Parameter	Value	Unit
A	467.771 557	hartree
B	−43.111 875	hartree·bohrs^{-1}
C	−509.601 417	hartree·bohrs
α	1.566 575	bohrs^{-1}
β	4.083 794	bohrs^{-1}
C_6	126.790 499	hartree·bohrs^6
C_8	5268.109 217	hartree·bohrs^8
A_{sh}	1296.0	hartree·bohrs
α_{sh}	3.067 950	bohrs^{-1}
β_{sh}	0.324 071 4	bohrs^{-2}

The analytical potential has a minimum at 4.01 Å where the interaction energy is −140.06 cm^{-1}. This estimate of the minimum interaction energy agrees very well with the value of −139.9 cm^{-1} for the empirical potential of Dham *et al.*,[48] while the previous *ab initio* studies of Haley and Cybulski[49] (−132.8 cm^{-1}) and Slavíček *et al.*[50] (−135.1 cm^{-1}) predicted a somewhat shallower van der Waals well. The comparison of the potential energy functions of Refs. 48–50 with the function developed in this work is presented in Figure 1. The values shown are percent deviations of the literature potentials relative to the $V(R)$ of Eq. (3) (thus, the deviations are naturally very large around 3.58 Å where $V(R)$ crosses zero). To estimate the uncertainties of the potential and of the computed quantities, we have constructed the lower- and upper-limit potentials $V^-(R)$ and $V^+(R)$ by fitting the same functional

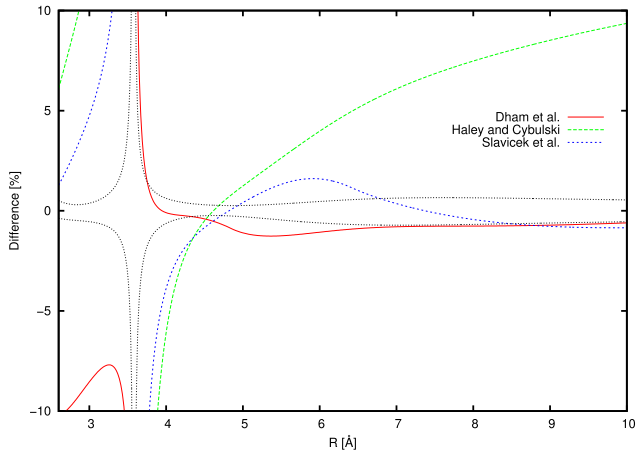


FIG. 1. Percent differences between the literature krypton-krypton potentials[48–50] and the potential developed in this work. The dotted black lines pertain to the lower- and upper-limit potentials $V^-(R)$ and $V^+(R)$.

form as Eq. (3) to $E_{int} - U(E_{int})$ and $E_{int} + U(E_{int})$, respectively. The deviations between these two potentials and $V(R)$ are also included in Figure 1. This figure shows that while the potentials of Haley and Cybulski[49] and Slavíček *et al.*[50] are quite far from our potential, the HFD-B2 function of Dham *et al.*[48] is either within or just outside our uncertainty range in the entire van der Waals minimum region. At shorter distances, all the literature potentials deviate from $V(R)$ by close to 10%.

As a first test of our analytic potential, we have calculated the vibrational transition frequencies for the most abundant $^{84}Kr_2$ isotopomer using a one-dimensional Schrödinger equation solver by Wolniewicz.[51] The potential supports 16 bound vibrational levels and the frequencies of transitions between the first 10 of them are presented in Table V. The uncertainty of each transition frequency was taken as the larger of the absolute differences between the results computed using $(V(R), V^+(R))$ and $(V(R), V^-(R))$. For comparison, Table V also lists the vibrational frequencies obtained by Dham *et al.*[48] (the HFD-B2 potential), Slaviček *et al.*[50] (their most accurate results, obtained with a relativistic effective core potential), and two experiments.[52,53] Our results are in a virtually perfect agreement with the empirical potential of Dham *et al.*, while the frequencies from the Slaviček *et al.* potential are somewhat underestimated. Our results are also within the (fairly wide) experimental error bars of Tanaka *et al.*[52] for all transitions except for the $5 \rightarrow 4$ one. The absolute agreement of our values with the two frequencies very precisely measured (for a different, $^{86}Kr_2$ isotopomer) by LaRocque *et al.*[53] is very good; both experimental values are within the theoretical uncertainty range. We conclude from Table V that our analytic potential gives vibrational transitions that are on par with the empirical potential by Dham *et al.*[48] and more accurate than the previous *ab initio* potential by Slaviček *et al.*[50]

IV. THE CALCULATION OF TRANSPORT PROPERTIES

The potential function fitted in Sec. III was used to compute the low-density transport properties of krypton in the classical kinetic theory by the following procedure.

A. Cross sections

First, a large number of cross sections $Q^{(l)}$ for collisions between two atoms are classically evaluated over a wide range of the collision energy E,[54]

$$Q^{(l)}(E) = 2\pi\left[1 - \frac{1+(-1)^l}{2(1+l)}\right]^{-1} \int_0^\infty (1 - \cos^l\theta)\, b\, db, \quad (6)$$

where l is the weighting parameter for molecular collisions and b the impact factor. The deflection angle θ is related to the potential energy function $V(R)$ and can be expressed as follows:

$$\theta(E,b) = \pi - 2b \int_{R_0}^\infty \frac{dR/R^2}{\sqrt{1 - b^2/R^2 - V(R)/E}} \quad (7)$$

in which R_0 represents the classical distance of closest approach during a collision and is determined from the formula

$$1 - b^2/R_0^2 - V(R_0)/E = 0. \quad (8)$$

B. Collision integrals

Next, the cross section $Q^{(l)}$ is integrated with respect to the collision energy E to derive the collision integral $\Omega^{(l,s)}$ as

TABLE V. Frequencies of vibrational transitions (in cm^{-1}) for the $^{84}Kr_2$ complex computed using the potential developed in this work compared to literature values.

Transition	This work	Empirical[48]	*Ab initio*[50,a]	Expt.[52]	Expt.[53,b]
$1 \rightarrow 0$	21.466 ± 0.081	21.41	21.02	21.56 ± 0.54	21.175 ± 0.010
$2 \rightarrow 1$	19.307 ± 0.077	19.30	18.86	19.09 ± 0.57	19.093 ± 0.020
$3 \rightarrow 2$	17.166 ± 0.073	17.20	16.73	16.76 ± 0.60	
$4 \rightarrow 3$	15.053 ± 0.069	15.11	14.62	14.76 ± 0.75	
$5 \rightarrow 4$	12.980 ± 0.065	13.02	12.56	12.23 ± 0.51	
$6 \rightarrow 5$	10.963 ± 0.062	10.97	10.56	10.49 ± 0.50	
$7 \rightarrow 6$	9.026 ± 0.060	9.01	8.65	8.92 ± 0.44	
$8 \rightarrow 7$	7.199 ± 0.063	7.17	6.86	6.92 ± 0.63	
$9 \rightarrow 8$	5.517 ± 0.064	5.49	5.19	5.54 ± 0.30	

[a]The ECP + aug-cc-pVQZ + *spdfg* frequencies computed using the average atomic weight of krypton at its natural isotopic composition.
[b]The transition frequency measured for the $^{86}Kr_2$ isotopomer. Our calculations for this isotopomer give 21.240 ± 0.080 and 19.131 ± 0.076 cm^{-1} for the $1 \rightarrow 0$ and $2 \rightarrow 1$ transitions, respectively.

a function of the temperature T,

$$\Omega^{(l,s)}(T) = \left[(s+1)!(kT)^{s+2}\right]^{-1} \int_0^\infty Q^{(l)}(E)\, e^{-E/kT} E^{s+1} dE, \tag{9}$$

where s is also the weighting parameter and k the Boltzmann constant ($1.380\,658 \times 10^{-23}$ J K^{-1}).

C. Transport properties

Finally, the transport properties for pure gases at low density are given by[55]

$$\eta = \frac{5}{16}(\pi m k T)^{1/2} \frac{f_\eta}{\Omega^{(2,2)}}, \tag{10}$$

$$\lambda = \frac{75}{64}(\pi k^3 T/m)^{1/2} \frac{f_\lambda}{\Omega^{(2,2)}}, \tag{11}$$

$$D = \frac{3}{8}(\pi k^3 T^3/m)^{1/2} \frac{f_D}{P\Omega^{(1,1)}}, \tag{12}$$

$$\alpha_T = \frac{15}{2} \frac{(6C^* - 5)(2A^* + 5)}{A^*(16A^* - 12B^* + 55)} (1 + \kappa_0), \tag{13}$$

where η represents the viscosity, λ the thermal conductivity, D the self-diffusion coefficient, α_T the thermal diffusion factor, m the atomic mass, and P the standard atmosphere (101.3 kPa). f_η and f_λ are the fifth-order approximations to the viscosity and thermal conductivity, respectively, which are given in detail in the Appendix of Viehland *et al.*[56] f_D and κ_0 denote the second-order corrections to the self-diffusion coefficient and thermal diffusion factor, respectively,[55]

$$f_D = 1 + \frac{1}{8}(6C^* - 5)^2 (2A^* + 5)^{-1}, \tag{14}$$

$$\kappa_0 = \frac{1}{9}(7 - 8E^*) \left\{ \frac{2A^*}{35/4 + 7A^* + 4F^*} \left\{ H^* + \frac{[A^*(7 - 8E^*) - 7(6C^* - 5)](35/8 + 28A^* - 6F^*)}{42A^*(2A^* + 5)} \right\} - \frac{5}{7}\left[H^* + \frac{7(6C^* - 5)}{5(2A^* + 5)} - \frac{3}{10}(7 - 8E^*) \right] \right\}. \tag{15}$$

The quantities A^*–H^* are derived by combining different collision integrals $\Omega^{(l,s)}$,

$$A^* = \Omega^{(2,2)}/\Omega^{(1,1)}, \tag{16}$$

$$B^* = \left(5\Omega^{(1,2)} - 4\Omega^{(1,3)}\right)/\Omega^{(1,1)}, \tag{17}$$

$$C^* = \Omega^{(1,2)}/\Omega^{(1,1)}, \tag{18}$$

$$E^* = \Omega^{(2,3)}/\Omega^{(2,2)}, \tag{19}$$

$$F^* = \Omega^{(3,3)}/\Omega^{(1,1)}, \tag{20}$$

$$H^* = (3B^* + 6C^* - 35/4)/(6C^* - 5). \tag{21}$$

V. RESULTS AND DISCUSSION

The transport properties of krypton were calculated over the temperature range between 116 and 5000 K. In this work, the classical calculations are valid because the ratio of the thermal de Broglie wavelength $h/\sqrt{2\pi m k T}$ to the atomic diameter σ(~4 Å) is much less than 1 for the complete temperature range. It should also be noted that 83.798 was used as the relative atomic mass for krypton. The uncertainty in a given transport property was estimated as half the absolute difference of properties computed with the lower- and upper-limit potentials, $V^-(R)$ and $V^+(R)$.

The experimental data from the literature were employed to check the quality of the *ab initio* values of transport properties. Since the data for the thermal diffusion factor have a comparably large uncertainty, the comparison for this property was not investigated in the present work. We also examined the performance of the empirical potential by Dham *et al.*[48] and the older *ab initio* potential by Slavíček *et al.*[50] for the transport properties, comparing them with the theoretical values from our new potential. It should be noted here that most measurements at low density were performed at or near to atmospheric pressure, whereas the viscosity and thermal conductivity were computed in the limit of zero density. At the temperatures near the normal boiling point of krypton, the effect of the change in density is 2%–3%, which has to be taken into account in the comparisons. Therefore, according to the Rainwater-Friend theory,[57,58] the data at zero density X_0 were derived from the density virial expansion of the low-density experimental data X in the first order,

$$X_0 = \frac{X}{1 + B_X \rho}, \tag{22}$$

$$B_X = N_A \sigma^3 B_X^*, \tag{23}$$

$$B_X^* = \sum_{i=0}^{n} a_i \left(\sqrt{T^*}\right)^{-i} \tag{24}$$

in which $X = \eta, \lambda$, $T^* = T/(\varepsilon/k)$, ρ is the mass density, and N_A is the Avogadro constant ($6.022\,141\,29 \times 10^{23}$ mol^{-1}). In Ref. 58, which used a Lennard-Jones description for the interatomic potential, the atomic diameter σ and the well depth ε/k were given in Table 5.2, whereas the resulting coefficients a_i were listed in Table 5.3.

A. Viscosity

A large body of experimental data for the viscosity of krypton can be found in the literature with different levels of accuracy. Here, we considered the data[59–63] recommended by Table I of Bich *et al.*[64] as well as some data[65–67] reported after 2000. The data were corrected using the Rainwater-Friend

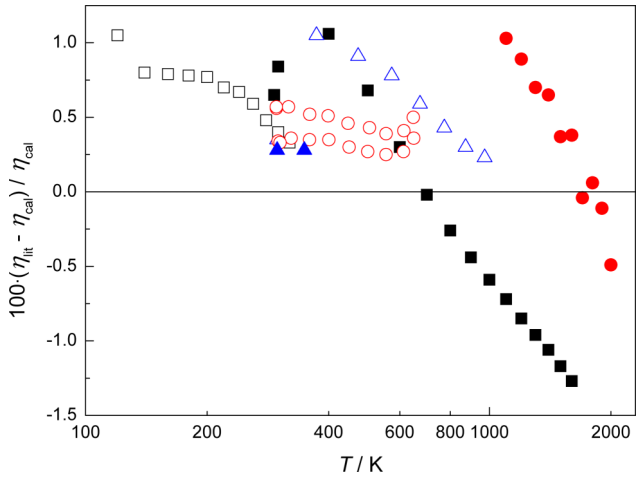


FIG. 2. Relative deviations of the literature viscosity data, $\eta_{\rm lit}$, from the values calculated in this work, $\eta_{\rm cal}$, for krypton. Data sources: (black filled square) Dawe and Smith;[59] (red filled circle) Goldblatt *et al.*;[60] (blue open triangle) Kestin *et al.*;[61] (black open square) Gough *et al.*;[62] (red open circle) Vogel;[63] and (blue filled triangle) Wilhelm and Vogel.[65]

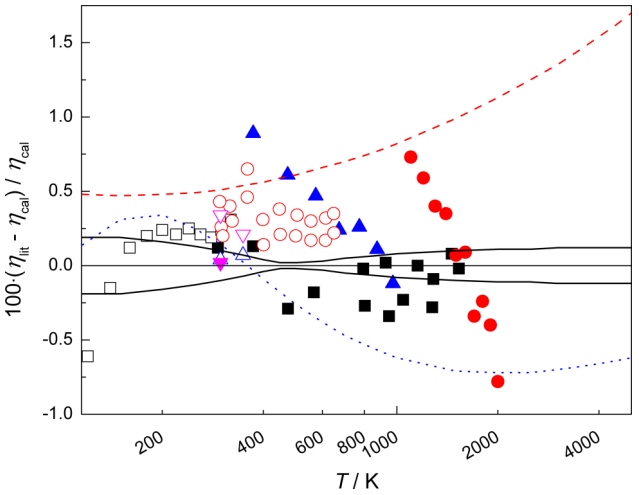


FIG. 3. Relative deviations of the literature viscosity data, $\eta_{\rm lit}$, from the values calculated in this work, $\eta_{\rm cal}$, for krypton. Data sources: (black filled square) Dawe and Smith,[59] re-evaluated data; (red filled circle) Goldblatt *et al.*,[60] re-evaluated data; (blue filled triangle) Kestin *et al.*,[61] re-evaluated data; (black open square) Gough *et al.*[62] re-evaluated data; (red open circle) Vogel,[63] re-evaluated data; (blue open triangle) Wilhelm and Vogel,[65] re-evaluated data; (magenta unfilled inverted triangle) Evers *et al.*;[66] (magenta filled inverted triangle) Berg and Burton;[67] (red dashed line) calculated using the potential by Dham *et al.*;[48] (blue dotted line) calculated using the potential by Slavíček *et al.*;[50] and (solid line) uncertainty of $\eta_{\rm cal}$.

theory except the zero-density data by Wilhelm and Vogel[65] as well as Berg and Burton.[67] Figure 2 reveals substantial disagreement between the viscosity data sets by relative measurements[59–63,65] and the theoretical values of this work. Thus, 40 from 66 data points deviate by more than ±0.4% from the calculated values. At room temperature, the differences fall into the interval of 0.3% to 0.7% and increase to −1.3% for temperatures up to 2000 K. Most of the viscosity data are not suitable for a reasonable comparison to judge the quality of the calculated values.

On the other hand, the recent data by Berg and Burton[67] based on a calibration with one *ab initio* helium value agree perfectly with the present theoretical result. The relative deviation at 25 °C gives the value of only 0.02% which is distinctly smaller than those for other relative measurements. The viscometers[59–63,65] were calibrated on older viscosity data which are characterized by larger uncertainties than those of recent *ab initio* results. Consequently, we re-evaluated the viscosity data of krypton to better accuracy by the following procedure. First, the ratio $\eta_{\rm kr}(T)/\eta_{\rm ref}(T)$ was calculated for the viscosity of krypton and one reference gas from the same paper. Next, the new value for the reference gas $\eta_{ab\ initio}(T)$ was interpolated from the *ab initio* results by Cencek *et al.*[1] for helium, Mehl[68] for argon, or Hellmann[69] for nitrogen. Finally, the new viscosity data for krypton were determined by the expression $\eta_{ab\ initio}\left(\frac{\eta_{\rm kr}}{\eta_{\rm ref}}\right)$. It should be stressed that this recalibration of the literature viscosity data includes no input from the present work (and no *ab initio* krypton data in general) and is simply an attempt to correct a known deficiency of the older experimental data, which has been eliminated in the newer ones.[67]

In Figure 3, the re-evaluated data give considerably better agreement with the theoretical results than the original data. The viscosity data by Dawe and Smith[59] are characterized by deviations of ±(0%–0.3%) over a wide temperature range between 290 and 1500 K. The measurements by Goldblatt *et al.*[60] at high temperatures result in deviations from +0.7% at 1100 K down to −0.8% at 2000 K. The measurement by Kestin *et al.*[61] at 25 °C differs by +0.2% from the value calculated using the potential developed in this work. For the data recommended by Kestin *et al.* above the ambient temperature, the differences decrease with increasing temperature from +0.9% at 100 °C to −0.1% at 700 °C. The low-temperature measurements by Gough *et al.*[62] lead to results which agree within about ±0.3% over all temperatures apart from the lowest one. In the group of Vogel, measurements on krypton were carried out by means of oscillating-disk and vibrating-wire viscometers. The data by Vogel[63] in 1984 deviate by +(0.1%–0.7%), whereas that by Wilhelm and Vogel[65] in 2000 are only 0.04% and 0.07% higher than the theoretical values at 25 °C and 75 °C, respectively. Evers *et al.*[66] performed absolute measurements based on the rotating-cylinder method and obtained results which deviate by +0.3% at 298 K and +0.2% at 348 K.

Figure 3 additionally includes the comparison between the values computed from the potentials by Dham *et al.*[48] and Slavíček *et al.*[50] and the theoretical results of this work. The values for the two literature potentials show in most cases larger deviations than those for the experimental data, especially for temperatures above and below ambient temperature. The differences demonstrate that the interatomic potential developed in this work is more accurate than the empirical potential by Dham *et al.* and the older *ab initio* potential by Slavíček *et al.*

B. Thermal conductivity

As in the case of the viscosity, there is a large amount of experimental thermal conductivity data available in the literature for krypton. Figure 4 contains the percentage deviations of the data[70–72] collected in Table II of Bich *et al.*[64] as well as by two other sources[73,74] from the values calculated using

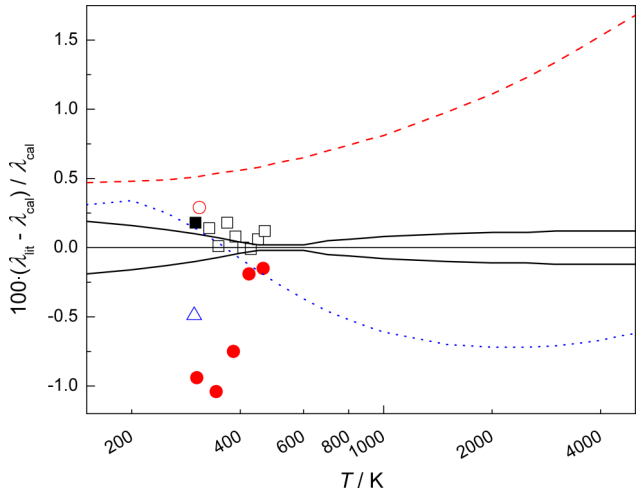


FIG. 4. Relative deviations of the literature thermal conductivity data, $\lambda_{\rm lit}$, from the values calculated in this work, $\lambda_{\rm cal}$, for krypton. Data sources: (black open square) Haarman;[70] (black filled square) Kestin *et al.*;[71] (red open circle) Assael *et al.*;[72] (red filled circle) Hemminger;[73] (blue open triangle) Le Neindre *et al.*;[74] (red dashed line) calculated using the potential by Dham *et al.*;[48] (blue dotted line) calculated using the potential by Slavíček *et al.*;[50] and (solid line) uncertainty of $\lambda_{\rm cal}$.

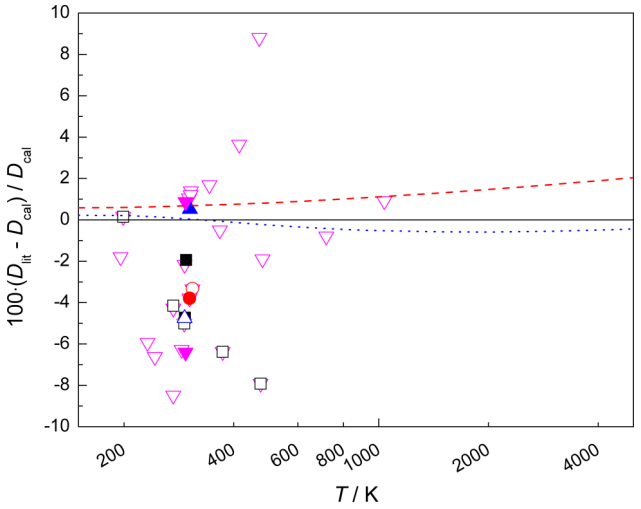


FIG. 5. Relative deviations of the literature self-diffusion coefficient data, $D_{\rm lit}$, from the values calculated in this work, $D_{\rm cal}$, for krypton. Data sources: (black filled square) Groth and Harteck;[75] (black open square) Schäfer and Schuhmann;[76] (red open circle) Durbin and Kobayashi;[77] (blue open triangle) Miller and Carman;[78] (red filled circle) Watts;[79] (blue filled triangle) Saran and Singh;[80] (magenta filled inverted triangle) Annis *et al.*;[81] (magenta unfilled inverted triangle) Weissman and DuBro;[82] (red dashed line) calculated using the potential by Dham *et al.*;[48] and (blue dotted line) calculated using the potential by Slavíček *et al.*[50] The small uncertainty of $D_{\rm cal}$ was omitted for clarity of the figure.

the potential constructed in this work. The corrections for the density effect were applied to three data sets,[70,73,74] while the data by Kestin *et al.*[71] and Assael *et al.*[72] were presented at zero density. The transient hot-wire technique was employed in the experimental studies by Haarman,[70] Kestin *et al.*[71] and Assael *et al.*[72] The data by Haarman between 328 and 468 K are characterized by deviations of ±(0%–0.2%). At ambient temperature, the differences are +0.2% and +0.3% from the calculated values for the data by Kestin *et al.* and Assael *et al.*, respectively. The data by Hemminger[73] differ by −0.9% at 30 °C and the deviations decrease to −0.2% at 190 °C, whereas the data by Le Neindre *et al.*[74] at 25 °C are consistent with the theoretical value within −0.5%.

C. Self-diffusion coefficient

In Figure 5, the self-diffusion coefficient data by Groth and Harteck[75] differ from the present calculated values by up to −5% around room temperature. The experimental results by Schäfer and Schuhmann[76] show deviations between +0.2% at 200 K and −8% at 470 K. The two-bulb method was employed by different researchers[77–82] to measure the self-diffusion coefficient of krypton. For the temperature range from 20 to 35 °C, the data in three papers[77–79] deviate from the calculated values by −(3%–5%), whereas a good agreement within +0.5% is found for the result by Saran and Singh.[80] Annis *et al.*[81] reported two values at the same temperature of 295 K which are characterized by deviations of +0.9% and −6%, respectively. Weissman and DuBro[82] obtained a total of 25 data covering a wide temperature range from about 200 to 1000 K with differences of ±(0.2%–9%) from the theoretically calculated values.

It should be noted that the thermal conductivity and self-diffusion coefficient usually have larger uncertainties than the viscosity. As mentioned above, the careful measurements by Roder *et al.*[22] indicated that the accuracy reached the level of 1%–2% for the low-density thermal conductivity of argon around room temperature. In a critical review by Marrero and Mason,[83] the uncertainty of the experimental self-diffusion coefficient for dilute krypton was estimated to be 1% at room temperature and increase to 5% at 1000 K. It is obvious that the comparably large deviations for the experimental data in Figures 4 and 5 do not allow one to distinguish between the empirical potential by Dham *et al.*,[48] the older *ab initio* potential by Slavíček *et al.*,[50] and our new *ab initio* potential.

VI. CONCLUSIONS

We have developed a new potential energy curve for the krypton dimer in a form of an analytical function, fitted to 25 *ab initio* data points, that exhibits quantitatively correct behavior in the limit $R \to \infty$ and qualitatively correct behavior for $R \to 0$. The *ab initio* interaction energies were obtained from explicitly correlated coupled-cluster calculations employing basis sets as large as aV5ZM (augmented correlation-consistent quintuple zeta with midbond functions) and extrapolated to the CBS limit. Moreover, the frozen-core CCSD(T)-level values were supplemented by corrections for core-core and core-valence correlation, higher-order coupled-cluster excitations up to CCSDT(Q), and relativistic effects. The total interaction energy at a near-minimum interatomic distance of 4.0 Å amounts to -140.12 ± 0.82 cm^{-1}.

The new krypton-krypton potential was subsequently used to compute several transport properties of krypton using the classical kinetic theory of dilute gases. The viscosity, thermal conductivity, self-diffusion coefficient, and thermal diffusion factor were studied in this work at low density over

TABLE VI. Low-density transport properties of krypton and their estimated uncertainties.

T/K	η/μPa s	λ/mW m^{-1} K^{-1}	D (101.3 kPa)/cm^2 s^{-1}	α_T
115.78	10.188(19)	3.7916(71)	0.015 453(36)	0.062 346(1314)
150	13.067(24)	4.8624(91)	0.025 753(53)	0.043 669(2083)
200	17.313(28)	6.4420(105)	0.045 376(75)	0.072 753(2845)
209.48	18.110(28)	6.7385(106)	0.049 677(78)	0.081 697(2938)
250	21.455(28)	7.9838(104)	0.070 029(89)	0.124 17(319)
273.15	23.314(26)	8.6763(99)	0.083 027(92)	0.149 31(325)
273.16	23.315(26)	8.6766(99)	0.083 033(92)	0.149 32(325)
293.15	24.887(25)	9.2623(94)	0.095 021(93)	0.170 64(327)
298.15	25.275(25)	9.4070(92)	0.098 127(93)	0.175 88(327)
300	25.419(24)	9.4604(92)	0.099 287(93)	0.177 80(327)
350	29.186(19)	10.865(7)	0.132 79(9)	0.226 97(320)
400	32.764(13)	12.200(5)	0.170 24(7)	0.269 90(305)
450	36.167(7)	13.471(3)	0.211 40(5)	0.306 65(287)
500	39.415(8)	14.684(3)	0.256 08(5)	0.337 84(267)
600	45.509(12)	16.963(4)	0.355 36(7)	0.386 66(230)
700	51.166(24)	19.079(9)	0.467 11(17)	0.421 89(197)
800	56.476(35)	21.066(13)	0.590 57(30)	0.447 62(169)
900	61.504(45)	22.947(16)	0.725 20(45)	0.466 65(146)
1000	66.300(54)	24.742(20)	0.870 54(61)	0.480 88(127)
1500	87.860(91)	32.806(34)	1.746 8(16)	0.514 16(73)
2000	106.85(12)	39.905(44)	2.852 4(30)	0.521 51(53)
2500	124.27(14)	46.412(53)	4.169 4(47)	0.520 59(46)
3000	140.59(16)	52.508(60)	5.685 7(66)	0.516 58(46)
3500	156.09(18)	58.294(68)	7.392 3(90)	0.511 36(49)
4000	170.94(20)	63.837(75)	9.282 2(116)	0.505 69(54)
4500	185.26(22)	69.181(82)	11.350(15)	0.499 91(59)
5000	199.14(24)	74.359(89)	13.590(18)	0.494 20(65)

the temperature range between 116 and 5000 K. The experimental data in the literature and the values computed by other krypton-krypton potentials were compared with the present theoretical values for the different transport properties. The experimental viscosity data re-evaluated using new *ab initio* reference values as well as the data of Berg and Burton[67] make it evident that the new *ab initio* potential developed in this work should be more accurate than the empirical potential by Dham *et al.*[48] and the older *ab initio* potential by Slavíček *et al.*[50] However, the comparably large uncertainties for the thermal conductivity and self-diffusion coefficient in the low-pressure region indicate that the two properties are not suitable to judge the quality of the different potentials. The careful analysis shows that the calculated values for the transport properties considered (Table VI) could be applied as standard values over the complete temperature range of the calculations.

ACKNOWLEDGMENTS

B.S. and X.W. are supported by the National Natural Science Foundation of China (Grant No. 51406154). J.W. and K.P. are supported by the U.S. National Science Foundation CAREER Award No. CHE-1351978.

[1]W. Cencek, M. Przybytek, J. Komasa, J. B. Mehl, B. Jeziorski, and K. Szalewicz, J. Chem. Phys. **136**, 224303 (2012).
[2]M. Przybytek, W. Cencek, J. Komasa, G. Łach, B. Jeziorski, and K. Szalewicz, Phys. Rev. Lett. **104**, 183003 (2010).
[3]J. W. Schmidt, R. M. Gavioso, E. F. May, and M. R. Moldover, Phys. Rev. Lett. **98**, 254504 (2007).
[4]J. M. L. Martin, Chem. Phys. Lett. **303**, 399 (1999).
[5]K. Patkowski, R. Podeszwa, and K. Szalewicz, J. Phys. Chem. A **111**, 12822 (2007).
[6]J. M. Merritt, V. E. Bondybey, and M. C. Heaven, Science **324**, 1548 (2009).
[7]K. Patkowski, V. Špirko, and K. Szalewicz, Science **326**, 1382 (2009).
[8]J. Koput, Phys. Chem. Chem. Phys. **13**, 20311 (2011).
[9]S. Sharma, T. Yanai, G. H. Booth, C. J. Umrigar, and G. K. Chan, J. Chem. Phys. **140**, 104112 (2014).
[10]M. Lesiuk, M. Przybytek, M. Musiał, B. Jeziorski, and R. Moszynski, Phys. Rev. A **91**, 012510 (2015).
[11]P. Jankowski, A. R. W. McKellar, and K. Szalewicz, Science **336**, 1147 (2012).
[12]P. Jankowski, L. A. Surin, A. Potapov, S. Schlemmer, A. R. W. McKellar, and K. Szalewicz, J. Chem. Phys. **138**, 084307 (2013).
[13]R. Hellmann, E. Bich, and E. Vogel, Mol. Phys. **106**, 133 (2008).
[14]E. Bich, R. Hellmann, and E. Vogel, Mol. Phys. **106**, 813 (2008).
[15]E. Vogel, B. Jäger, R. Hellmann, and E. Bich, Mol. Phys. **108**, 3335 (2010).
[16]B. Song, X. P. Wang, and Z. G. Liu, Mol. Simul. **2015**, 1.
[17]B. Jäger, R. Hellmann, E. Bich, and E. Vogel, Mol. Phys. **107**, 2181 (2009).
[18]B. Jäger, R. Hellmann, E. Bich, and E. Vogel, Mol. Phys. **108**, 105 (2010).
[19]K. Patkowski and K. Szalewicz, J. Chem. Phys. **133**, 094304 (2010).
[20]E. Vogel, Int. J. Thermophys. **31**, 447 (2010).
[21]R. F. Berg and M. R. Moldover, J. Phys. Chem. Ref. Data **41**, 043104 (2012).
[22]H. M. Roder, R. A. Perkins, A. Laesecke, and C. A. Nieto de Castro, J. Res. Natl. Inst. Stand. Technol. **105**, 221 (2000).
[23]M. J. Assael, K. D. Antoniadis, and W. A. Wakeham, Int. J. Thermophys. **31**, 1051 (2010).
[24]M. J. Assael, J.-A. M. Assael, M. L. Huber, R. A. Perkins, and Y. Takata, J. Phys. Chem. Ref. Data **40**, 033101 (2011).
[25]Y. J. Bomble, J. F. Stanton, M. Kállay, and J. Gauss, J. Chem. Phys. **123**, 054101 (2005).
[26]S. A. Kucharski and R. J. Bartlett, J. Chem. Phys. **108**, 9221 (1998).
[27]H.-J. Werner *et al.*, MOLPRO, version 2012.1, a package of *ab initio* programs, 2012, see http://www.molpro.net.
[28]J. Noga and R. J. Bartlett, J. Chem. Phys. **86**, 7041 (1987).
[29]N. Oliphant and L. Adamowicz, J. Chem. Phys. **94**, 1229 (1991).
[30]S. A. Kucharski and R. J. Bartlett, Theor. Chim. Acta **80**, 387 (1991).

[31]M. Kállay and P. R. Surján, J. Chem. Phys. **115**, 2945 (2001).
[32]T. H. Dunning, Jr., J. Chem. Phys. **90**, 1007 (1989).
[33]R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, J. Chem. Phys. **96**, 6796 (1992).
[34]A. K. Wilson, D. E. Woon, K. A. Peterson, and T. H. Dunning, Jr., J. Chem. Phys. **110**, 7667 (1999).
[35]S. F. Boys and F. Bernardi, Mol. Phys. **19**, 553 (1970).
[36]C. Hättig, W. Klopper, A. Köhn, and D. P. Tew, Chem. Rev. **112**, 4 (2012).
[37]L. Kong, F. A. Bischoff, and E. F. Valeev, Chem. Rev. **112**, 75 (2012).
[38]T. B. Adler, G. Knizia, and H.-J. Werner, J. Chem. Phys. **127**, 221106 (2007).
[39]G. Knizia, T. B. Adler, and H.-J. Werner, J. Chem. Phys. **130**, 054104 (2009).
[40]O. Marchetti and H.-J. Werner, J. Phys. Chem. A **113**, 11580 (2009).
[41]M. S. Marshall and C. D. Sherrill, J. Chem. Theory Comput. **7**, 3978 (2011).
[42]M. Douglas and N. M. Kroll, Ann. Phys. **82**, 89 (1974).
[43]B. A. Hess, Phys. Rev. A **33**, 3742 (1986).
[44]S. Coriani, T. Helgaker, P. Jørgensen, and W. Klopper, J. Chem. Phys. **121**, 6591 (2004).
[45]K. Patkowski, J. Chem. Phys. **137**, 034103 (2012).
[46]K. T. Tang and J. P. Toennies, J. Chem. Phys. **80**, 3726 (1984).
[47]A. Kumar and W. J. Meath, Mol. Phys. **54**, 823 (1985).
[48]A. K. Dham, A. R. Allnatt, W. J. Meath, and R. A. Aziz, Mol. Phys. **67**, 1291 (1989).
[49]T. P. Haley and S. M. Cybulski, J. Chem. Phys. **119**, 5487 (2003).
[50]P. Slavíček, R. Kalus, P. Paška, I. Odvárková, P. Hobza, and A. Malijevský, J. Chem. Phys. **119**, 2102 (2003).
[51]L. Wolniewicz, J. Chem. Phys. **45**, 515 (1966).
[52]Y. Tanaka, K. Yoshino, and D. E. Freeman, J. Chem. Phys. **59**, 5160 (1973).
[53]P. E. LaRocque, R. H. Lipson, P. R. Herman, and B. P. Stoicheff, J. Chem. Phys. **84**, 6627 (1986).
[54]G. C. Maitland, M. Rigby, E. B. Smith, and W. A. Wakeham, *Intermolecular Forces: Their Origin and Determination* (Clarendon Press, Oxford, 1981).
[55]J. Kestin, K. Knierim, E. A. Mason, B. Najafi, S. T. Ro, and M. Waldman, J. Phys. Chem. Ref. Data **13**, 229 (1984).
[56]L. A. Viehland, A. R. Janzen, and R. A. Aziz, J. Chem. Phys. **102**, 5444 (1995).
[57]J. C. Rainwater and D. G. Friend, Phys. Rev. A **36**, 4062 (1987).
[58]E. Bich and E. Vogel, in *Transport Properties of Fluids: Their Correlation, Prediction, and Estimation*, edited by J. Millat, J. H. Dymond, and C. A. Nieto de Castro (Cambridge University Press, Cambridge, 1996), p. 72.
[59]R. A. Dawe and E. B. Smith, J. Chem. Phys. **52**, 693 (1970).
[60]M. Goldblatt, F. A. Guevara, and B. B. McInteer, Phys. Fluids **13**, 2873 (1970).
[61]J. Kestin, W. A. Wakeham, and S. T. Ro, J. Chem. Phys. **56**, 4119 (1972).
[62]D. W. Gough, G. P. Matthews, and E. B. Smith, J. Chem. Soc., Faraday Trans. 1 **72**, 645 (1976).
[63]E. Vogel, Ber. Bunsenges. Phys. Chem. **88**, 997 (1984).
[64]E. Bich, J. Millat, and E. Vogel, J. Phys. Chem. Ref. Data **19**, 1289 (1990).
[65]J. Wilhelm and E. Vogel, Int. J. Thermophys. **21**, 301 (2000).
[66]C. Evers, H. W. Lösch, and W. Wagner, Int. J. Thermophys. **23**, 1411 (2002).
[67]R. F. Berg and W. C. Burton, Mol. Phys. **111**, 195 (2013).
[68]J. B. Mehl, private communication (2012).
[69]R. Hellmann, Mol. Phys. **111**, 387 (2013).
[70]J. W. Haarman, AIP Conf. Proc. **11**, 193 (1973).
[71]J. Kestin, R. Paul, A. A. Clifford, and W. A. Wakeham, Physica A **100**, 349 (1980).
[72]M. J. Assael, M. Dix, A. Lucas, and W. A. Wakeham, J. Chem. Soc., Faraday Trans. 1 **77**, 439 (1981).
[73]W. Hemminger, Int. J. Thermophys. **8**, 317 (1987).
[74]B. Le Neindre, Y. Garrabos, and R. Tufeu, Physica A **156**, 512 (1989).
[75]W. Groth and P. Harteck, Z. Elektrochem. **47**, 167 (1941), available at http://onlinelibrary.wiley.com/doi/10.1002/bbpc.19410470219/abstract.
[76]K. Schäfer and K. Schuhmann, Z. Elektrochem. **61**, 246 (1957), available at http://onlinelibrary.wiley.com/doi/10.1002/bbpc.19570610205/abstract.
[77]L. Durbin and R. Kobayashi, J. Chem. Phys. **37**, 1643 (1962).
[78]L. Miller and P. C. Carman, Trans. Faraday Soc. **60**, 33 (1964).
[79]H. Watts, Trans. Faraday Soc. **60**, 1745 (1964).
[80]A. Saran and Y. Singh, Can. J. Chem. **44**, 2222 (1966).
[81]B. K. Annis, A. E. Humphreys, and E. A. Mason, Phys. Fluids **12**, 78 (1969).
[82]S. Weissman and G. A. DuBro, Phys. Fluids **13**, 2689 (1970).
[83]T. R. Marrero and E. A. Mason, J. Phys. Chem. Ref. Data **1**, 3 (1972).

Appendix B

Accurate virial coefficients of gaseous krypton from state-of-the-art *ab initio* potential and polarizability of the krypton dimer

THE JOURNAL OF CHEMICAL PHYSICS **148**, 024306 (2018)

Accurate virial coefficients of gaseous krypton from state-of-the-art *ab initio* potential and polarizability of the krypton dimer

Bo Song,[1,a),b)] Jonathan M. Waldrop,[2,a)] Xiaopo Wang,[1] and Konrad Patkowski[2,c)]
[1]*Key Laboratory of Thermo-Fluid Science and Engineering, Ministry of Education, School of Energy and Power Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China*
[2]*Department of Chemistry and Biochemistry, Auburn University, Auburn, Alabama 36849, USA*

(Received 29 September 2017; accepted 22 December 2017; published online 11 January 2018)

We have developed a new krypton–krypton interaction-induced isotropic dipole polarizability curve based on high-level *ab initio* methods. The determination was carried out using the coupled-cluster singles and doubles plus perturbative triples method with very large basis sets up to augmented correlation-consistent sextuple zeta as well as the corrections for core-core and core-valence correlation and relativistic effects. The analytical function of polarizability and our recently constructed reference interatomic potential [J. M. Waldrop *et al.*, J. Chem. Phys. **142**, 204307 (2015)] were used to predict the thermophysical and electromagnetic properties of krypton gas. The second pressure, acoustic, and dielectric virial coefficients were computed for the temperature range of 116 K–5000 K using classical statistical mechanics supplemented with high-order quantum corrections. The virial coefficients calculated were compared with the generally less precise available experimental data as well as with values computed from other potentials in the literature {in particular, the recent highly accurate potential of Jäger *et al.* [J. Chem. Phys. **144**, 114304 (2016)]}. The detailed examination in this work suggests that the present theoretical prediction can be applied as reference values in disciplines involving thermophysical and electromagnetic properties of krypton gas. *Published by AIP Publishing.* https://doi.org/10.1063/1.5006970

I. INTRODUCTION

The macroscopic properties of fluids depend on the microscopic interactions between molecules and in principle they can be calculated from them.[1–4] In this way, with the help of modern computational techniques, it is now possible to obtain thermophysical properties of gases to high accuracy purely from theory.[5–12] This process involves two lengthy computations. At first, the potential energy curve $V(R)$, where R is the interatomic distance, is computed from first principles. Then, $V(R)$ is employed in expressions based on statistical mechanics and kinetic theory of dilute gases to yield thermophysical properties. The best known example is the theoretical results for helium[9] which have uncertainties that are significantly smaller (sometimes by nearly two orders of magnitude) than those of the corresponding measured quantities. Standard theoretical values of thermophysical properties are becoming important for two areas: instrument calibration[13–15] and metrology.[16–18] Other studies such as the development of thermodynamic and transport equations could also benefit from more accurate thermophysical data (e.g., Refs. 19–21).

We have recently constructed a new pair potential for krypton using coupled-cluster calculations up to the singles, doubles, triples, and perturbative quadruples level, including the corrections for core-core and core-valence correlation as well as relativistic effects.[10] Comparisons with the literature indicated that this potential is more accurate than any curve published previously. In this work, we have used the new potential together with classical statistical mechanics with quantum corrections to compute some thermophysical properties of krypton gas: the second pressure virial coefficient B and the second acoustic virial coefficient β_a for the range of 115.78 K (the triple-point temperature of naturally abundant krypton[22]) to 5000 K. The range of temperatures investigated is wide enough for most scientific and engineering purposes. In addition to thermophysical properties, the electromagnetic properties of simple gases are also valuable for pressure and temperature metrology.[23] For these reasons, we have also calculated the second dielectric virial coefficient B_ε of krypton using a semi-classical method for the same aforementioned temperature range. The calculations of B_ε require not only the pair potential but also the interaction-induced isotropic dipole polarizability $\Delta\alpha_{ave}(R)$ of the weakly bound dimer.

The rest of the article is structured as follows. We present in Sec. II the determination of the krypton–krypton interaction-induced polarizability as well as its analytical representation $\Delta\alpha_{ave}^{fit}(R)$ and the expressions for the lower- and upper-limits of its values. In Sec. III, our recent interatomic potential of the weakly bound krypton dimer[10] is briefly described for the sake of completeness. Section IV provides the detailed formulations for the computation of virial coefficients of krypton gas and the procedure to evaluate the uncertainty of theoretically calculated values. Section V contains a comparison of the experimental data and the values from some literature potentials with the present calculated values for the virial coefficients of krypton to assess carefully the performance of this work. It

a)B. Song and J. M. Waldrop contributed equally to this work.
b)E-mail: song.bo@xjtu.edu.cn
c)E-mail: patkowsk@auburn.edu

should be noted that, very recently, the Rostock group reported their fourth reference potential for pure noble gas molecules, namely, one for krypton,[12] following ones for helium,[6] neon,[7] and argon.[8] Jäger *et al.* used the best *ab initio* methods suitable for krypton to develop the potential energy curve for the two-body as well as the three-body interaction between krypton atoms. In Sec. V, considerable attention is paid to the agreement between the new *ab initio* potentials of Jäger *et al.*[12] and our work for the virial coefficients of krypton gas. Finally, we summarize in Sec. VI the main conclusions and give some perspectives of this work.

II. DEVELOPMENT OF INTERACTION-INDUCED POLARIZABILITY

The interaction-induced isotropic pair polarizability (the trace of the interaction-induced polarizability tensor) $\Delta\alpha_{\rm ave}(R)$ and its anisotropy $\Delta\alpha_{\rm aniso}(R)$ of the krypton dimer were calculated for the same interatomic distances as those used in the development of our high-accuracy Kr–Kr potential energy curve.[10] All calculations employed the counterpoise correction for the basis set superposition error.[24] Specifically, $\Delta\alpha_{\rm ave}(R)$ and $\Delta\alpha_{\rm aniso}(R)$ are defined by

$$\Delta\alpha_{\rm ave}(R) = \frac{1}{3}\left[\Delta\alpha_{\parallel}(R) + 2\Delta\alpha_{\perp}(R)\right], \tag{1}$$

$$\Delta\alpha_{\rm aniso}(R) = \left[\Delta\alpha_{\parallel}(R) - \Delta\alpha_{\perp}(R)\right], \tag{2}$$

$$\Delta\alpha_{\parallel}(R) = \alpha_{\parallel}^{\rm dimer}(R) - 2\alpha_{\parallel}^{\rm monomer}(R), \tag{3}$$

$$\Delta\alpha_{\perp}(R) = \alpha_{\perp}^{\rm dimer}(R) - 2\alpha_{\perp}^{\rm monomer}(R), \tag{4}$$

where $\alpha_{\parallel}^{\rm dimer}(R)$ and $\alpha_{\perp}^{\rm dimer}(R)$ are the components of the dimer polarizability that are parallel and perpendicular, respectively, to the internuclear axis at a given interatomic distance R and $\alpha_{\parallel}^{\rm monomer}(R)$ and $\alpha_{\perp}^{\rm monomer}(R)$ are the corresponding monomer components of polarizability calculated in the dimer basis.[8] All polarizabilities were calculated at the coupled-cluster singles and doubles plus perturbative triples [CCSD(T)] level and include corrections for core-core and core-valence correlation ($\Delta\alpha_{\rm ave}^{\rm AE-FC}$, $\Delta\alpha_{\rm aniso}^{\rm AE-FC}$) and relativistic effects ($\Delta\alpha_{\rm ave}^{\rm rel}$, $\Delta\alpha_{\rm aniso}^{\rm rel}$) so that

$$\Delta\alpha_{\rm ave} = \Delta\alpha_{\rm ave}^{\rm CCSD(T)/FC} + \Delta\alpha_{\rm ave}^{\rm AE-FC} + \Delta\alpha_{\rm ave}^{\rm rel}, \tag{5}$$

$$\Delta\alpha_{\rm aniso} = \Delta\alpha_{\rm aniso}^{\rm CCSD(T)/FC} + \Delta\alpha_{\rm aniso}^{\rm AE-FC} + \Delta\alpha_{\rm aniso}^{\rm rel}. \tag{6}$$

Calculations were performed using the CFOUR and MOLPRO codes.[25–27] Augmented correlation-consistent Dunning basis sets with X = D, T, Q, 5, 6 were used in all calculations: they included the polarized valence series,

TABLE I. Frozen-core CCSD(T) components of the interaction-induced pair polarizability and all-electron and relativistic corrections (a_0^3) for the krypton dimer near van der Waals minimum (R = 4.0 Å), computed in augmented Dunning basis sets aVXZ/aCVXZ/. . . as indicated.

	X =				
Component	D	T	Q	5	6
$\alpha_{\parallel}^{\rm CCSD(T)/FC}$/aV$X$Z	1.9897	2.5673	2.6037	2.6195	2.6246
$\alpha_{\perp}^{\rm CCSD(T)/FC}$/aV$X$Z	−0.9923	−1.2590	−1.2445	−1.2380	−1.2374
$\alpha_{\parallel}^{\rm CCSD(T)/FC}$/aCV$X$Z	2.0051	2.5805	2.6031	2.6192	
$\alpha_{\perp}^{\rm CCSD(T)/FC}$/aCV$X$Z	−0.9991	−1.2611	−1.2440	−1.2377	
$\alpha_{\parallel}^{\rm CCSD(T)/FC}$/awCV$X$Z	2.0440	2.5689	2.5998	2.6177	
$\alpha_{\perp}^{\rm CCSD(T)/FC}$/awCV$X$Z	−1.0144	−1.2564	−1.2421	−1.2369	
$\alpha_{\parallel}^{\rm AE-FC}$/aV$X$Z	−0.0030	−0.0183	−0.0246	−0.0410	
$\alpha_{\perp}^{\rm AE-FC}$/aV$X$Z	0.0015	0.0091	0.0121	0.0193	
$\alpha_{\parallel}^{\rm AE-FC}$/aCV$X$Z	−0.0077	−0.0357	−0.0472	−0.0559	
$\alpha_{\perp}^{\rm AE-FC}$/aCV$X$Z	0.0042	0.0185	0.0240	0.0277	
$\alpha_{\parallel}^{\rm AE-FC}$/awCV$X$Z	−0.0237	−0.0439	−0.0541	−0.0596	
$\alpha_{\perp}^{\rm AE-FC}$/awCV$X$Z	0.0120	0.0227	0.0272	0.0294	
$\alpha_{\parallel}^{\rm Rel}$/aV$X$Z	−0.0469	−0.0167	−0.0160	−0.0165	
$\alpha_{\perp}^{\rm Rel}$/aVXZ	0.0251	0.0128	0.0143	0.0127	
$\alpha_{\parallel}^{\rm Rel}$/aCV$X$Z	−0.0040	−0.0178	0.0066		
$\alpha_{\perp}^{\rm Rel}$/aCVXZ	0.0089	0.0137	0.0040		
$\alpha_{\parallel}^{\rm Rel}$/aCV$X$Z-DK	0.0237	0.0148	0.0117		
$\alpha_{\perp}^{\rm Rel}$/aCVXZ-DK	−0.0030	0.0007	0.0014		
$\alpha_{\parallel}^{\rm Rel}$/decontracted aV$X$Z	0.0185	0.0128	0.0103		
$\alpha_{\perp}^{\rm Rel}$/decontracted aVXZ	−0.0008	0.0015	0.0020		

aug-cc-pV*X*Z (aV*X*Z), polarized core and valence, aug-cc-pCV*X*Z (aCV*X*Z), weighted core-valence, aug-cc-pwCV*X*Z (awCV*X*Z), and aug-cc-pCV*X*Z recontracted for relativistic calculations, aug-cc-pCV*X*Z-DK (aCV*X*Z-DK).[12,28–32]

An initial assessment of the basis sets and corrections was performed on the same near van der Waals minimum distance as examined for the potential energy curve (R = 4.0 Å). The results of this examination can be seen in Table I. The uncertainty of a value in a particular basis is defined as the difference between that value and the corresponding result in the preceding basis. As such, the frozen-core results converge well in the aV6Z basis with values of 2.6246 ± 0.0051 a_0^3 in the parallel component and −1.2374 ± 0.0006 a_0^3 in the perpendicular component. Likewise, the core-core and core-valence correlation correction beyond the frozen-core approximation is reasonably well converged in the awCVQZ basis with values of −0.0541 ± 0.0102 and 0.0272 ± 0.0045 a_0^3 for the parallel and perpendicular components, respectively. It was determined that a basis set increase to awCV5Z was unnecessary given the minor improvement and the overall small size of the AE-FC correction. We have also checked if higher-order coupled-cluster excitations are important for the interaction-induced polarizability by performing a frozen-core CCSDT calculation in the aVDZ basis set (using the analytical implementation in CFOUR). At R = 4.0 Å, the difference between the CCSDT and CCSD(T) results is 0.0023 a_0^3 for the parallel component and −0.0003 a_0^3 for the perpendicular component. Thus, the full triples correction to the isotropic interaction-induced polarizability is just 0.0005 a_0^3 at this distance and we did not consider this correction any further.

The relativistic corrections were calculated analytically using the exact two-component (X2C) method.[33] This method is available in CFOUR,[34] though commented out in the public release due to the lack of picture change effects on the dipole moment operator. To validate the results obtained from the analytical code with a slightly incomplete treatment of relativity, we performed finite field calculations using the CFOUR and MOLPRO programs. The $\alpha_\parallel$ and $\alpha_\perp$ values for the monomer and dimer were calculated as numerical first derivatives of the analytical dipole moment (CFOUR) and second derivatives of the total energy (MOLPRO) with respect to a change in a uniform electric field. MOLPRO calculations of the relativistic effects used the second-order Douglas-Kroll-Hess (DKH) Hamiltonian, which provides an alternative to the X2C results.[35,36] As can be seen in Fig. 1, the numerical and analytical results from CFOUR agree extraordinarily well and the MOLPRO results provide satisfactory agreement given the difference in methods. As a side note, it was observed in the comparison of $\alpha_\parallel^{\text{dimer}}$ that the second-order DKH correction was around twice that of X2C, as shown in Fig. 2. This difference is small in absolute terms and cancels out in $\Delta\alpha_\parallel$ but is difficult to rationalize on its own. Due to the consistency between methods, the analytical approach was used to calculate the correction. The decontracted aV*X*Z results display the fastest basis set convergence, so we took the decontracted aVQZ values as the preferred ones, with values of 0.0103 ± 0.0025 a_0^3 for the parallel component and 0.0020 ± 0.0005 a_0^3 for the perpendicular component.

With the selection of the basis sets and levels of theory defined above, the interaction-induced isotropic pair polarizability $\Delta\alpha_{\text{ave}}$ at the near van der Waals minimum separation is 0.054 84 ± 0.002 97 a_0^3, where the uncertainty is the square root of the quadratic sum of the uncertainties of the contributing terms. This best level of theory was used to calculate $\Delta\alpha_{\text{ave}}(R)$ and $\Delta\alpha_{\text{aniso}}(R)$ at 25 values of R from 2.6 Å to 12.0 Å. The *ab initio* values for each term and the total $\Delta\alpha_{\text{ave}}$ can be found in Table II, while the corresponding results for $\Delta\alpha_{\text{aniso}}$ are presented in Table SI in the supplementary material. The near cancellation of $\Delta\alpha_\parallel$ and $\Delta\alpha_\perp$ results in values of $\Delta\alpha_{\text{ave}}$ very close to zero, as found previously for the argon dimer.[8]

Using a weighted least-squares routine,[37] the total $\Delta\alpha_{\text{ave}}$ and $\Delta\alpha_{\text{aniso}}$ were fitted to functions with forms[38–40]

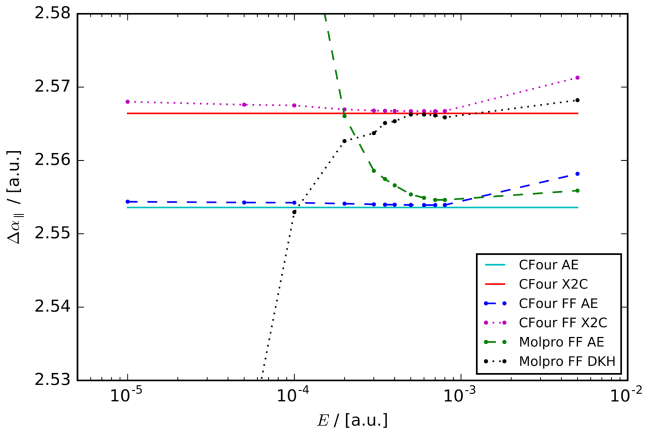


FIG. 1. The parallel component of the interaction-induced pair polarizability, $\Delta\alpha_\parallel$, as a function of electric field strength, E. AE denotes calculations including core-core and core-valence correlation and no relativistic correction. X2C and DKH denote calculations using the respective relativistic correction, while FF denotes calculations using a finite field method. All calculations for this test were performed in the decontracted aVTZ basis set.

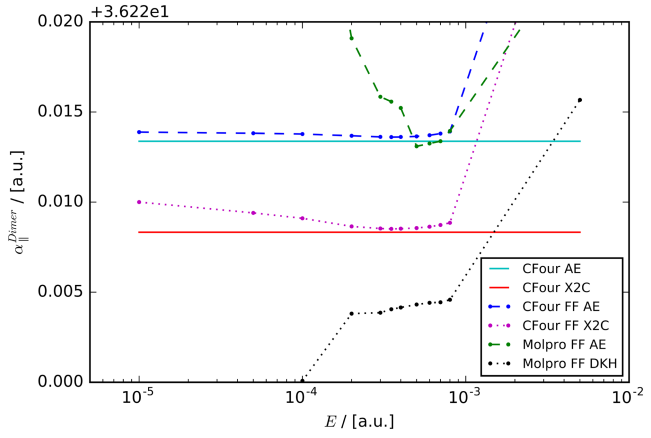


FIG. 2. The parallel component of the dimer polarizability, $\alpha_\parallel^{\text{Dimer}}$, as a function of electric field strength, E in atomic units. AE denotes calculations including core-core and core-valence correlation and no relativistic correction. X2C and DKH denote calculations using the respective relativistic correction, while FF denotes calculations using a finite field method. All calculations for this test were performed in the decontracted aVTZ basis set.

TABLE II. Contributions to the interaction-induced isotropic polarizability $\Delta\alpha_{\rm ave}$ for the krypton dimer, the total $\Delta\alpha_{\rm ave}$, and the total uncertainty $U(\Delta\alpha_{\rm ave})$ in units of a_0^3. The column $\Delta\alpha_{\rm ave}^{\rm fit}$ contains the values obtained from the fitted expression in Eq. (7).

R (Å)	$\Delta\alpha_{\rm ave}^{\rm CCSD(T)/FC}$	$\Delta\alpha_{\rm ave}^{\rm AE-FC}$	$\Delta\alpha_{\rm ave}^{\rm Rel}$	$\Delta\alpha_{\rm ave}$	$U(\Delta\alpha_{\rm ave})$	$\Delta\alpha_{\rm ave}^{\rm fit}$
2.60	−0.175 174 733	−0.020 902 900	−0.018 060 067	−0.214 137 700	0.013 476 933	−0.215 636 912
2.80	−0.312 638 567	−0.005 541 267	−0.000 928 067	−0.319 107 900	0.009 993 667	−0.315 336 594
3.00	−0.281 941 933	0.001 092 167	0.007 473 133	−0.273 376 633	0.008 011 233	−0.275 332 625
3.20	−0.194 681 200	0.003 224 533	0.010 414 167	−0.181 042 500	0.006 781 333	−0.183 014 852
3.40	−0.103 253 400	0.003 155 233	0.010 266 100	−0.089 832 067	0.005 577 233	−0.090 066 202
3.60	−0.029 548 700	0.002 200 600	0.008 675 167	−0.018 672 933	0.004 612 133	−0.017 982 438
3.70	−0.001 491 533	0.001 632 867	0.007 680 067	0.007 821 400	0.004 163 500	0.008 577 516
3.80	0.020 769 433	0.001 080 300	0.006 668 133	0.028 517 867	0.003 739 100	0.029 166 655
3.90	0.037 696 533	0.000 574 767	0.005 693 933	0.043 965 233	0.003 341 767	0.044 415 041
4.00	0.049 917 100	0.000 134 167	0.004 791 967	0.054 843 233	0.002 974 000	0.055 068 809
4.06	0.055 279 367	−0.000 095 467	0.004 293 867	0.059 477 767	0.002 767 800	0.059 576 680
4.10	0.058 127 900	−0.000 233 867	0.003 981 600	0.061 875 633	0.002 636 067	0.061 897 302
4.20	0.063 029 633	−0.000 528 767	0.003 270 300	0.065 771 167	0.002 330 267	0.065 634 625
4.30	0.065 282 433	−0.000 754 633	0.002 658 400	0.067 186 200	0.002 053 967	0.066 946 194
4.40	0.065 480 967	−0.000 918 767	0.002 140 600	0.066 702 800	0.001 806 900	0.066 412 705
4.60	0.061 698 800	−0.001 094 300	0.001 352 633	0.061 957 133	0.001 390 300	0.061 690 744
4.80	0.054 859 900	−0.001 122 400	0.000 829 033	0.054 566 533	0.001 061 133	0.054 412 776
5.00	0.047 011 267	−0.001 061 200	0.000 494 000	0.046 444 067	0.000 805 967	0.046 414 633
5.50	0.029 333 767	−0.000 764 333	0.000 120 933	0.028 690 367	0.000 426 233	0.028 796 038
6.00	0.017 528 567	−0.000 483 733	0.000 025 867	0.017 070 700	0.000 284 100	0.017 116 974
7.00	0.006 518 233	−0.000 181 367	0.000 002 433	0.006 339 300	0.000 160 700	0.006 311 046
8.00	0.002 761 167	−0.000 075 567	0.000 001 467	0.002 687 067	0.000 077 867	0.002 671 768
9.00	0.001 315 300	−0.000 035 533	0.000 000 733	0.001 280 500	0.000 039 433	0.001 277 486
10.00	0.000 684 067	−0.000 018 533	0.000 000 700	0.000 666 233	0.000 019 700	0.000 666 128
12.00	0.000 223 933	−0.000 006 100	0.000 000 233	0.000 218 067	0.000 006 067	0.000 217 420

$$\Delta\alpha_{\rm ave}^{\rm fit}(R) = \left(\frac{A^{\rm (ave)}}{R} + B^{\rm (ave)} + C^{\rm (ave)}R + D^{\rm (ave)}R^2\right) {\rm e}^{-\alpha^{\rm (ave)}R} + \sum\nolimits_{n=6,8} f_n\left(\beta^{\rm (ave)}R\right)\frac{C_n^{\rm (ave)}}{R^n}, \quad (7)$$

$$\Delta\alpha_{\rm aniso}^{\rm fit}(R) = \left(A^{\rm (aniso)}R + B^{\rm (aniso)}R^2\right) {\rm e}^{-\alpha^{\rm (aniso)}R} + \sum\nolimits_{n=3,6,8} f_n\left(\beta^{\rm (aniso)}R\right)\frac{C_n^{\rm (aniso)}}{R^n}, \quad (8)$$

where $f_n(x)$ is the Tang-Toennies damping function[41]

$$f_n(x) = 1 - {\rm e}^{-x}\sum\nolimits_{k=0}^{n}\frac{x^k}{k!}. \quad (9)$$

The weight for each *ab initio* point was the inverse square of the uncertainty and the fitted functions pass within the uncertainties of all calculated points. The $\Delta\alpha_{\rm ave}^{\rm fit}(R)$ function has MUEσ = 0.140 and $\Delta\alpha_{\rm aniso}^{\rm fit}(R)$ has MUEσ = 0.216, where the mean unsigned error with respect to uncertainty (MUEσ) is defined as

$${\rm MUE}\sigma = \frac{1}{N_R}\sum\nolimits_R \frac{|Fit(R) - Calc(R)|}{U(Calc(R))}. \quad (10)$$

In Eq. (10), $Fit(R)$ is either $\Delta\alpha_{\rm ave}^{\rm fit}(R)$ or $\Delta\alpha_{\rm aniso}^{\rm fit}(R)$ and $Calc(R)$ is the corresponding *ab initio* value. The long range C_n terms were fitted first to the data with $R \geq 8.0$ Å assuming no damping, then frozen throughout the rest of the procedure. The values of the fit parameters for $\Delta\alpha_{\rm ave}^{\rm fit}$ can be found in Table III; analogous parameters for the polarizability anisotropy are given in Table SII in the supplementary material.

Additionally, the upper- and lower-limit functions for both $\Delta\alpha_{\rm ave}^{\rm fit}(R)$ ($\Delta\alpha_{\rm ave}^{\rm fit,+}(R)$ and $\Delta\alpha_{\rm ave}^{\rm fit,-}(R)$) and $\Delta\alpha_{\rm aniso}^{\rm fit}(R)$

TABLE III. Parameters of the polarizability functions of $\Delta\alpha_{\rm ave}^{\rm fit}(R)$, $\Delta\alpha_{\rm ave}^{\rm fit,+}(R)$, and $\Delta\alpha_{\rm ave}^{\rm fit,-}(R)$ for the krypton dimer in Eq. (7).

Parameter	$\Delta\alpha_{\rm ave}^{\rm fit}(R)$	$\Delta\alpha_{\rm ave}^{\rm fit,+}(R)$	$\Delta\alpha_{\rm ave}^{\rm fit,-}(R)$	Units
$A^{\rm (ave)}$	−131 248.569 521	−144 967.965 213	−118 621.208 123	a_0^4
$B^{\rm (ave)}$	80 067.715 588	87 897.681 211	72 838.349 627	a_0^3
$C^{\rm (ave)}$	−15 649.670 075	−17 086.345 415	−14 319.251 185	a_0^2
$D^{\rm (ave)}$	958.404 374	1 040.655 286	882.044 335	a_0
$\alpha^{\rm (ave)}$	1.336 794	1.348 786	1.324 710	a_0^{-1}
$\beta^{\rm (ave)}$	0.857 610	0.867 624	0.845 646	a_0^{-1}
$C_6{}^{\rm (ave)}$	27 649.313 556 109 817	28 401.128 468 681 782	26 897.498 643 404 404	a_0^9
$C_8{}^{\rm (ave)}$	992 472.153 870 260 93	1 039 598.256 973 147 2	945 346.050 803 672 29	a_0^{11}

($\Delta\alpha^{\mathrm{fit},+}_{\mathrm{aniso}}(R)$ and $\Delta\alpha^{\mathrm{fit},-}_{\mathrm{aniso}}(R)$) were produced for the measurement of the uncertainty of the corresponding functions and the quantities derived from them. The functions $\Delta\alpha^{\mathrm{fit},+}_{\mathrm{ave}}(R)$ and $\Delta\alpha^{\mathrm{fit},-}_{\mathrm{ave}}(R)$ were fitted to $\Delta\alpha_{\mathrm{ave}} + U(\Delta\alpha_{\mathrm{ave}})$ and $\Delta\alpha_{\mathrm{ave}} - U(\Delta\alpha_{\mathrm{ave}})$, respectively, using the same functional form as $\Delta\alpha^{\mathrm{fit}}_{\mathrm{ave}}(R)$, and the equivalent procedure was used for the production of $\Delta\alpha^{\mathrm{fit},+}_{\mathrm{aniso}}(R)$ and $\Delta\alpha^{\mathrm{fit},-}_{\mathrm{aniso}}(R)$. The relation

$$\Delta\alpha^{\mathrm{fit},-}(R) < \Delta\alpha^{\mathrm{fit}}(R) < \Delta\alpha^{\mathrm{fit},+}(R) \tag{11}$$

is true for both quantities within the range of the fitted points but not necessarily outside of this range. The MUEσ values of the functions $\Delta\alpha^{\mathrm{fit},+}_{\mathrm{ave}}(R)$, $\Delta\alpha^{\mathrm{fit},-}_{\mathrm{ave}}(R)$, $\Delta\alpha^{\mathrm{fit},+}_{\mathrm{aniso}}(R)$, and $\Delta\alpha^{\mathrm{fit},-}_{\mathrm{aniso}}(R)$ are 0.163, 0.119, 0.212, and 0.222, respectively. The parameters for the upper- and lower-limit functions can be found in Table III for $\Delta\alpha^{\mathrm{fit}}_{\mathrm{ave}}$ and Table SII in the supplementary material for $\Delta\alpha^{\mathrm{fit}}_{\mathrm{aniso}}$.

III. POTENTIAL ENERGY CURVE

In our previous work,[10] we developed a new krypton–krypton interatomic potential based on high-level *ab initio* calculations. The analytic potential energy curve, used in this work in the same form, is given by

$$V(R) = \left(A + BR + \frac{C}{R}\right)e^{-\alpha R} - \sum_{n=3}^{4} f_{2n}(\beta R)\frac{C_{2n}}{R^{2n}}, \tag{12}$$

where $R \geq 1.8$ Å, A, B, C, α, β, C_6, and C_8 are the fit parameters and $f_{2n}(x)$ denote the Tang-Toennies damping functions,[41] Eq. (9). To avoid unphysical behavior at short distances ($R < 1.8$ Å), Eq. (12) is spliced continuously with a simpler expression

$$V(R) = \left(\frac{A_{\mathrm{sh}}}{R}\right)e^{-\alpha_{\mathrm{sh}}R+\beta_{\mathrm{sh}}R^2}. \tag{13}$$

In addition, we have constructed potentials $V^+(R)$ and $V^-(R)$ by fitting the same expression as $V(R)$ to $E + U(E)$ and $E - U(E)$, respectively. Here, the uncertainty $U(E)$ of the potential energy E was inferred by a careful examination of basis set convergence patterns.[10] The potential of Ref. 10 was computed within the Born-Oppenheimer approximation (which works extremely well for nuclei as heavy as krypton) and is exactly the same for all isotopologues of the krypton dimer. All fitted parameters in the analytical representations of $V(R)$, $V^+(R)$, and $V^-(R)$ are listed in Table IV.

TABLE IV. Parameters of the potential energy functions $V(R)$, $V^+(R)$, and $V^-(R)$ for the krypton dimer in Eqs. (12) and (13).

Parameter	$V(R)$	$V^+(R)$	$V^-(R)$	Unit
A	467.771 557	511.688	596.938	E_{h}
B	−43.111 875	−45.622	−56.519	$E_{\mathrm{h}} \cdot a_0^{-1}$
C	−509.601 417	−787.134	−997.849	$E_{\mathrm{h}} \cdot a_0$
α	1.566 575	1.558	1.572	a_0^{-1}
β	4.083 794	1.832	1.285	a_0^{-1}
C_6	126.790 499	126.498	127.083	$E_{\mathrm{h}} \cdot a_0^6$
C_8	5268.109 217	5096.285	5439.933	$E_{\mathrm{h}} \cdot a_0^8$
A_{sh}	1296.0	1296.0	1296.0	$E_{\mathrm{h}} \cdot a_0$
α_{sh}	3.067 950	2.744	2.900	a_0^{-1}
β_{sh}	0.324 0714	0.239	0.280	a_0^{-2}

IV. THEORETICAL EVALUATION OF VIRIAL COEFFICIENTS

In this section, we calculate the second pressure, acoustic, and dielectric virial coefficients and estimate their uncertainties.

A. Second pressure virial coefficient

We employ the classical statistical-mechanics formulas with second-order or third-order quantum corrections to compute the virial coefficients of krypton. The calculations of the second pressure virial coefficient B by statistical mechanics have been studied extensively and the explicit expressions are presented in a large number of scientific publications (for example, in Ref. 42). For the convenience of the reader, we still list the details of the formulations necessary to compute B up to the third-order quantum corrections,

$$B = B_{\mathrm{cl}} + \lambda B_{\mathrm{qm},1} + \lambda^2 B_{\mathrm{qm},2} + \lambda^3 B_{\mathrm{qm},3}, \tag{14}$$

where $\lambda = \hbar^2\beta/12m$, $\hbar = h/2\pi$, and $\beta = 1/k_B T$. m is the relative molecular mass (83.798 for krypton in this work), h is the Planck constant, k_{B} is the Boltzmann constant,[43] and T is the temperature. In the case of a completely isotropic potential $V(R)$ (no angular dependence), the classical and quantum contributions can be written as follows:

$$B_{\mathrm{cl}} = -2\pi N_{\mathrm{A}} \int_0^\infty \left[\exp(-\beta V) - 1\right] R^2 dR, \tag{15}$$

$$B_{\mathrm{qm},1} = 2\pi N_{\mathrm{A}} \int_0^\infty (\beta V')^2 \exp(-\beta V)\, R^2 dR, \tag{16}$$

$$B_{\mathrm{qm},2} = -2\pi N_{\mathrm{A}} \int_0^\infty \left[\frac{6}{5}(\beta V'')^2 + \frac{12}{5R^2}(\beta V')^2 + \frac{4}{3R}(\beta V')^3 - \frac{1}{6}(\beta V')^4\right]\exp(-\beta V)\, R^2 dR, \tag{17}$$

$$B_{\mathrm{qm},3} = 2\pi N_{\mathrm{A}} \int_0^\infty \left[\frac{36}{35}(\beta V''')^2 + \frac{216}{35R^2}(\beta V'')^2 + \frac{24}{21}(\beta V'')^3 + \frac{24}{5R}(\beta V')(\beta V'')^2 + \frac{288}{315R^3}(\beta V')^3 \right.$$
$$\left. - \frac{6}{5}(\beta V')^2(\beta V'')^2 - \frac{2}{15R^2}(\beta V')^4 - \frac{2}{5R}(\beta V')^5 + \frac{1}{30}(\beta V')^6\right]\exp(-\beta V)\, R^2 dR, \tag{18}$$

in which N_A denotes the Avogadro constant and $V' = dV/dR$, $V'' = d^2V/dR^2$, and $V''' = d^3V/dR^3$.

B. Second acoustic virial coefficient

The expressions for the second acoustic virial coefficient β_a are related to those for the second pressure virial coefficient B and its first and second temperature derivatives B' and B''. The semi-classical expansion of the second acoustic virial coefficient in powers of λ is[8]

$$\beta_a = \beta_{a,cl} + \lambda\beta_{a,qm,1} + \lambda^2\beta_{a,qm,2}, \tag{19}$$

where the individual terms are formulated as

$$\beta_{a,cl} = 4\pi N_A \int_0^\infty \left[1 - \exp(-\beta V)\left(1 + \frac{2}{5}\beta V + \frac{2}{15}(\beta V)^2\right)\right] R^2 dR, \tag{20}$$

$$\beta_{a,qm,1} = 4\pi N_A \int_0^\infty \left[\frac{3}{5} - \frac{2}{5}\beta V + \frac{2}{15}(\beta V)^2\right] (\beta V')^2 \exp(-\beta V)\, R^2 dR, \tag{21}$$

$$\begin{aligned}\beta_{a,qm,2} = 4\pi N_A \int_0^\infty \Bigg\{ &\left[-\frac{6}{5}(\beta V'')^2 - \frac{12}{5R^2}(\beta V')^2 - \frac{20}{9R}(\beta V')^3 + \frac{13}{30}(\beta V')^4\right] \\ &+ \left[\frac{4}{5}(\beta V'')^2 + \frac{8}{5R^2}(\beta V')^2 + \frac{56}{45R}(\beta V')^3 - \frac{1}{5}(\beta V')^4\right]\beta V \\ &+ \left[-\frac{4}{25}(\beta V'')^2 - \frac{8}{25R^2}(\beta V')^2 - \frac{8}{45R}(\beta V')^3 + \frac{1}{45}(\beta V')^4\right](\beta V)^2\Bigg\} \exp(-\beta V)\, R^2 dR.\end{aligned} \tag{22}$$

C. Second dielectric virial coefficient

The formulas for the dielectric virial coefficients can be derived from the expansion of the Clausius-Mossotti function. The second dielectric virial coefficient can again be approximated as the sum of a classical term and of quantum corrections up to second order[40]

$$B_\varepsilon = B_{\varepsilon,cl} + \lambda B_{\varepsilon,qm,1} + \lambda^2 B_{\varepsilon,qm,2}, \tag{23}$$

where

$$B_{\varepsilon,cl} = \frac{8\pi^2 N_A^2}{3} \int_0^\infty \Delta\alpha_{ave} \exp(-\beta V)\, R^2 dR, \tag{24}$$

$$\begin{aligned}B_{\varepsilon,qm,1} = -\frac{8\pi^2 N_A^2}{3} \int_0^\infty &\left[\Delta\alpha_{ave}(\beta V')^2 - 2\Delta\alpha'_{ave}\beta V'\right] \\ &\times \exp(-\beta V)\, R^2 dR,\end{aligned} \tag{25}$$

$$B_{\varepsilon,qm,2} = \frac{16\pi^2 N_A^2}{5} \int_0^\infty (\Delta\alpha_{ave} f + g) \exp(-\beta V)\, R^2 dR, \tag{26}$$

$$f = (\beta V'')^2 + \frac{2}{R^2}(\beta V')^2 + \frac{10}{9R}(\beta V')^3 - \frac{5}{36}(\beta V')^4, \tag{27}$$

$$\begin{aligned}g = \Delta\alpha'_{ave}&\left[-\frac{4}{R^2}\beta V' - \frac{10}{3R}(\beta V')^2 + \frac{5}{9}(\beta V')^3\right] \\ &- 2\Delta\alpha''_{ave}\beta V''.\end{aligned} \tag{28}$$

Here,

$$\Delta\alpha'_{ave} = d\Delta\alpha_{ave}/dR, \tag{29}$$

$$\Delta\alpha''_{ave} = d^2\Delta\alpha_{ave}/dR^2. \tag{30}$$

D. Uncertainty

There are several possible sources of uncertainty in the present calculation. First, quantum effects become important at low temperatures depending on the molecule in question. The lower-limit temperature for the calculations considered in this work is 115.78 K, the triple-point temperature of krypton. At this temperature, the corresponding ratio of the thermal de Broglie wavelength to the atomic diameter is much smaller than one, i.e., $\left(h/\sqrt{2\pi m k_B T}\right)/\sigma = 0.05 \ll 1$. This guarantees that we can use classical statistical-mechanics formulae with second-order or third-order quantum corrections to compute the different virials of krypton for the whole temperature range.

In addition, Moszynski *et al.*[44] suggested that the use of Padé approximants could better represent full quantum-mechanical results at lower temperatures. Our earlier unpublished work on the second dielectric virial coefficient of neon at 116 K showed that the Padé approximant of order [1/1] reproduced the semi-classical value up to five digits of precision. This agreement further justifies the correctness of classical treatments supplemented by quantum expansions in powers of λ.

The other major contribution to the uncertainty of our calculated values is attributed to the uncertainty of the potential and polarizability of the krypton dimer. As mentioned above, we generated $V^+(R)$, $V^-(R)$, $\Delta\alpha_{ave}^{fit,+}(R)$, and $\Delta\alpha_{ave}^{fit,-}(R)$ expressions that account for the corresponding lower- and upper-limit ranges of potential energies and interaction-induced polarizabilities. Following the work of Hurly *et al.*,[45,46] we attempted to estimate the uncertainty U of the present theoretical results in the following manner:

$$U(X) = \frac{|X_{V^+} - X_{V^-}|}{2}, \tag{31}$$

where X denotes virial coefficients calculated from V^+ and V^-. The calculation of Cencek *et al.*[38] suggested that the influence of the uncertainties in $V(R)$ on the values of B_ε was negligible. Hence, for the second dielectric virial coefficient, the uncertainty of the theoretical values was estimated as

$$U(B_\varepsilon) = \frac{\left| B_{\varepsilon,\Delta\alpha_{\rm ave}^{\rm fit,+}} - B_{\varepsilon,\Delta\alpha_{\rm ave}^{\rm fit,-}} \right|}{2}. \quad (32)$$

V. COMPARISON WITH VIRIAL COEFFICIENT VALUES FROM THE LITERATURE

Here we compare the results of our calculations with values published elsewhere for the second pressure, acoustic, and dielectric virial coefficients.

A. Second pressure virial coefficient

Very recently, Jäger *et al.*[12] reported the calculation of the second pressure virial coefficient of krypton from their new potential. Figure 3 of the present article illustrates that the *ab initio* pressure virial values of Jäger *et al.* are perfectly consistent with our theoretical values over a wide temperature range. The difference between the two *ab initio* values decreases from 1 $cm^3 \cdot mol^{-1}$ at 150 K down to 0.3 $cm^3 \cdot mol^{-1}$ at room temperature and further down to below 0.1 $cm^3 \cdot mol^{-1}$ above 2000 K. Figure 3 also shows the uncertainty of our calculated values, which was evaluated from the potentials $V^+(R)$ and $V^-(R)$ by the procedure discussed above. It is clear that the differences between the second virial coefficient of Jäger *et al.* and the results of this work never exceed the theoretically estimated uncertainty for the complete temperature range.

There is a large number of experimental data for the second virial coefficient of krypton, published more than 30 years ago. The information on the sources of literature was summarized in the compilation of Dymond *et al.*[47] It should be noted that an experimental determination of a strictly two-body quantity such as $B(T)$ requires caution to eliminate three- and higher-body effects even at low densities of the gas.[48] In their Fig. 1, Jäger *et al.*[12] compared their computed values with measured data sets for the second pressure virial coefficient. Because of the excellent agreement between the values of Jäger *et al.* and this work, the same pattern could be drawn for the comparison of the experimental data with our computed values. In many cases, a deviation of more than ±1 $cm^3 \cdot mol^{-1}$ could be observed between the theory and measurements for the second pressure virial coefficient, the maximum deviation being −16 $cm^3 \cdot mol^{-1}$ at the low temperature of 120 K. Aziz and Slaman[49] did not consider the second pressure virial coefficient of krypton as primary data in the development of their empirical potentials due to the inconsistency of experimental data, which is partly supported by the rather large disagreements in Fig. 1 of Jäger *et al.*[12] Moreover, considering the fact that the uncertainty of theoretical values for krypton is normally lower than that of the experimental data,[47] we conclude that both the present results and those of Jäger *et al.* may be employed as recommended values for the second pressure virial coefficient of this gas.

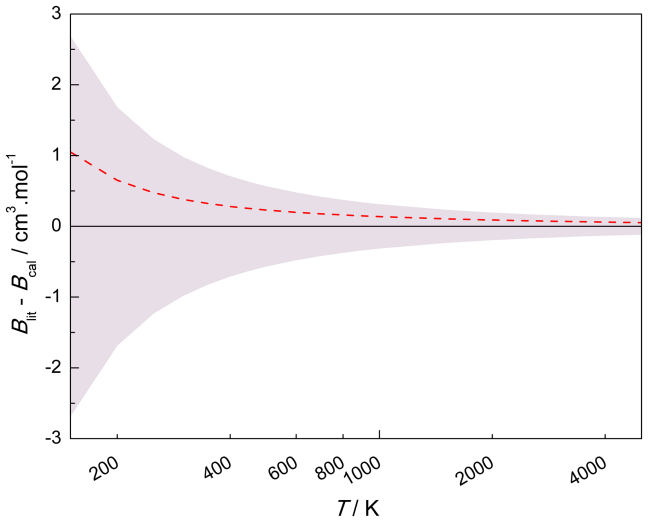


FIG. 3. Absolute deviations of the literature second pressure virial coefficient data, $B_{\rm lit}$, from the values calculated in this work, $B_{\rm cal}$, for krypton as a function of temperature. Data sources: (red dashed line) calculated from the potential by Jäger *et al.*[12] The shaded area corresponds to the uncertainty of $B_{\rm cal}$.

B. Second acoustic virial coefficient

To the best of our knowledge, only one publication with measurements of the second acoustic virial coefficient of krypton is to be found in the literature. Ewing *et al.*[50] built a cylindrical interferometer to measure the speed of sound of gases. Values of the second acoustic virial coefficient of krypton were reported for 285 K, 305 K, and 320 K, with a reported standard uncertainty of around ±1 $cm^3 \cdot mol^{-1}$. Figure 4 shows that the experimental data of Ewing *et al.* are in close agreement of ±(0.5–1.0) $cm^3 \cdot mol^{-1}$ with the theoretically calculated values of this work.

Figure 4 also presents the differences between values computed in this work and those from several krypton potentials in the literature, namely, the well-recognized empirical potential of Dham *et al.*,[51] the old *ab initio* potential of Slavíček *et al.*,[52] and the new, high-accuracy potential of Jäger *et al.*[12] The second acoustic virial coefficient from the potential of Dham *et al.* agrees with the theoretically calculated values of this work at $T \leq 500$ K. However, the absolute deviation begins to exceed the estimated uncertainty of our theoretical values at higher temperatures. The computed values from the

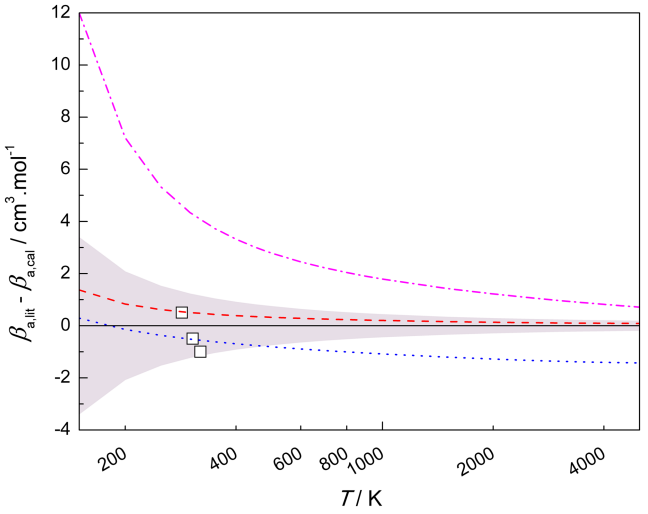


FIG. 4. Absolute deviations of the literature second acoustic virial coefficient data, $\beta_{\rm a,lit}$, from the values calculated in this work, $\beta_{\rm a,cal}$, for krypton as a function of temperature. Data sources: (black open squares) Ewing *et al.*;[50] (blue dotted line) calculated from the potential by Dham *et al.*;[51] (magenta dot line) calculated from the potential by Slavíček *et al.*;[52] (red dashed line) calculated from the potential by Jäger *et al.*[12] The shaded area corresponds to the uncertainty of $\beta_{\rm a,cal}$.

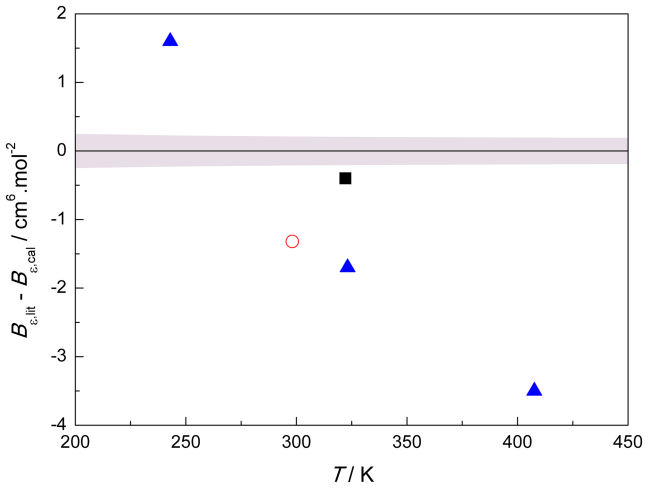


FIG. 5. Absolute deviations of the literature second dielectric virial coefficient data, $B_{\varepsilon,\mathrm{lit}}$, from the values calculated in this work, $B_{\varepsilon,\mathrm{cal}}$, for krypton as a function of temperature. Data sources: (black filled square) Orcutt and Cole;[53] (red open circle) Vidal and Lallemand;[54] (blue filled triangles) Huot and Bose.[55] The shaded area corresponds to the uncertainty of $B_{\varepsilon,\mathrm{cal}}$.

potential of Slavíček *et al.* exhibit a relatively large disagreement with those of this work: the differences are larger than the corresponding uncertainties of our theoretical values over the whole temperature range. On the other hand, one can see in Fig. 4 that there is good agreement for the second acoustic virial coefficient between the new *ab initio* potential of Jäger *et al.* and that of this work. The absolute difference lies within the uncertainty range of our theoretical calculation for temperatures up to 5000 K. We suggest that the present computed values can be used with confidence in different fields wherever values of the second acoustic virial coefficient of krypton are required.

C. Second dielectric virial coefficient

Three sets of experimental data are available in the literature for the second dielectric virial coefficient of krypton.[53–55] Figure 5 shows a comparison of the experimental data with the values calculated here. Orcutt and Cole[53] reported one data point at ambient temperature, where the difference slightly exceeds our theoretical uncertainty. The data point by Vidal and Lallemand[54] at 25 °C shows a disagreement of −1.3 cm^6·mol^{-2}, which is six times larger than the estimated uncertainty of our calculated value. A larger scatter of ±(2–4) cm^6·mol^{-2} can be observed for the deviation of the data by Huot and Bose[55] with respect to the values calculated here. The inconsistency of measurements from different laboratories implies that the prediction of this work may be applied as reference values for the second dielectric virial coefficient.

TABLE V. Virial coefficients of krypton and their estimated uncertainties as a function of temperature. The temperatures 115.78 K and 209.48 K correspond, respectively, to the triple- and critical-point temperatures of krypton. The virial coefficient values at other temperatures are available upon request for the range between 115.78 K and 5000 K.

T (K)	B (cm^3 · mol^{-1})	$U(B)$ (cm^3 · mol^{-1})	β_a (cm^3 · mol^{-1})	$U(\beta_a)$ (cm^3 · mol^{-1})	B_ε (cm^6 · mol^{-2})	$U(B_\varepsilon)$ (cm^6 · mol^{-2})
115.78	−321.30	4.44	−347.47	6.11	10.923	0.384
150	−198.04	2.68	−191.03	3.39	8.7080	0.2976
200	−116.34	1.68	−94.934	2.071	7.2778	0.2447
209.48	−106.55	1.57	−83.747	1.934	7.1080	0.2386
250	−75.164	1.23	−48.270	1.52	6.5637	0.2199
273.15	−62.372	1.09	−33.980	1.37	6.3408	0.2125
273.16	−62.368	1.09	−33.975	1.37	6.3407	0.2125
293.15	−53.306	1.00	−23.911	1.26	6.1817	0.2075
298.15	−51.271	0.98	−21.658	1.23	6.1458	0.2064
300	−50.539	0.98	−20.848	1.22	6.1329	0.2060
350	−34.244	0.82	−2.9210	1.0372	5.8418	0.1973
400	−22.716	0.71	9.6263	0.9079	5.6296	0.1916
450	−14.164	0.62	18.835	0.813	5.4663	0.1876
500	−7.5918	0.56	25.832	0.740	5.3354	0.1847
600	1.7896	0.47	35.644	0.635	5.1354	0.1810
700	8.1039	0.41	42.069	0.561	4.9861	0.1789
800	12.596	0.37	46.498	0.507	4.8676	0.1778
900	15.922	0.34	49.663	0.466	4.7693	0.1772
1000	18.459	0.31	51.983	0.433	4.6851	0.1770
1500	25.184	0.23	57.344	0.333	4.3807	0.1784
2000	27.779	0.19	58.588	0.282	4.1717	0.1811
2500	28.906	0.17	58.517	0.250	4.0089	0.1839
3000	29.382	0.15	57.940	0.228	3.8745	0.1867
3500	29.528	0.14	57.157	0.212	3.7600	0.1894
4000	29.495	0.13	56.297	0.201	3.6604	0.1919
4500	29.360	0.12	55.419	0.191	3.5726	0.1943
5000	29.166	0.11	54.551	0.184	3.4945	0.1965

VI. CONCLUSIONS

We have developed a new interaction-induced isotropic polarizability for the krypton dimer and used it together with our recent interaction potential to calculate the second pressure, acoustic, and dielectric virial coefficients of krypton gas. The coupled-cluster method at the CCSD(T) level as well as basis sets up to aV6Z was selected to determine the interaction-induced isotropic polarizability $\Delta\alpha_{\rm ave}(R)$. The corrections for core-core correlation, core-valence correlation, and relativistic effects were also included to improve the quality of $\Delta\alpha_{\rm ave}(R)$. As a result, values of the interaction-induced isotropic polarizability were determined with low uncertainty for the krypton dimer at 25 different interatomic distances R covering the range of 2.6–12.0 Å. An analytical expression of $\Delta\alpha^{\rm fit}_{\rm ave}(R)$ was obtained by fitting the individual values in order to facilitate the calculation of the second dielectric virial coefficient.

Using the newly developed interaction-induced polarizability, together with our highly accurate interatomic potential,[10] we computed the second pressure, acoustic, and dielectric virial coefficients of krypton gas. Calculations were performed using classical statistical mechanics supplemented with quantum corrections up to second or third orders. The theoretically calculated values of virial coefficients are listed in Table V for the temperature range of 115.78 K–5000 K. The corresponding uncertainty given in Table V is estimated from the difference between the values calculated using $V^{+}(R)$, $V^{-}(R)$, $\Delta\alpha^{\rm fit+}_{\rm ave}(R)$, and $\Delta\alpha^{\rm fit-}_{\rm ave}(R)$.

Comparisons of the literature data were performed with the values computed in this work. Some inconsistencies were found with respect to the relatively small number of experimental data. On the other hand, the two new *ab initio* potentials, the one of Jäger *et al.*[12] and the one of this work, exhibit excellent agreement for the calculations of the second pressure and acoustic virial coefficients. All things considered, we believe that the present theoretically predicted values can be used as reference values for the different virial coefficients of krypton gas. Given the scarcity of reliable experimental data, accurate measurements of thermophysical and electromagnetic properties of krypton gas are desirable to further check the validity of the present predictions.

SUPPLEMENTARY MATERIAL

See supplementary material for the results of the interaction-induced anisotropy polarizability of the krypton dimer.

ACKNOWLEDGMENTS

The authors thank Dr. Mark Plimmer, of the Conservatoire National des Arts et Métiers, for his kind reading of the manuscript. We also thank Dr. Lan Cheng, of the Johns Hopkins University, for his help with the X2C calculations. This work was supported by the National Key R&D Program of China (Grant No. 2016YFE0204200), the U.S. National Science Foundation CAREER Award No. CHE-1351978, and the 111 Project (No. B16038).

[1] J. O. Hirschfelder, C. F. Curtiss, and R. B. Bird, *Molecular Theory of Gases and Liquids* (Wiley, New York, 1954).
[2] S. Chapman and T. G. Cowling, *The Mathematical Theory of Non-Uniform Gases: An Account of the Kinetic Theory of Viscosity, Thermal Conduction and Diffusion of Gases* (Cambridge University Press, Cambridge, 1970).
[3] G. C. Maitland, M. Rigby, E. B. Smith, and W. A. Wakeham, *Intermolecular Forces: Their Origin and Determination* (Clarendon Press, Oxford, 1981).
[4] F. R. W. McCourt, J. J. M. Beenakker, W. E. Köhler, and I. Kuščer, *Non-Equilibrium Phenomena in Polyatomic Gases* (Clarendon Press, Oxford, 1990).
[5] R. Hellmann, "Thermophysical properties of industrially relevant fluids and fluid mixtures from pure theory," in *19th European Conference on Thermophysical Properties* (Aristotle University of Thessaloniki, Greece, Thessaloniki, 2011).
[6] E. Bich, R. Hellmann, and E. Vogel, Mol. Phys. **105**, 3035–3049 (2007).
[7] E. Bich, R. Hellmann, and E. Vogel, Mol. Phys. **106**, 813–825 (2008); **106**, 1107–1122 (2008).
[8] E. Vogel, B. Jäger, R. Hellmann, and E. Bich, Mol. Phys. **108**, 3335–3352 (2010).
[9] W. Cencek, M. Przybytek, J. Komasa, J. B. Mehl, B. Jeziorski, and K. Szalewicz, J. Chem. Phys. **136**, 224303 (2012).
[10] J. M. Waldrop, B. Song, K. Patkowski, and X. Wang, J. Chem. Phys. **142**, 204307 (2015).
[11] B. Song, X. Wang, and Z. Liu, Mol. Simul. **42**, 9–13 (2016).
[12] B. Jäger, R. Hellmann, E. Bich, and E. Vogel, J. Chem. Phys. **144**, 114304 (2016).
[13] R. F. Berg and M. R. Moldover, J. Phys. Chem. Ref. Data **41**, 043104 (2012).
[14] D. Seibt, S. Herrmann, E. Vogel, E. Bich, and E. Hassel, J. Chem. Eng. Data **54**, 2626–2637 (2009).
[15] M. Schäfer, M. Richter, and R. Span, J. Chem. Thermodyn. **89**, 7–15 (2015).
[16] R. M. Gavioso, D. M. Ripa, P. P. M. Steur, C. Gaiser, T. Zandt, B. Fellmuth, M. de Podesta, R. Underwood, G. Sutton, L. Pitre, F. Sparasci, L. Risegari, L. Gianfrani, A. Castrillo, and G. Machin, Philos. Trans. R. Soc., A **374**, 20150046 (2016).
[17] C. Gaiser, B. Fellmuth, and T. Zandt, Int. J. Thermophys. **35**, 395–404 (2014).
[18] R. M. Gavioso, D. M. Ripa, P. P. M. Steur, C. Gaiser, D. Truong, C. Guianvarc'h, P. Tarizzo, F. M. Stuart, and R. Dematteis, Metrologia **52**, S274–S304 (2015).
[19] M. J. Assael, J. A. Assael, M. L. Huber, R. A. Perkins, and Y. Takata, J. Phys. Chem. Ref. Data **40**, 033101 (2011).
[20] B. Jäger, R. Hellmann, E. Bich, and E. Vogel, J. Chem. Phys. **135**, 084308 (2011).
[21] K. R. Shaul, A. J. Schultz, D. A. Kofke, and M. R. Moldover, Chem. Phys. Lett. **531**, 11–17 (2012).
[22] K. D. Hill, AIP Conf. Proc. **1552**, 198–203 (2013).
[23] J. W. Schmidt, R. M. Gavioso, E. F. May, and M. R. Moldover, Phys. Rev. Lett. **98**, 254504 (2007).
[24] S. Boys and F. Bernardi, Mol. Phys. **19**, 553–566 (1970).
[25] J. F. Stanton, J. Gauss, M. E. Harding, P. G. Szalay, A. A. Auer, R. J. Bartlett, U. Benedikt, C. Berger, D. E. Bernholdt, Y. J. Bomble, O. Christiansen, M. Heckert, O. Heun, C. Huber, T.-C. Jagau, D. Jonsson, J. Jusélius, K. Klein, W. J. Lauderdale, D. A. Matthews, T. Metzroth, D. P. O'Neill, D. R. Price, E. Prochnow, K. Ruud, F. Schiffmann, S. Stopkowicz, A. Tajti, J. Vázquez, F. Wang, and J. D. Watts, CFOUR, a quantum chemical program package, containing the integral packages MOLECULE (J. Almlöf and P. R. Taylor), PROPS (P. R. Taylor), ABACUS (T. Helgaker, H. J. Aa. Jensen, P. Jørgensen, and J. Olsen), and ECP routines by A. V. Mitin and C. van Wüllen. For the current version, see http://www.cfour.de.
[26] H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, and M. Schütz, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **2**, 242–253 (2012).
[27] H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, P. Celani, W. Györffy, D. Kats, T. Korona, R. Lindh, A. Mitrushenkov, G. Rauhut, K. R. Shamasundar, T. B. Adler, R. D. Amos, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, E. Goll, C. Hampel, A. Hesselmann, G. Hetzer, T. Hrenar, G. Jansen, C. Köppl, Y. Liu, A. W. Lloyd, R. A. Mata, A. J. May, S. J. McNicholas, W. Meyer, M. E. Mura, A. Nicklaß, D. P. O'Neill, P. Palmieri, D. Peng, K. Pflüger, R. Pitzer, M. Reiher, T. Shiozaki, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, and M. Wang, MOLPRO, version 2012.1, a package of *ab initio* programs, 2012, see http://www.molpro.net.

[28]T. H. Dunning, Jr., J. Chem. Phys. **90**, 1007 (1989).
[29]R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, J. Chem. Phys. **96**, 6796 (1992).
[30]A. K. Wilson, D. E. Woon, K. A. Peterson, and T. H. Dunning, Jr., J. Chem. Phys. **110**, 7667 (1999).
[31]N. B. Balabanov and K. A. Peterson, J. Chem. Phys. **123**, 064107 (2005).
[32]N. J. DeYonker, K. A. Peterson, and A. K. Wilson, J. Phys. Chem. A **111**, 11383–11393 (2007).
[33]K. G. Dyall, J. Chem. Phys. **115**, 9136 (2001).
[34]L. Cheng, S. Stopkowicz, and J. Gauss, Int. J. Quantum Chem. **114**, 1108–1127 (2014).
[35]M. Douglas and N. M. Kroll, Ann. Phys. **82**, 89–155 (1974).
[36]B. A. Hess, Phys. Rev. A **33**, 3742 (1986).
[37]E. Jones, T. Oliphant, P. Peterson *et al.*, SciPy: Open Source Scientific Tools for Python, 2001, http://www.scipy.org/ (Online, accessed September 22, 2017).
[38]W. Cencek, J. Komasa, and K. Szalewicz, J. Chem. Phys. **135**, 014301 (2011).
[39]C. Hättig, H. Larsen, J. Olsen, P. Jørgensen, H. Koch, B. Fernández, and A. Rizzo, J. Chem. Phys. **111**, 10099 (1999).
[40]R. Moszynski, T. G. A. Heijmen, P. E. S. Wormer, and A. van der Avoird, J. Chem. Phys. **104**, 6997 (1996).
[41]K. T. Tang and J. P. Toennies, J. Chem. Phys. **80**, 3726 (1984).
[42]M. E. Boyd, S. Y. Larsen, and J. E. Kilpatrick, J. Chem. Phys. **50**, 4034 (1969).
[43]P. J. Mohr, D. B. Newell, and B. N. Taylor, J. Phys. Chem. Ref. Data **45**, 043102 (2016).
[44]R. Moszynski, T. G. A. Heijmen, and A. van der Avoird, Chem. Phys. Lett. **247**, 440–446 (1995).
[45]J. J. Hurly and M. R. Moldover, J. Res. Natl. Inst. Stand. Technol. **105**, 667–688 (2000).
[46]J. J. Hurly and J. B. Mehl, J. Res. Natl. Inst. Stand. Technol. **112**, 75–94 (2007).
[47]J. H. Dymond, R. C. Wilhoit, K. N. Marsh, K. C. Wong, and M. Frenkel, in *Landolt-Börnstein: Numerical Data and Functional Relationships in Science and Technology: New Series, Group IV: Physical Chemistry*, edited by M. Frenkel and K. N. Marsh (Springer, New York, 2002), Vol. 21A, Chap. 2, pp. 64–67.
[48]J. Wiebke, E. Pahl, and P. Schwerdtfeger, J. Chem. Phys. **137**, 064702 (2012).
[49]R. A. Aziz and M. J. Slaman, Chem. Eng. Commun. **78**, 153–165 (1989).
[50]M. B. Ewing, M. L. McGlashan, and J. P. M. Trusler, J. Chem. Thermodyn. **17**, 549–559 (1985).
[51]A. K. Dham, A. R. Allnatt, W. J. Meath, and R. A. Aziz, Mol. Phys. **67**, 1291–1307 (1989).
[52]P. Slavíček, R. Kalus, P. Paška, I. Odvárková, P. Hobza, and A. Malijevský, J. Chem. Phys. **119**, 2102 (2003).
[53]R. H. Orcutt and R. H. Cole, J. Chem. Phys. **46**, 697 (1967).
[54]D. Vidal and M. Lallemand, J. Chem. Phys. **64**, 4293 (1976).
[55]J. Huot and T. K. Bose, J. Chem. Phys. **95**, 2683 (1991).

Appendix C

Interactions of CO_2 with Metal-Organic Framework Models **[Undergoing Revision]**

Journal Name

ARTICLE TYPE

Cite this: DOI: 00.0000/xxxxxxxxxx

Interactions of CO_2 with Metal-Organic Framework Models†

Jonathan M. Waldrop[a] and Konrad Patkowski*[a]

Received Date
Accepted Date

DOI: 00.0000/xxxxxxxxxx

The interactions between carbon dioxide and model systems of metal-organic frameworks were studied using a variety of *ab initio* methods. Three metal species and four organic linkers were considered in these models as a representation of the tunable nature of metal-organic frameworks and the potential multireference character of such systems. Common single-reference methods, such as MP2 and CCSD(T), were compared with multireference methods based on complete active space self consistent field theory, going as far as MRCISD. Special consideration is taken to avoid issues of size inconsistency in the CI results, where an alternate reference is used in the interaction energy definition. The benchmark values are used to judge the adequacy of a selection of density functionals for the current systems. Symmetry-adapted perturbation theory decomposition was performed to elucidate the important effects that comprise the binding interactions. The systems proved to have very limited multireference character, and MP2 values were closer to the CCSD(T) benchmark than the more difficult MRCISD results. Though the SAPT total energies prove to be relatively poor approximations to the benchmark interaction energies, the energy decompositions indicate that the CO_2 binding is driven by electrostatics and induction for the benzenedicarboxylate and furandicarboxylate complexes but has a significant contribution of dispersion for the diformate and paddlewheel systems.

Introduction

Metal-organic frameworks (MOFs) are crystalline structures formed by the coordination of organic linker molecules to metal centers or clusters. The resulting structures are highly porous and can have open coordination sites on the metal centers that allow for improved interactions with gas molecules. The interactions of MOFs with various gases are of great interest in the areas of gas storage, separation, and catalysis due to the tunable nature of the MOF structures and the effect that this tunablilty plays on the selectivity of the MOF's interactions [1–3]. More specifically, these compounds have garnered interest in areas focused on the generation of renewable energy and/or green alternatives and improvements to fossil fuel based energy production. MOFs have been considered as potential storage media for H_2 for usage as an alternative fuel source [4]. There has also been interest in using MOFs as selective gas filters to sequester CO_2 from flue and exhaust gases [5–7].

The design of MOFs with a desired level of selectivity for a particular gas requires a detailed understanding of the interactions between the adsorbent and adsorbate. Computational modeling can provide insight into these interactions, but requires the balancing of accuracy to sample size. Given the periodic, bulk nature of MOFs, they are complicated systems to treat accurately from a computational perspective. Methods that allow for a large scale treatment of a bulk representative structure, such as periodic density functional theory (DFT), have been found incapable of accounting for dispersion interactions which are crucial for the physisorption process [8,9]. More accurate methods, such as the wide array of wavefunction based *ab initio* ones, are capable of a more complete description of these interactions, but are not trivial to apply to large-scale simulations.

An ideal method to study these systems would be applicable to large-scale models while also capturing the important components of their interactions. This goal requires benchmarking potential candidates against high accuracy *ab initio* calculations. In the process, an adequate treatment of the possible multireference character of these systems is of potential key importance. Transition metal atoms and ions with different d-electron counts and spin states may require the consideration of several important electronic configurations, resulting in non-dynamical correlation.

Symmetry-adapted perturbation theory (SAPT) [10–12] is a powerful approach for the calculation of the interaction energy between molecules. SAPT is a perturbative method that directly

[a] *Department of Chemistry and Biochemistry, Auburn University, Auburn, AL 36849, United States. E-mail: patkowsk@auburn.edu*

† Electronic Supplementary Information (ESI) available: a detailed explanation of the algorithm used to determine geometries, Cartesian coordinates for all systems considered, and additional figures. See DOI: 00.0000/00000000.

calculates the interaction energy as opposed to the indirect determination of this quantity via the supermolecular method. The most important advantage of SAPT is that it provides a physically meaningful breakdown of the interaction energy in terms of the electrostatic, exchange, induction, and dispersion components. An analysis of this decomposition can provide insight on the selectivity of gas adsorption and can also inform the development of more approximate methodologies to accurately model these systems, such as first principles-based force fields[13]. Perhaps even more importantly, the SAPT decomposition provides a route to a rational design of MOFs and other adsorbent materials via amplifying or reducing particular interaction energy terms[14].

The benchmarking of DFT for interactions involving MOFs has already been considered by several groups. Grajciar et al. investigated the suitability of DFT for the calculation of interactions of water with the uncoordinated sites of a MOF[15]. In comparison to their benchmark coupled-cluster singles and doubles with perturbative triples calculations at the complete basis set limit (CCSD(T)/CBS), DFT was found unreliable in the reproduction of interaction energies. The addition of empirical dispersion was not sufficient to correct the shortcomings of this method. The authors determined that the composite DFT/CC method[16,17], which adds an empirical pairwise correction parameterized on CCSD(T) results to the DFT results in a fashion similar to the empirical dispersion corrections[18–20], was capable of attaining acceptable agreement with experimental and higher-level theoretical results[21].

Taking a different approach, Howe et al. looked at the interactions of several gases with the external surfaces of MOFs[22]. Their calculations utilized a number of density functionals and wavefuction methods up to CCSD(T) to consider different sizes of model systems of the MOF, taking into account the importance of periodicity. Their results show good agreement between different model sizes and predict greater interactions for the Zn and Co containing systems over their Cu analogs.

Dzubak et al. produced a method for the development of force fields for interactions between MOFs with open coordination sites and gas molecules[23]. Their methodology uses a decomposition of second-order Møller-Plesset perturbation theory interactions as the basis for the terms of a new force field. The newly produced force field showed a favorable improvement over the universal force field (UFF) for the simulation of adsorption isotherms of the systems in question and also proved to be transferable to systems other than the fitted one. McDaniel and Schmidt[13] used DFT based SAPT (DFT-SAPT) for the decomposition of interactions to produce physically motivated transferable force fields, and applied these force fields to the calculation of gas adsorption to MOFs.

In another investigation of the interactions of gas molecules with MOFs, Goings et al. produced SAPT and DFT-SAPT decompositions for the potential energy curves between H_2 and a MOF model containing various metals[24]. The study considered Mg, Ca, Sr, Mn, and Zn metal centers for their models. Geometries and binding energies were produced for all metals, but the SAPT calculations were restricted to the closed-shell dications of Mg, Ca, and Zn. In comparison to the CCSD(T) benchmark, the DFT-SAPT binding energies showed a much greater discrepancy than the SAPT results due to differences in the exchange contributions. The SAPT decompositions also show a comparably small, but important, contribution from dispersion interactions.

Recently, Sladek and Tvaroška studied the interactions between divalent cations of Mn and Mg with a number of small molecules representing the components of amino acids that are most likely to bind with these metals in biomolecules[25]. The authors use SAPT decompositions in an attempt to highlight the difference between the two metals as catalytic components in enzymes, where Mn can often serve as an acceptable stand-in for Mg while the reverse is not necessarily true. Their results show that the induction components have a greater magnitude in the interactions involving Mn^{2+} and that the preference over Mg^{2+} in certain systems may be due to preferential interactions between Mn^{2+} and ligands that bind through N or S.

In the current study, the interactions of MOF mimic systems with CO_2 are studied at several levels of *ab initio* theory including spin-restricted CCSD(T) (RCCSD(T)), the complete active space self consistent field (CASSCF), and post-CASSCF levels of theory to determine appropriate benchmark treatments of metal centers with different occupations of the *d* shell. An alternative reference for the interaction energy is defined to circumvent issues of size inconsistency in MRCISD calculations. Then, the SAPT decompositions for the closed and open-shell systems are performed to elucidate the importance of different interactions for different model systems. These results are compared to DFT data to determine the adequacy of various functionals and dispersion corrections for calculations of interaction energies in systems containing metal ions with various spin states.

Methodology and Computational Details

The MOLPRO code was used for supermolecular interaction energy calculations and all geometry optimizations[26,27]. SAPT calculations were performed with the PSI4 code[28]. Augmented correlation-consistent Dunning basis sets were used throughout and included aug-cc-pVDZ through aug-cc-pVQZ (a*X*Z)[29,30]. Density fitting (DF) was utilized for Hartree-Fock (HF), second order Møller-Plesset perturbation theory (MP2), and Symmetry Adapted Perturbation Theory (SAPT)[10] calculations, as well as Density Functional Theory (DFT) calculations in the aQZ basis set. Due to the limited availability of auxiliary basis sets for the selected metals, the aXZ/MP2FIT[31–33] basis set was used in place of the aXZ/JKFIT[34] basis set and aTZ/MP2FIT was employed with both aDZ and aTZ orbital bases. The DFT functionals used in these calculations were BLYP[35], B3LYP[36,37], PBE[38], PBE0[39,40], M05-2X[41], and M06-2X[42]. Grimme's DFT-D3 code was used to calculate the dispersion corrections[18–20], using both the original and the revised damping parameters[43]. For the open-shell SAPT decompositions, both restricted open-shell Hartree-Fock (ROHF) and unrestricted Hartree-Fock (UHF) based methods were used[44–46].

Model Structures and Geometry Optimization

Four types of systems were utilized as models for MOFs interacting with CO_2. The MOF fragment monomers each contain

one or two metal centers and a number of representative organic ligands. The overall charge of each system is zero. The metal species used throughout are the dicationic forms of Mn, Co, and Zn. The metals were chosen to provide models for completely filled (Zn^{2+}), halfway filled (Mn^{2+}), and intermediately filled (Co^{2+}) *d* shell systems. The first two MOF models contain a single benzenedicarboxylate (BDC) or furandicarboxylate (FDC) linker coordinated to a single metal center. These systems were optimized at the B3LYP/aDZ level. The third model consists of a single metal center coordinated to two formate ($HCOO^-$) linkers. This system was first optimized at the same level of theory as the previously mentioned structures, then additionally reoptimized at the MP2/aTZ level. The fourth model has four formate anions coordinated simultaneously to two metal centers, referred to as a paddlewheel structure[24], and was optimized at the MP2/aDZ level. The representative structures for each complex can be seen in Figure 1 and several important geometric parameters can be found in Table 1.

Table 1 Geometric parameters (in Angstrom) of the optimized structures for each complex. Values in parentheses are the optimized values for the MOF model without CO_2. $R_{Ligand-M^{2+}}$ is defined as the distance between the metal cation and the closest carbon of the ligand, and $R_{M^{2+}-CO_2}$ is the distance between the metal anion and the CO_2 carbon. For the paddlewheel model, X refers to the center of mass of the MOF model and $R_{M^{2+}-CO_2}$ is the distance between the CO_2 and the nearer of the two metal centers. The FDC system is the only one where the ligand-M^{2+}-CO_2 angle is not fixed by symmetry.

Parameter	Mn^{2+}	Co^{2+}	Zn^{2+}
	BDC		
$R_{Ligand-M^{2+}}$	2.35 (2.36)	2.35 (2.29)	2.31 (2.27)
$R_{M^{2+}-CO_2}$	3.32	3.17	3.38
	FDC		
$R_{Ligand-M^{2+}}$	2.33 (2.32)	2.30 (2.27)	2.28 (2.26)
$R_{M^{2+}-CO_2}$	3.32	3.35	3.21
$\Theta_{Ligand-M^{2+}-CO_2}$	179.0°	155.7°	178.9°
	Diformate		
$R_{Ligand-M^{2+}}$	2.46 (2.45)	2.40 (2.38)	2.35 (2.34)
$R_{M^{2+}-CO_2}$	3.58	3.70	3.59
	Paddlewheel		
$R_{X-M^{2+}}$	1.53 (1.52)	1.48 (1.47)	1.36 (1.34)
$R_{X-Ligand}$	2.65 (2.64)	2.64 (2.63)	2.56 (2.56)
$R_{M^{2+}-CO_2}$	3.47	3.65	3.36

For systems where the metal is Mn^{2+} or Zn^{2+}, the dimer structures were optimized in both a fully unrestricted way as well as with the monomers restricted to their individually optimized structures. This latter, rigid configuration was used for scans of the radial potential energy curve. The full geometry optimizations were initially not performed for systems containing Co^{2+} due to the potentially multireference nature of these systems. Instead, the corresponding Mn^{2+} geometry for each system was used as a starting point for simple scans at the complete active space self consistent field (CASSCF) level[47,48] in the aDZ basis set. The details of these scans can be found in the supplementary information. The scanned geometries were also compared to the results of a full optimization and were found to be acceptable representatives.

The BDC-M^{2+} dimers were found to be of nearly C_{2v} symmetry and were symmetrized for computational advantage. The FDC-M^{2+} structures were of C_s symmetry. The initial unrestricted optimizations of the diformate complex showed notable deformations from the planar geometries of the MOF mimics. It was determined that these geometries would not serve as suitable models for MOFs, given the general rigidity of the larger systems, and the optimizations were rerun restricting the $(HCOO^-)_2M^{2+}$ unit to a planar orientation.

Two minima were found for the paddlewheel dimers, the key difference between them being the position of CO_2. A visual comparison of the structures can be found in Figure 2. The lower energy geometry, seen in the bottom of Figure 2, has CO_2 coordinated to one metal center, while also showing signs of interaction with two of the formate ligands. The other geometry has CO_2 positioned on the axis passing through the two metal centers, having an overall C_{4v} symmetry. While the lower energy result is also representative of interactions that can occur between gas molecules and MOF linkers, the higher energy geometry was chosen for symmetry reasons.

Single Point Calculations and Interaction Energies

The MOF-CO_2 interaction energies were calculated in the counterpoise (CP) corrected fashion, that is, the monomer calculations utilized the full dimer basis set[49]. This scheme gives an interaction energy defined as

$$E_{int}^{CP} = E^{AB}(AB) - E^{AB}(A) - E^{AB}(B) \quad (1)$$

where the superscript denotes the usage of the dimer basis set and the letters in parentheses denote the full dimer system or a monomer subsystem. E_{int}^{CP} was calculated with the previously mentioned DFT functionals as well as MP2 and RCCSD(T). Conventional MP2 was used in the case of the paddlewheel calculations, while density fitted MP2 was used with all others. Additionally, E_{int}^{CP} was calculated at the CASSCF level of theory along with the following post-CASSCF methods: second-order multi-reference Rayleigh-Schrödinger perturbation theory (CASPT2)[50,51], second-order *n*-electron valence state perturbation theory (NEVPT2)[52–54], and multi-reference configuration interaction with single and double excitations (MRCISD)[55,56] including the Davidson correction[57,58]. All CAS based calculations used the aDZ basis set. Because of the size inconsistency in the MRCISD calculations, an alternative form of the interaction energy was defined as

$$E_{int}^{CP*} = E^{R}(AB) - E^{R\infty}(AB) \quad (2)$$

where the first term is the dimer energy at the optimized intermolecular distance and the second is the energy of the complex at an exceedingly large intermolecular distance ($R \approx 1000$ Å) approximating the infinite separation of the monomers. Additionally, ghost molecules are positioned at dimer-optimized intermolecular separations near each monomer to maintain the counterpoise correction. A depiction of such a configuration can be

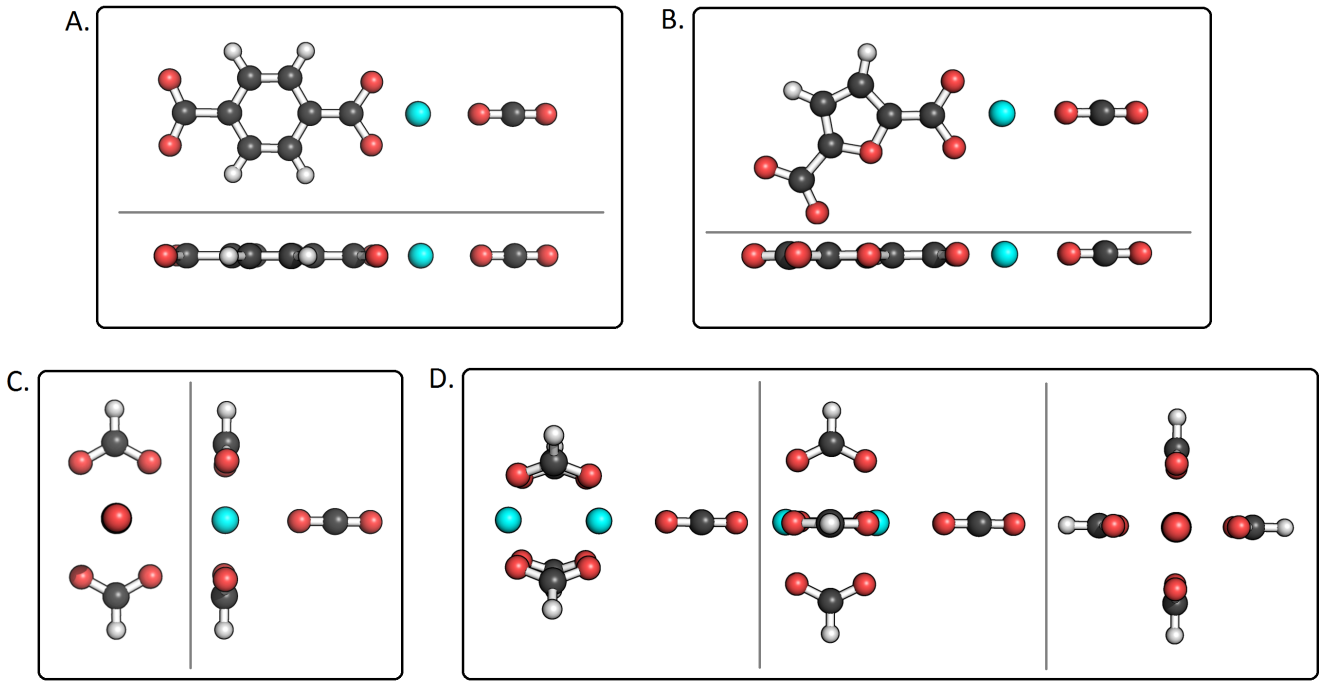


Fig. 1 A. Terephthalate (benzenedicarboxylate) model. B. Furandicarboxylate model. C. Diformate model. D. Paddlewheel model.

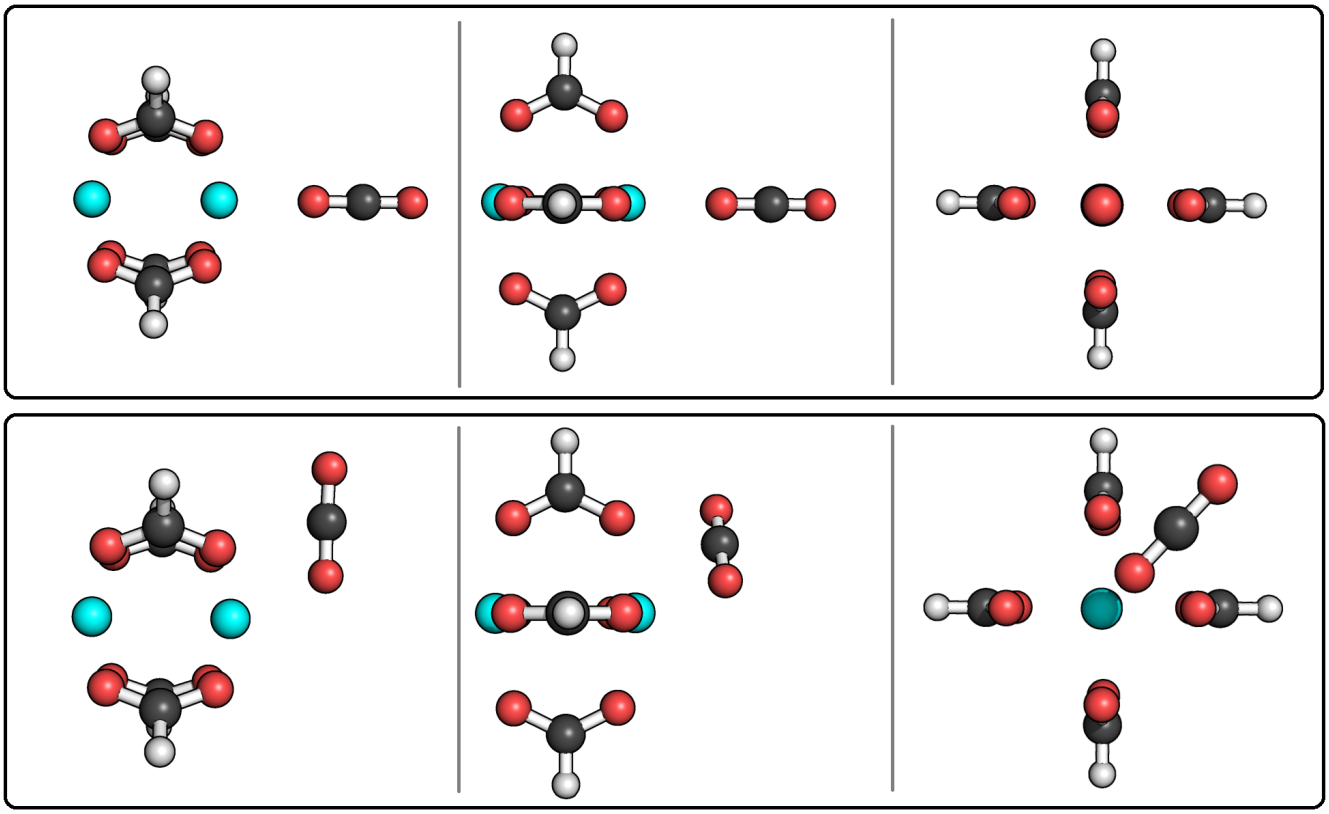


Fig. 2 Top: Axially constrained geometry for the Paddlewheel-CO_2 complex. Bottom: Unconstrained geometry.

seen in Figure 3. The change in the structure of the monomers in the complex with respect to their individually optimized structures, called the monomer deformation correction, was calculated at the same level of theory as the optimized geometries. This correction has the form

$$\Delta E_{int}^{flex} = [E^A(A) - E_0^A(A)] + [E^B(B) - E_0^B(B)] \quad (3)$$

where the superscript denotes that the energy was calculated in the monomer basis set and the subscript 0 denotes that the geometry was the optimized geometry for an isolated monomer.

The 3d, 4s, and 4p orbitals of the metals were considered in the active space. As such, the active space for the BDC, FDC, and diformate models was either (n,6) (without 4p) or (n,9) (with 4p) where n is the d-electron count of the metal cation. The Mn^{2+} model is considered a sextet, the Co^{2+} a quartet, and Zn^{2+} a singlet. Since the paddlewheel models contain two metal centers, the active space included orbitals on both metals. This results in a ($2n$,12) active space if the 3d and 4s orbitals are considered, and ($2n$,18) if 4p are included. The active space including the 4p orbitals was deemed too large for the purposes of this study, so only the ($2n$,12) active space was considered for these models. Only the ferromagnetically coupled paddlewheel configurations were considered for the Mn^{2+} (an undecaplet) and Co^{2+} (a septet) systems, following the reasoning of Grajciar et al.[15] that the interaction between the MOF models and a gas molecule was not strongly dependent on the spin state. The 1s, 2s, and 2p orbitals of the metal centers and the 1s orbitals of the second row elements were frozen in all post-HF calculations. Due to limitations of the CASPT2 implementation, the 3s and 3p orbitals on the metals and the 2s orbitals on the second row atoms were also frozen in the CASPT2 calculations. The point-group representation for all systems was A_1 in the C_{2v} geometries and A′ for the C_s FDC. Natural orbital occupation numbers (NOONs) were calculated at the CASSCF/aDZ level for all considered active spaces.

Potential Energy Curves

A study of E_{int}^{CP} as a function of the intermolecular distance was performed on all systems at the B3LYP/aDZ, MP2/aDZ, and RCCSD(T)/aDZ levels of theory. The intermolecular distance is defined here as the distance between the CO_2 carbon and the metal center of the MOF model (the closer of the two metals in the case of the paddlewheel systems). The intermolecular distances, R_0, for these calculations were determined by optimizing the separation of the rigid monomer structures, so as to avoid the monomer deformation effects skewing the curve. For the Co^{2+} systems, the geometries used were a combination of the scanned dimer structure and the individually optimized CO_2, with the intermolecular distance between the monomers allowed to adjust to the altered CO_2. The distances are represented in terms of a dimensionless parameter $z = R/R_0$, so that $z = 1$ is the optimized structure.

SAPT Calculations

Closed-shell SAPT energy decomposition was performed on each Zn^{2+} containing system at the level of SAPT0 (the sum of first and second-order terms including response and the δ(HF) correction for higher-order induction, but neglecting intramonomer correlation) and SAPT2+3 (including terms up to third-order in the intermolecular interaction and second order in the intramolecular correlation, omitting δ(MP2))[59]. The SAPT0 calculations were done in both aDZ and aTZ basis sets, while the SAPT2+3 calculations were performed with aDZ only. High-spin open-shell ROHF-based SAPT0 calculations were performed for the systems containing the other two metals with the aDZ and aTZ bases. We have also computed UHF-based SAPT for the Mn^{2+} containing systems. For the ROHF-based SAPT results, the induction terms currently implemented in Psi4 are not able to include the Coupled-Perturbed Hartree-Fock (CPHF) response. CPHF induction is available for the UHF based SAPT, but was excluded in this study because of CPHF convergence problems. Both methods include the δ(HF) correction.

Numerical Results and Discussion

Single Point Calculations and Interaction Energies

Table 2 shows the results for the multireference calculations on the diformate systems in both active spaces, along with the RHF/aDZ, MP2/aDZ, and RCCSD(T)/aDZ results for comparison. These systems provide the opportunity to compare the consistency between E_{int}^{CP} and E_{int}^{CP*} and establish the importance of the size consistency. Most of the results are in perfect agreement at the presented accuracy. The differences between E_{int}^{CP} and E_{int}^{CP*} in these systems are never greater than 0.02 kcal/mol, excluding the considerable differences of the MRCISD results. The issue of size inconsistency in the conventional E_{int}^{CP} result at the MRCISD/aDZ level leads to results that are incongruous with the other values, even with the Davidson correction. The results for E_{int}^{CP*} are much more reasonable.

Table 2 Comparison of E_{int}^{CP} (Eq. (1)) and E_{int}^{CP*} (Eq. (2)) for the M^{2+}-diformate-CO_2 model systems (kcal/mol).

Metal	Mn^{2+}		Co^{2+}		Zn^{2+}	
Method	E_{int}^{CP}	E_{int}^{CP*}	E_{int}^{CP}	E_{int}^{CP*}	E_{int}^{CP}	E_{int}^{CP*}
RHF	-3.18	-3.18	-1.53	-1.53	-1.96	-1.96
MP2	-5.32	-5.32	-2.82	-2.82	-3.87	-3.88
RCCSD(T)	-5.26	-5.26	-2.66	-2.66	-3.93	-3.93
(n,6)			Active Space		(3d4s)	
CASSCF	-3.18	-3.18	-1.44	-1.44	-1.13	-1.13
CASPT2	-5.37	-5.35	-3.03	-3.03	-2.78	-2.78
NEVPT2	-5.13	-5.15	-1.54	-1.54	-2.50	-2.50
MRCISD	37.69	-4.95	42.09	-2.65	43.69	-3.54
(n,9)			Active Space		(3d4s4p)	
CASSCF	-2.99	-2.99	-1.05	-1.05	-1.65	-1.65
CASPT2	-5.51	-5.51	-2.61	-2.61	-3.82	-3.82
NEVPT2	-4.74	-4.74	-2.25	-2.25	-3.31	-3.31
MRCISD	37.21	-4.90	42.24	-2.69	43.64	-3.67

The interactions between $(HCOO^-)_2Co^{2+}$ and CO_2 are slightly weaker than the $(HCOO^-)_2Zn^{2+}$-CO_2 interactions, and both are weaker than the $(HCOO^-)_2Mn^{2+}$-CO_2 interactions. MP2 is consistently in good agreement with the RCCSD(T) results, differing by no more than 0.16 kcal/mol. Looking at the multirefer-

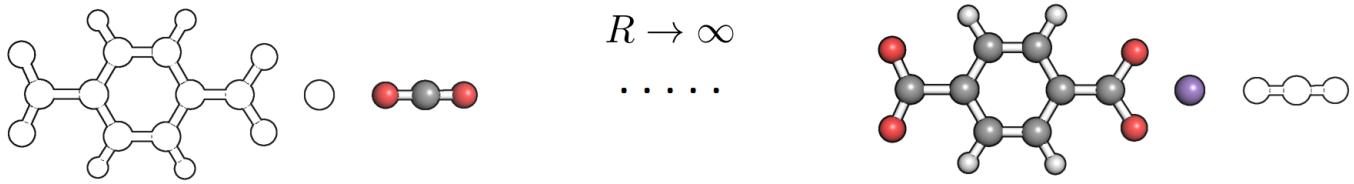


Fig. 3 Visualization of the counterpoise corrected, size-consistent reference (Eq. (2)) for the MRCISD calculations. Unfilled contours denote ghost atoms.

ence methods in the (n,6) active space, the Mn^{2+} results show no change from HF to CASSCF. Likewise, the CASPT2 results are only 0.03-0.05 kcal/mol different from the MP2 results. The NEVPT2 results have a slightly larger difference at 0.07-0.09 kcal/mol. The two multireference perturbation methods bookend the RCCSD(T) results at about 0.1 kcal/mol on either side. The MRCISD result in this active space is less negative than the values for RCCSD(T). In the larger (n,9) space, CASSCF and the perturbation theory methods show a greater degree of difference from the single reference methods than the results in the more limited active space. The CASPT2 and NEVPT2 results are inconsistent in the direction of the change from the smaller active space, with the CASPT2 interaction energy being more attractive in the larger active space while the NEVPT2 result is less attractive. Generally, the Mn^{2+} results in the small active space are closer to the RCCSD(T) results. The Co^{2+} results have some similarity with the Mn^{2+} results. The effect of active space sizes on CASSCF results is more pronounced for Co^{2+} and the step to the larger active space is beneficial for CASPT2 and NEVPT2. The MRCISD accuracy remains about the same in either active space. The Zn^{2+} system is the most effected by the move from single reference to multireference for this model. The smaller active space is consistently further away from the CC results than the larger active space. In the diformate models, MRCISD slightly underestimates the interaction energy compared to the RCCSD(T) results.

Figure 4 shows DFT results, with and without the -D3(BJ) dispersion correction, for the three diformate systems, along with the HF, MP2, and RCCSD(T) results. Again, the agreement between RCCSD(T) and MP2 in the aDZ basis set is quite good for all three metals, with MP2/aTZ resulting in slightly stronger binding. The DFT results are presented in the aDZ basis: though aTZ and aQZ results were obtained, the effect of basis set incompleteness on the aDZ results was minimal. In general, the DFT functionals without the dispersion correction give a poor agreement with the coupled-cluster results, though the hybrid functionals are better than the generalized gradient approximation based functionals (GGAs). M05-2X and M06-2X overestimate the interaction energy in all three cases. BLYP, B3LYP, PBE, and PBE0 are improved considerably with the addition of the dispersion corrections. The selection of damping parameters appears to be of little importance and only the -D3(BJ) results are shown here. The results for the other damping parameters for these and the following systems can be found in the supplementary information.

For the paddlewheel models, Table 3, it can be seen that some difficulties were encountered. The E_{int}^{CP} and E_{int}^{CP*} values for CASSCF for the Mn^{2+} and Zn^{2+} systems show the greatest discrepancies here due to difficulties associated with the CASSCF convergence. Additionally, the NEVPT2 E_{int}^{CP*} value for the Co^{2+} system was not obtained due to difficulty in the calculation of the infinite reference. The accuracy of the NEVPT2 results, compared to RCCSD(T), fluctuates throughout this work, and CASPT2 proved incapable of any useful predictive capacity past the diformate systems and was discarded. We focus here on the E_{int}^{CP*} results, when available, based on their consistency with the single-reference results. The results in Table 3 lead to similar conclusions as for the diformate systems. The paddlewheel fragment has the same order of preference for the metal cations, Mn>Zn>Co (implying Mn has the strongest interaction and Co has the weakest). Again, the transition to multireference methods has a greater effect for the Zn^{2+} systems than Co^{2+} and MP2 is in very good agreement with RCCSD(T). The paddlewheel DFT results found in Figure 5 show that dispersion is also important to the accuracy of DFT for these systems. For these models, the DFT results without -D3 are roughly of HF quality or worse. As above, the hybrid functionals agree with RCCSD(T) better than the GGAs and the Minnesota functionals overestimate the interaction energies. Their overestimation is quite severe for the Mn^{2+} case, but less so for the other two metals. The agreement between the dispersion corrected hybrid functionals and RCCSD(T) is very good and the selection of damping parameters is less important.

The results for the BDC and FDC models are presented in Tables 4 and 5, respectively. Interestingly, the RCCSD(T) result for

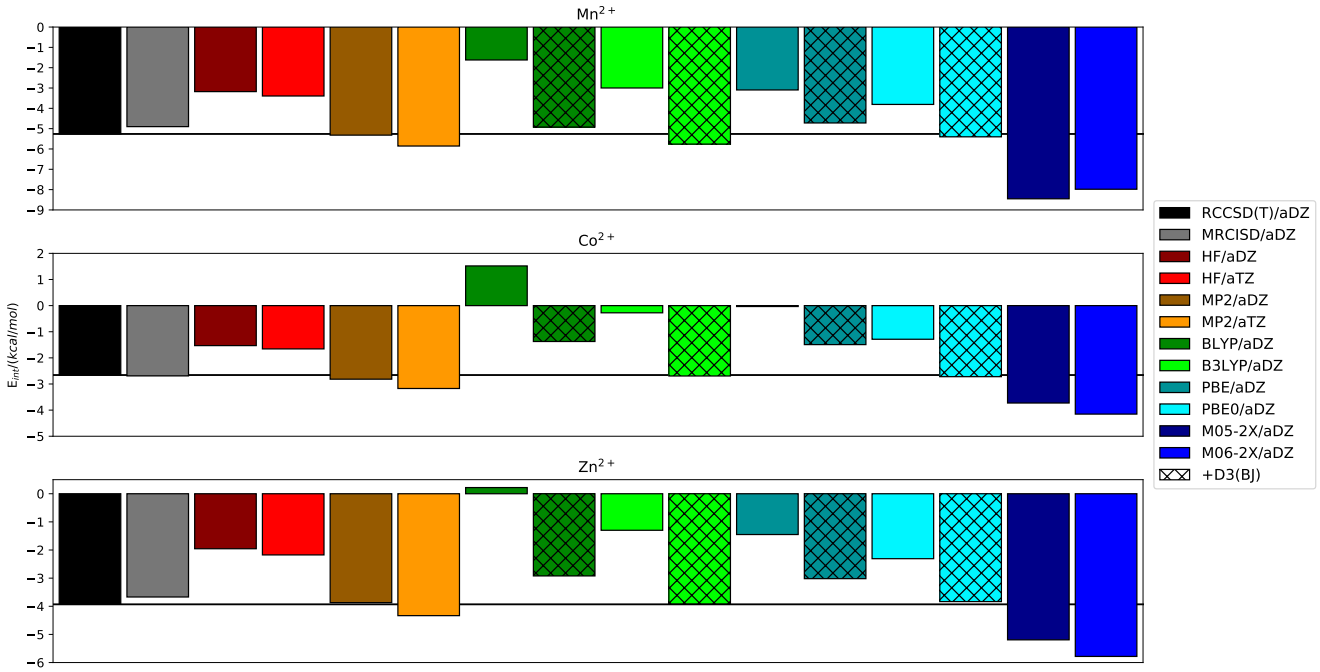


Fig. 4 Counterpoise corrected interaction energies (kcal/mol) for diformate MOF model-CO_2 systems.

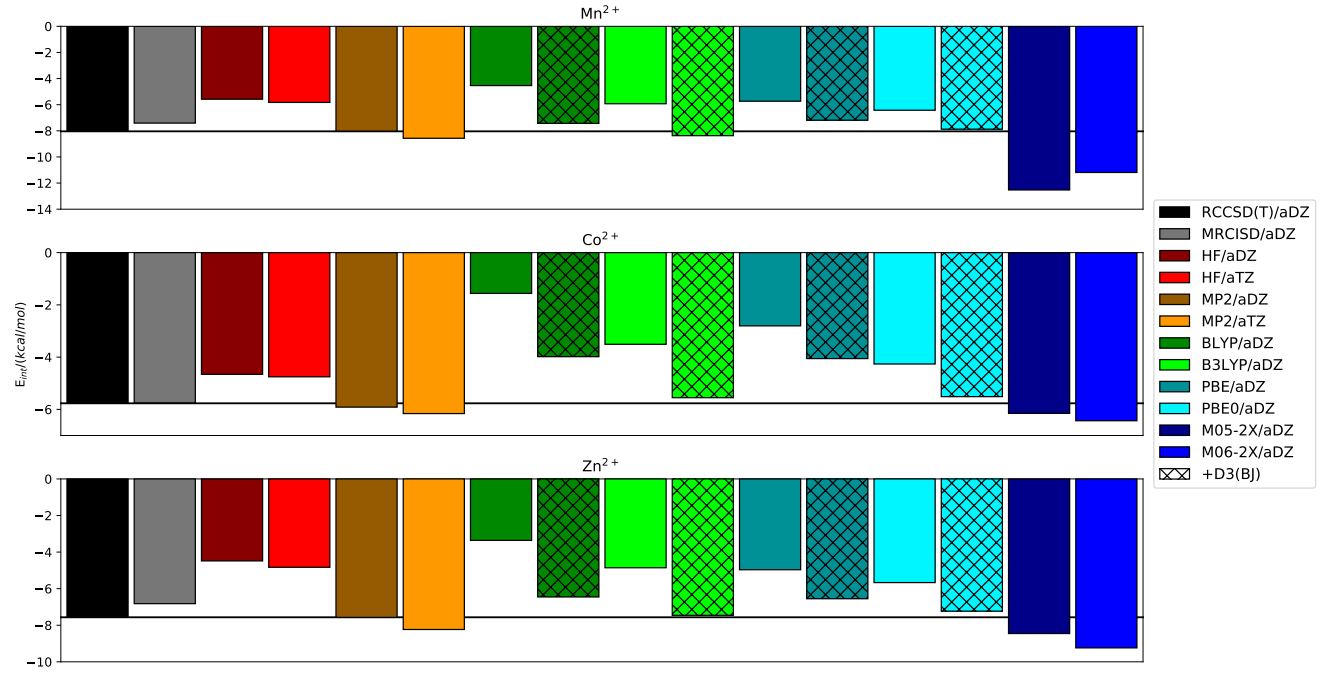


Fig. 5 Counterpoise corrected interaction energies (kcal/mol) for paddlewheel MOF model-CO_2 systems.

the Zn^{2+}-BDC system had to be replaced by RCCSD due to a highly suspicious result for the triples correction (the discarded RCCSD(T) interaction energy was −14.97 kcal/mol). The non-CC results in Tables 4 and 5 generally show varying but reasonable agreement with the CC results. Figures 6 and 7 show the DFT+D values of E_{int}^{CP} for the BDC and FDC systems, which are similar to the previous systems. The dispersion corrections are very small and imply that the interactions of these systems are not highly dispersion dependent. Given the obvious dipole character of these MOF models, it would stand to reason that electrostatic and in-

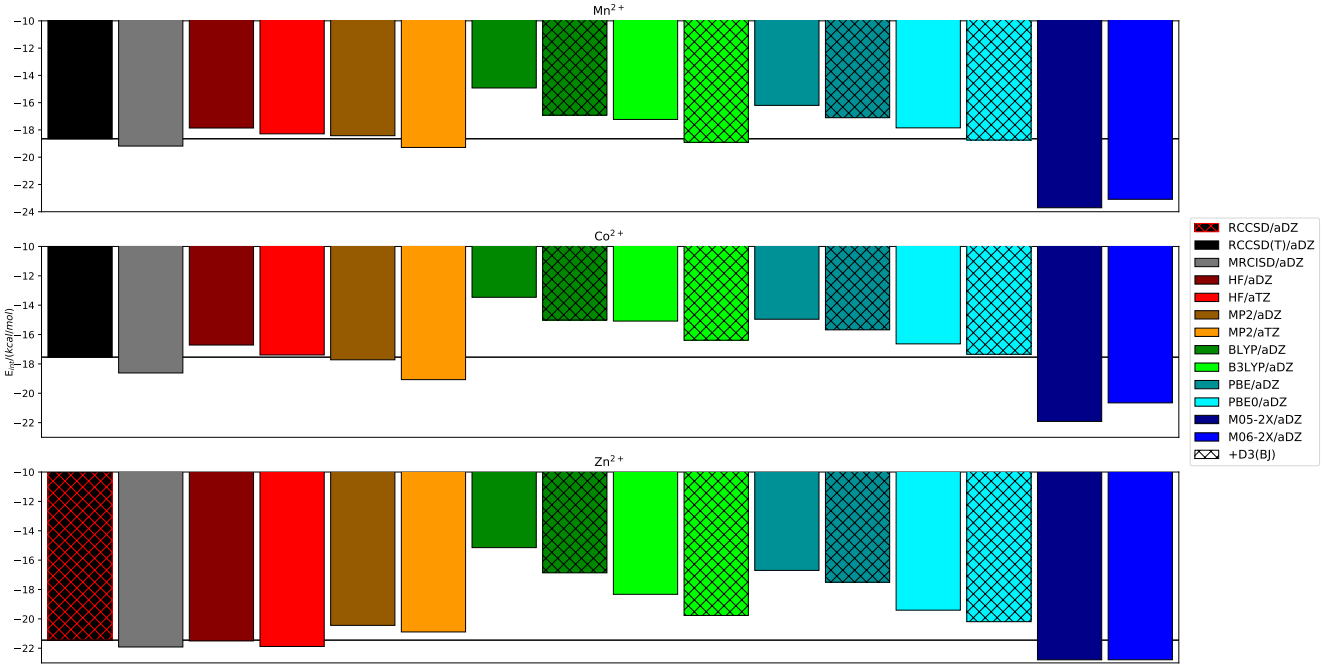


Fig. 6 Counterpoise corrected interaction energies (kcal/mol) for benzenedicarboxylate MOF model-CO_2 systems.

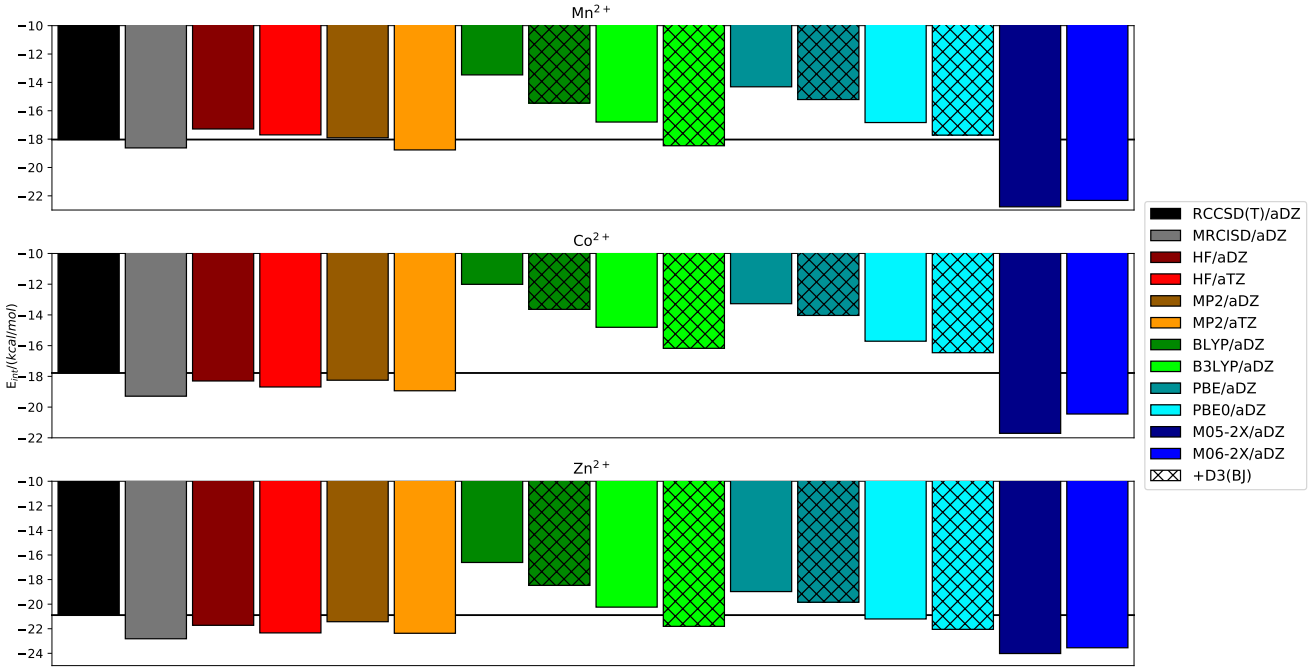


Fig. 7 Counterpoise corrected interaction energies (kcal/mol) for furandicarboxylate MOF model-CO_2 systems.

duction forces would contribute more to the overall interaction. We will revisit this issue while discussing the SAPT calculations. Despite their limited effect, the addition of the atom-pairwise -D3 corrections does improve the agreement of the DFT results with the benchmark values.

Tables 6–9 present the CASSCF/aDZ natural orbital occupation numbers of the diformate, BDC, FDC, and paddlewheel systems, respectively. The results show that these systems have only small amounts of multireference character, with the greatest deviations from the expected single reference occupation values (2,1, or 0)

Table 3 Comparison of E_{int}^{CP} (Eq. (1)) and E_{int}^{CP*} (Eq. (2)) for the M^{2+}-paddlewheel-CO_2 model systems (kcal/mol).

Metal	Mn^{2+}		Co^{2+}		Zn^{2+}	
Method	E_{int}^{CP}	E_{int}^{CP*}	E_{int}^{CP}	E_{int}^{CP*}	E_{int}^{CP}	E_{int}^{*CP}
RHF	-5.58	-5.58	-4.66	-4.66	-4.48	-4.48
MP2	-8.02	-8.02	-5.91	-5.91	-7.57	-7.57
RCCSD(T)	-8.00		-5.77		-7.57	
(*2n*,12)			Active Space		(3d4s)	
CASSCF	-3.34	-5.58	-4.57	-4.57	-6.69	-6.78
NEVPT2	-12.18	-7.76	-5.74		-19.81	-20.11
MRCISD	67.91	-7.41	72.62	-5.75	75.14	-6.82

being around 0.01. Though the deviations are small, they do correlate well with the observed trends in the multireference calculations. For the diformate, BDC, and FDC systems, the Co^{2+} NOONs show the largest deviations from a single reference character with the Zn^{2+} ones following after.

The flexible interaction energies $E_{int}^{CP}+\Delta E_{int}^{flex}$ (Eq. (3)) in Table 10 lead to similar conclusions as the rigid interaction energies. There are no Co^{2+} values here because the fully optimized structures were not considered for this system. The BDC ΔE_{int}^{flex} corrections are the most considerable, especially for the Mn^{2+} system. The change in the MOF fragment geometry contributes the majority of these corrections, with CO_2 being more limited in its possibility for deformation. The limited change in $R_{Ligand-M^{2+}}$ for this model in its bound and unbound states also points to the change in the BDC structure as the largest contributor to this correction. None of the other ΔE_{int}^{flex} results constitute a change greater than 1 kcal/mol, with the FDC systems exhibiting the next largest changes. The minuscule diformate and paddlewheel ΔE_{int}^{flex} results agree with the smaller E_{int}^{CP} values for these systems. These structures are also more rigid in their configuration and have ligands that have less possibility for deformation. The important geometric parameters for all systems can be seen in Table 1. The changes in these parameters between the optimized dimer and monomer systems are generally small, with the exception of the Co^{2+}-BDC and Co^{2+}-FDC $R_{Ligand-M^{2+}}$ values. The Co^{2+}-FDC Ligand-M^{2+}-CO_2 angle also differs from the corresponding values for the other metals, which was also seen in the full optimization that was previously discussed. For the Mn^{2+}-BDC system, the changes that lead to the value of ΔE_{int}^{flex} appear to be the O-C-O angles of the carboxylate groups. Both angles contract by about 1.5° in the dimer structure. A similar occurence is seen in the Zn^{2+}-BDC system, though the angles expand and only by about 0.5°.

Potential Energy Curves

The dependence of E_{int}^{CP} on intermolecular separation can be seen in Figures 8, 9, and 10 for models containing Mn^{2+}, Co^{2+}, and Zn^{2+}, respectively. For the BDC and FDC systems, the dispersion-uncorrected DFT curves give quantitatively similar results to RCCSD(T) and MP2. As seen in the single point calculations for these systems, the dispersion corrections lead to a generally small but beneficial change in the accuracy of the B3LYP fucntional. The results for the diformate and paddlewheel models are also similar to each other, with the addition of the empirical dispersion to DFT being necessary to obtain near-RCCSD(T) accuracy. The CASSCF results for the paddlewheel systems, lacking dynamical correlation, do not perform any better than non-D3 DFT. It is worth restating that the geometries used for these curves are constructed from rigid monomers and therefore may differ from the results of the previous subsection, which used the flexible geometries. Also, z is defined by the rigidly optimized total dimer energy, not interaction energy, and z=1 does not necessarily correspond to the CP-corrected interaction energy minimum.

SAPT Calculations

The results of the closed-shell SAPT energy decomposition for each of the Zn^{2+}-containing systems are found in Table 11. As has been the case for the preceding results, the BDC and FDC models are similar in their trends. As inferred by the small effect of the empirical dispersion correction on the DFT results, the decomposition of the interaction energy shows that induction has a much greater effect on the interactions of these models than dispersion. Dispersion accounts for around 20% of the interaction energy for these systems, while induction accounts for approximately 60%. Electrostatics plays a larger role in the SAPT0 results, where the exchange term is noticeably smaller in magnitude than its attractive electrostatic counterpart. With the move to SAPT2+3 and the accompanying jump in the exchange values, electrostatics is mostly canceled out in the BDC model and completely canceled out in the FDC model. The diformate and paddlewheel results also agree with the role of dispersion noted in analyzing the DFT data. In these models, dispersion plays a greater role in the interactions, being similar in magnitude to induction in the paddlewheel and greater than induction in the diformate model. Together, induction and dispersion account for nearly 120% of the total binding energy, overcoming the repulsive first-order contribution as exchange overpowers electrostatics in these models. At the SAPT2+3 level, exchange is sufficient in magnitude to effectively cancel out electrostatics and induction (or dispersion in the case of the paddlewheel structure) at the minimum intermolecular separation.

For the SAPT0 calculations in aDZ and aTZ, the rate of convergence with regard to basis set size can be inferred. The difference in the total SAPT interaction energies is around 1 kcal/mol in all cases, with the FDC model having the largest difference in consecutive basis sets at 1.21 kcal/mol. Regarding the convergence of individual components, the exchange contribution shows the least change. The electrostatics and dispersion generally contribute the most to the overall change, with induction appearing to be fairly well converged in the smaller basis. The comparison between the results at the SAPT0 and SAPT2+3 levels shows a much greater change. The exchange contributions increase considerably in each system with the addition of intramolecular correlation while induction and dispersion decrease only slightly and electrostatics increases slightly. These changes result in a total energy that is more positive than at the lower level of theory, though only by about 1 kcal/mol at most. The SAPT2+3 results are consistently around 2 kcal/mol more negative than the RCCSD(T) or

Table 4 Comparison of E_{int}^{CP} (Eq. (1)) and E_{int}^{CP*} (Eq. (2)) for the M^{2+}-BDC-CO_2 model systems (kcal/mol).

Metal	Mn^{2+}		Co^{2+}		Zn^{2+}	
Method	E_{int}^{CP}	E_{int}^{CP*}	E_{int}^{CP}	E_{int}^{CP*}	E_{int}^{CP}	E_{int}^{CP*}
RHF	-17.86	-17.86	-16.71	-16.71	-21.51	-21.51
MP2	-18.42	-18.42	-17.72	-17.72	-20.45	-20.45
RCCSD(T)	-18.65		-17.54		-21.47[a]	
(n,6)			Active Space		(3d4s)	
CASSCF	-17.86	-17.86	-17.12	-17.12	-21.31	-21.31
NEVPT2	-18.02	-18.02	-18.35	-18.35	-6.58	-6.58
MRCISD	45.97	-19.17	48.36	-18.52	46.31	-22.07
(n,9)			Active Space		(3d4s4p)	
CASSCF	-17.67	-17.67	-16.82	-16.82	-21.01	-21.01
NEVPT2	-14.29	-14.29	-12.18	-12.18	-17.99	-17.99
MRCISD	45.97	-19.18	48.25	-18.62	44.85	-21.92

a RCCSD value replacing the RCCSD(T) one with a suspect triples contribution (see text).

Table 5 Comparison of E_{int}^{CP} (Eq. (1)) and E_{int}^{CP*} (Eq. (2)) for the M^{2+}-FDC-CO_2 model systems (kcal/mol).

Metal	Mn^{2+}		Co^{2+}		Zn^{2+}	
Method	E_{int}^{CP}	E_{int}^{CP*}	E_{int}^{CP}	E_{int}^{CP*}	E_{int}^{CP}	E_{int}^{CP*}
RHF	-17.28	-17.28	-18.30	-18.30	-21.73	-21.73
MP2	-17.91	-17.91	-18.24	-18.24	-21.44	-21.44
RCCSD(T)	-18.03		-17.78		-20.89	
(n,6)			Active Space		(3d4s)	
CASSCF	-17.28	-17.28	-18.48	-18.48	-21.50	-21.53
NEVPT2	-17.50	-17.50	-17.87	-17.87	-8.88	-21.02
MRCISD	43.58	-18.61	44.69	-19.29	42.74	-22.95
(n,9)			Active Space		(3d4s4p)	
CASSCF	-17.09	-17.09	-18.09	-18.09	-21.21	-21.21
NEVPT2	-13.76	-13.76	-20.62	-20.62	-17.71	
MRCISD	43.16	-18.62	44.69	-19.29	41.14	-22.81

MRCISD benchmarks for each system.

Tables 12 and 13 contain the SAPT0 results in aDZ and aTZ bases for the Mn^{2+} models based on UHF and ROHF, respectively, while Table 14 contains the ROHF based results for the Co^{2+} models. The blanks in Table 14 are due to convergence issues. For Mn^{2+}, the UHF and ROHF results show good agreement with each other. For a given basis set, the electrostatic, exchange, and induction terms are almost identical between the two methods. The only considerable change occurs in the dispersion term, which is always larger in the ROHF based results. All SAPT components for the Mn^{2+} and Co^{2+} systems, where applicable, are again indifferent to the basis set choice.

The trends seen in the SAPT0 results for the Zn^{2+} systems are mostly repeated for each model, though there are some changes associated with the change in metal. The BDC model has a notable change in results between Zn^{2+} and Mn^{2+}. The difference in the electrostatic contribution between the Zn^{2+} system and the Mn^{2+} one is around 4 kcal/mol, while the difference in the exchange terms is close to -9 kcal/mol. The Zn^{2+}-BDC exchange term at the SAPT0 level is around half that of the Mn^{2+}-BDC system. The differences between the induction and dispersion terms, 1.92 and 1.27 kcal/mol respectively, negate some of the differences between the total interaction energies, leaving the total interaction energy in the Zn^{2+} system only 2.19 kcal/mol more negative than in the Mn^{2+} one. The differences, as well as the relatively large magnitudes, of the electrostatic and induction terms are supported by the PBE0/aDZ dipole moments of the MOF fragments that appear in Table 15. The BDC and FDC systems exhibit large dipole moments, while the diformate and paddlewheel systems are effectively non-dipolar. For BDC, the slight differences in the z-component of the dipole in the Zn^{2+} and Mn^{2+} systems support the more negative electrostatic and induction terms of Mn^{2+} over Zn^{2+}. In the FDC systems, all of the attractive terms are slightly more negative for Zn^{2+} than Mn^{2+}, and the exchange energy is only slightly more positive. Again, the dipole moments explain the relative sizes of the electrostatic and induction terms between these systems. In this case, the relationship between Co^{2+} and the other two metals can also be seen in the dipole moments. The smaller dipole moment of Co^{2+}-containing model corresponds to less negative electrostatic and induction terms, but the exchange term is not as positive, adding up to a lower total interaction energy.

The same kind of lower exchange value seen in Co^{2+}-FDC is also observed in the Co^{2+}-paddlewheel system, where it compensates for the electrostatic and induction terms being considerably weaker for this metal than for the other two. In the diformate

Table 6 CASSCF/aDZ natural orbital occupation numbers for the active space orbitals of M^{2+}-diformate-CO_2. Numbers outside of parentheses are for the (n,9) active space, and numbers in parentheses are for the (n,6) active space. The color of the orbital index cell denotes the dominant character of the orbital in the following way: green for s, red for p, blue for d.

Orbital	Mn^{2+}	Co^{2+}	Zn^{2+}
20.1	0.99678 (1.00000)	1.98903 (1.99990)	1.99153 (1.99999)
21.1	0.99626 (1.00000)	1.98753 (1.99305)	1.99058 (1.99672)
22.1	0.00376 (0.00000)	0.01233 (0.00705)	0.00948 (0.00339)
23.1	0.00321	0.01072	0.00847
12.2	0.99646 (1.00000)	0.99553 (1.00000)	1.99103 (1.99999)
13.2	0.00355	0.00470	0.00902
10.3	0.99643 (1.00000)	0.99544 (1.00000)	1.99099 (1.99993)
11.3	0.00359	0.00479	0.00904
6.4	0.99996 (1.00000)	0.99995 (1.00000)	1.99985 (1.99998)

Table 7 CASSCF/aDZ natural orbital occupation numbers for the active space orbitals of M^{2+}-BDC-CO_2. Numbers outside of parentheses are for the (n,9) active space, and numbers in parentheses are for the (n,6) active space. The color of the orbital index cell denotes the dominant character of the orbital in the following way: green for s, red for p, blue for d.

Orbital	Mn^{2+}	Co^{2+}	Zn^{2+}
34.1	0.99676 (1.00000)	1.98897 (1.99995)	1.99153 (2.00000)
35.1	0.99624 (1.00000)	1.98755 (1.99232)	1.99105 (1.99627)
36.1	0.00378 (0.00000)	0.01231 (0.00773)	0.00900 (0.00386)
37.1	0.00323	0.01080	0.00847
9.2	0.99640 (1.00000)	0.99550 (1.00000)	1.99111 (1.99992)
10.2	0.00362	0.00473	0.00893
20.3	0.99646 (1.00000)	0.99573 (1.00000)	1.99127 (1.99999)
21.3	0.00355	0.00445	0.00878
4.4	0.99996 (1.00000)	0.99996 (1.00000)	1.99985 (1.99997)

systems, the balance between electrostatics and exchange is similar for all three metals. The dispersion terms in the diformate and paddlewheel systems, though weaker for Co^{2+}, are of similar magnitude. Between the different metals in the diformate model, the difference in the electrostatic terms is nearly canceled out by the difference in the exchange terms, the dispersion difference is small, and the induction difference consistutes about 66% of the total difference in interaction energy. The total interaction energies of the Zn^{2+}-paddlewheel and Mn^{2+}-paddlewheel systems differ by less than 0.1 kcal/mol because of the interplay in the differences of the SAPT components. The differences between the electrostatic and dispersion terms favor Zn^{2+}, but the exchange difference favors Mn^{2+}. The induction terms are almost equal. In general, the SAPT results in this work overestimate their benchmarks by around 2-5 kcal/mol. SAPT does a good job of replicating the metal preference seen in the RCCSD(T) results for the diformate and paddlewheel systems. For the FDC systems, SAPT finds correctly that the interaction of the Zn^{2+} model with CO_2 is notably more attractive than for the other two metals, but flips the energetic ordering for Mn^{2+}-FDC and Co^{2+}-FDC.

Conclusions

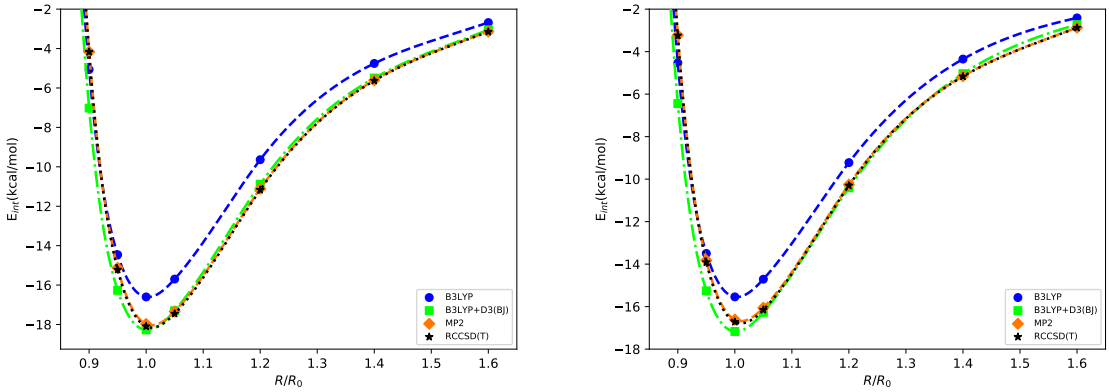
In this work, *ab initio* benchmark interaction energy values, including RCCSD(T) and MRCISD, were obtained for model fragments of metal-organic frameworks interacting with CO_2. These models contained either Mn^{2+}, Co^{2+}, or Zn^{2+} as representatives of systems with half filled, intermediately filled, and completely filled d shell cations, respectively. The potential importance of a multireference treatment of these systems was investigated at the CASSCF and post-CASSCF levels and considered the effects of active space size. A selection of DFT functionals was compared to these benchmarks, both with and without the empirical dispersion correction. The effect of intermolecular separation was also considered for selected levels of theory. Last but not least, SAPT interaction energy decompositions were performed to examine the importance of individual interaction types for these systems.

The SAPT results obtained on the MOF model systems corroborate the observed importance of the empirical dispersion corrections for the DFT results. These corrections improve the overall results slightly for BDC and FDC model systems, while providing considerable improvement to the diformate and paddlewheel systems. Additionally, the hybrid functionals perform much better than the GGAs and PBE0 generally outperforms B3LYP. The selection of damping parameters for the dispersion correction was, in general, less important than the mere addition of this correction. The M05-2X and M06-2X functionals tend to overestimate the RCCSD(T) values, sometimes quite considerably.

The consideration of the multireference nature of the models showed consistently that the Mn^{2+}-containing systems were the least affected by the switch from HF to CASSCF, while Co^{2+}-containing systems showed moderate change and Zn^{2+}-containing ones were the most affected, but still the multireference character was small as indicated by the NOONs. For the paddlewheel-Zn^{2+} system, the change from HF to CASSCF was greater than 2 kcal/mol and CASSCF proved to be a good approximation to MRCISD, though both methods fall short of the

Table 8 CASSCF/aDZ natural orbital occupation numbers for the active space orbitals of M^{2+}-FDC-CO_2. Numbers outside of parentheses are for the (n,9) active space, and numbers in parentheses are for the (n,6) active space. The color of the orbital index cell denotes the dominant character of the orbital in the following way: green for s, red for p, blue for d.

Orbital	Mn^{2+}	Co^{2+}	Zn^{2+}
50.1	0.99676 (1.00000)	1.98879 (1.99922)	1.99150 (1.99999)
51.1	0.99645 (1.00000)	1.98760 (1.99155)	1.99124 (1.99999)
52.1	0.99624 (1.00000)	0.99574 (1.00077)	1.99099 (1.99618)
53.1	0.00378 (0.00000)	0.01198 (0.00846)	0.00906 (0.00395)
54.1	0.00356	0.01077	0.00881
55.1	0.00324	0.00443	0.00850
12.2	0.99996 (1.00000)	1.00022 (1.00000)	1.99984 (1.99996)
13.2	0.99639 (1.00000)	0.99571 (1.00000)	1.99108 (1.99992)
14.2	0.00362	0.00475	0.00897



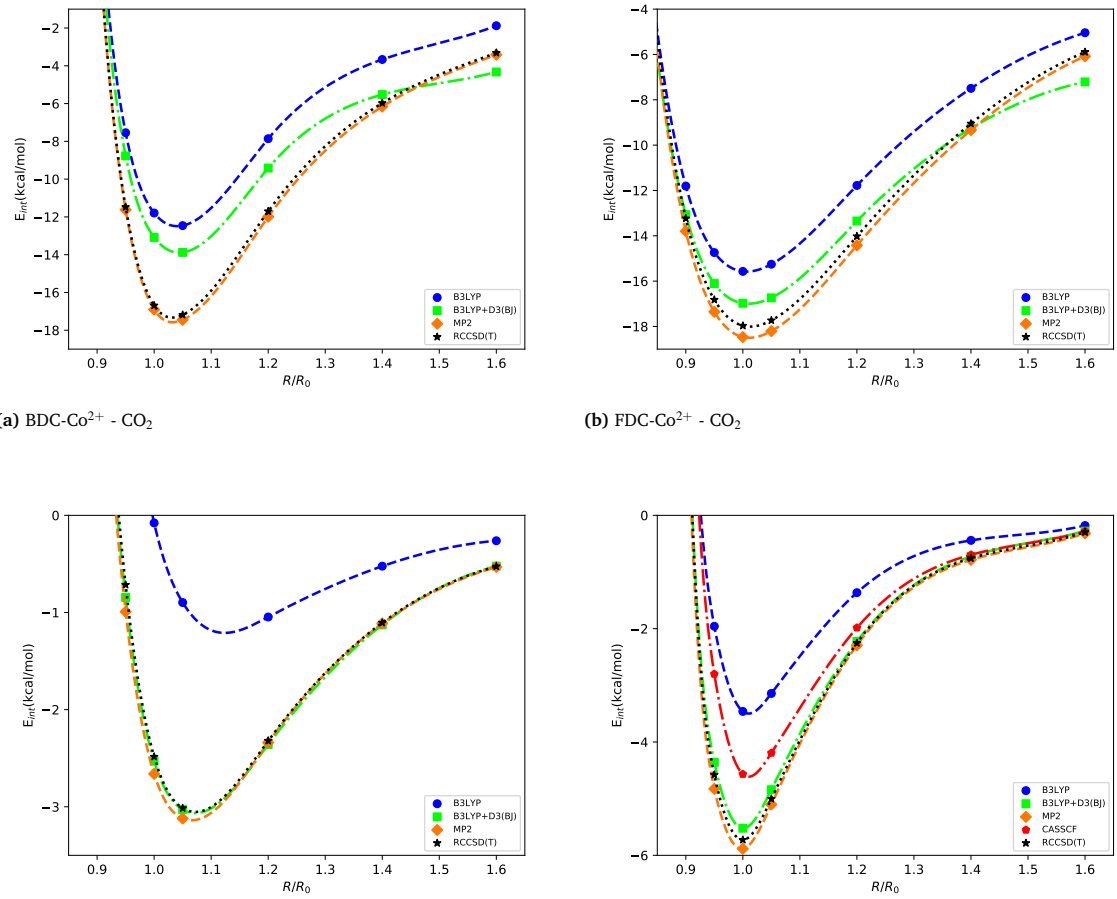
(a) BDC-Mn^{2+} - CO_2 **(b)** FDC-Mn^{2+} - CO_2

(c) $(HCOO^-)_2Mn^{2+}$ - CO_2 **(d)** $(HCOO^-)_4(Mn^{2+})_2$ - CO_2

Fig. 8 Interaction energy curves (kcal/mol) for rigid dimer structures containing Mn^{2+} as a function of the intermolecular separation for DFT, DFT+D, and selected wavefunction methods. Distances are represented relative to the optimal separation distance R_0. The values of R_0 for the BDC, FDC, diformate, and paddlewheel models are 3.312, 3.321, 3.595, and 4.999 angstroms, respectively.

RCCSD(T) result by about 1 kcal/mol. For the MRCISD results, the interaction energy based on the infinitely separated reference overcomes the issues of size inconsistency and provides more meaningful results than the typical supermolecular definition of the interaction energy while it is still possible to retain the CP correction for the basis set superposition error. MRCISD showed results similar to RCCSD(T), but MP2 was consistently closer to the benchmark RCCSD(T) values.

The SAPT decomposition shows that the binding in the BDC and FDC systems is induction dominated, as expected from the dipole character of these systems, while the diformate and paddlewheel systems show a more even mix of induction and disper-



(a) BDC-Co^{2+} - CO_2 **(b)** FDC-Co^{2+} - CO_2

(c) $(HCOO^-)_2Co^{2+}$ - CO_2 **(d)** $(HCOO^-)_4(Co^{2+})_2$ - CO_2

Fig. 9 Interaction energy curves (kcal/mol) for rigid dimer structures containing Co^{2+} as a function of the intermolecular separation for DFT, DFT+D, and selected wavefunction methods. Distances are represented relative to the optimal separation distance R_0. The values of R_0 for the BDC, FDC, diformate, and paddlewheel models are 3.182, 2.070, 3.677, and 5.110 angstroms, respectively.

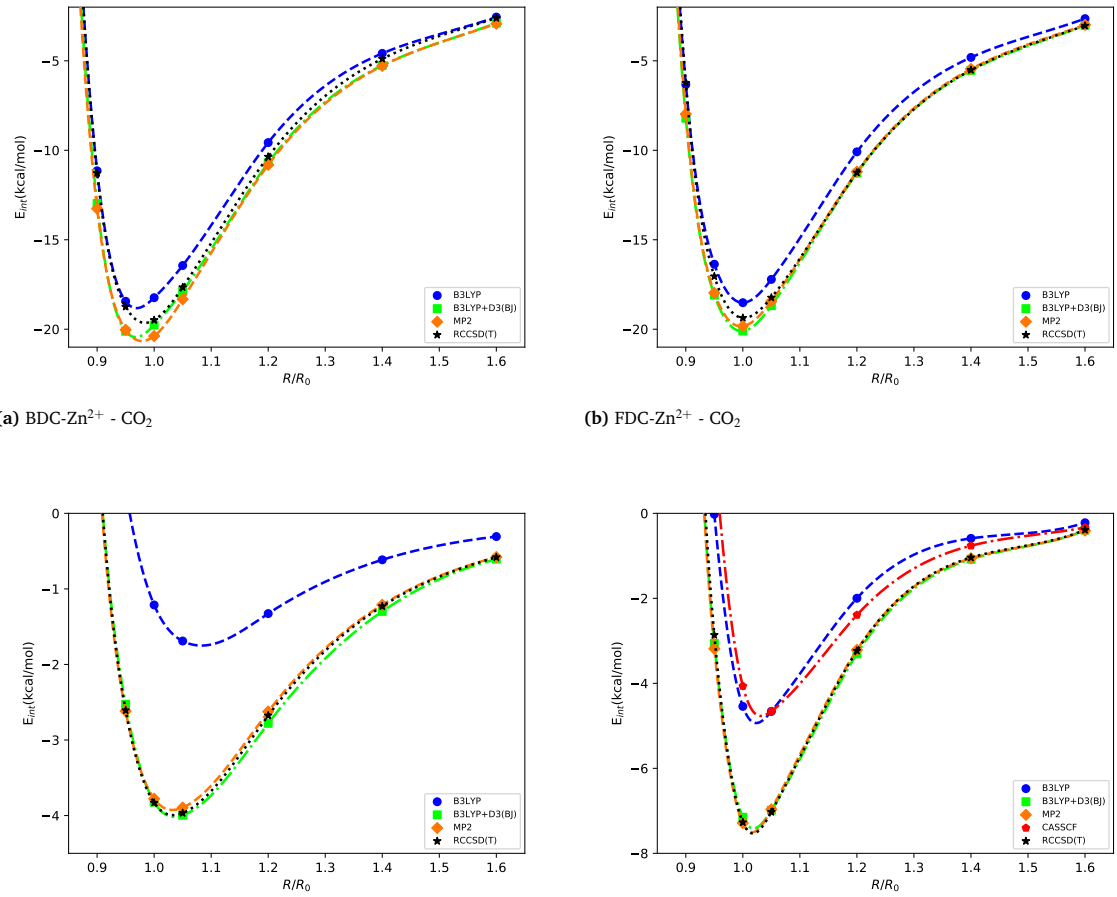
Table 9 CASSCF/aDZ natural orbital occupation numbers for the active space orbitals of M^{2+}-paddlewheel-CO_2 in the (n,12) active space. The color of the orbital index cell denotes the dominant character of the orbital in the following way: green for s, blue for d.

Orbital	Mn^{2+}	Co^{2+}	Zn^{2+}
38.1	1.00000	1.99353	1.99997
39.1	1.00000	1.99243	1.99996
40.1	1.00000	1.00000	1.99827
41.1	1.00000	1.00000	1.99401
42.1	0.00000	0.00799	0.00567
43.1	0.00000	0.00608	0.00256
19.2	1.00000	1.00000	1.99999
20.2	1.00000	1.00000	1.99980
19.3	1.00000	1.00000	1.99999
20.3	1.00000	1.00000	1.99979
5.4	1.00000	1.99998	2.00000
6.4	1.00000	1.99998	1.99999

Table 10 Monomer relaxation corrections to interaction energy, ΔE_{int}^{flex} (kcal/mol), for the Mn^{2+} and Zn^{2+} containing systems. B3LYP/aDZ was used for the BDC and FDC systems, MP2/aDZ for the paddlewheel systems, and MP2/aTZ for the diformate systems.

	Systems			
Metal	BDC	FDC	Diformate	Paddlewheel
Mn^{2+}	4.40	0.55	0.26	0.09
Zn^{2+}	1.41	0.81	0.08	0.15

sion. For the three systems for which we were able to converge the Co^{2+} SAPT0 calculations, the induction and dispersion energies are less attractive than for the Mn^{2+} and Zn^{2+}-containing counterparts, but the total binding energy is only slightly less attractive thanks to the reduction in the exchange repulsion. The ROHF and UHF-based open-shell SAPT results have good agreement, but the UHF-based variant provides consistently more negative dispersion values for all systems. The SAPT total energies overestimate the benchmarks, but the energy decomposition provides useful insight and replicates the key features of the metal preference order for each system.



(a) BDC-Zn^{2+} - CO_2

(b) FDC-Zn^{2+} - CO_2

(c) $(HCOO^-)_2Zn^{2+}$ - CO_2

(d) $(HCOO^-)_4(Zn^{2+})_2$ - CO_2

Fig. 10 Interaction energy curves (kcal/mol) for rigid dimer structures containing Zn^{2+} as a function of the intermolecular separation for DFT, DFT+D, and selected wavefunction methods. Distances are represented relative to the optimal separation distance R_0. The values of R_0 for the BDC, FDC, diformate, and paddlewheel models are 3.257, 3.206, 3.600, and 4.714 angstroms, respectively.

Conflicts of interest

There are no conflicts to declare.

Acknowledgments

This work was supported by the U.S. National Science Foundation CAREER Award No. CHE-1351978. The authors thank Dr. Piotr Żuchowski and Dr. Evangelos Miliordos for valuable discussions.

Notes and references

1 J.-R. Li, J. Sculley and H.-C. Zhou, *Chem. Rev.*, 2012, **112**, 869–932.

2 S. O. Odoh, C. J. Cramer, D. G. Truhlar and L. Gagliardi, *Chem. Rev.*, 2015, **115**, 6051–6111.

3 J. Lee, O. K. Farha, J. Roberts, K. A. Scheidt, S. T. Nguyen and J. T. Hupp, *Chem. Soc. Rev.*, 2009, **38**, 1450–1459.

4 L. J. Murray, M. Dinca and J. R. Long, *Chem. Soc. Rev.*, 2009, **38**, 1294–1314.

5 L.-C. Lin, A. H. Berger, R. L. Martin, J. Kim, J. A. Swisher, K. Jariwala, C. H. Rycroft, A. S. Bhown, M. W. Deem, M. Haranczyk and B. Smit, *Nat. Mater.*, 2012, **11**, 633–641.

6 A. O. Yazaydin, R. Q. Snurr, T.-H. Park, K. Koh, J. Liu, M. D. LeVan, A. I. Benin, P. Jakubczak, M. Lanuza, D. B. Galloway, J. J. Low and R. R. Willis, *J. Am. Chem. Soc.*, 2009, **131**, 18198–18199.

7 J. Yu, L.-H. Xie, J.-R. Li, Y. Ma, J. M. Seminario and P. B. Balbuena, *Chem. Rev.*, 2017, **117**, 9674–9754.

8 A. K. Rappé and E. R. Bernstein, *J. Phys. Chem. A*, 2000, **104**, 6117–6128.

9 S. F. Sousa, P. A. Fernandes and M. J. Ramos, *J. Phys. Chem. A*, 2007, **111**, 10439–10452.

10 B. Jeziorski, R. Moszynski and K. Szalewicz, *Chem. Rev.*, 1994, **94**, 1887–1930.

11 K. Szalewicz, K. Patkowski and B. Jeziorski, *Struct. Bonding (Berlin)*, 2005, **116**, 43–117.

12 E. G. Hohenstein and C. D. Sherrill, *WIREs Comput. Mol. Sci.*, 2012, **2**, 304–326.

13 J. G. McDaniel and J. Schmidt, *J. Phys. Chem. A*, 2013, **117**,

Table 11 SAPT decomposition (kcal/mol) for complexes containing Zn^{2+}.

Method	Basis Set	electrostatics	exchange	induction	dispersion	total
			BDC			
SAPT0	AVDZ	-16.28	9.46	-14.69	-3.42	-24.92
	AVTZ	-16.58	9.46	-14.76	-3.76	-25.64
SAPT2+3	AVDZ	-15.12	13.06	-16.99	-5.21	-24.27
			FDC			
SAPT0	AVDZ	-22.27	19.52	-18.98	-4.96	-26.68
	AVTZ	-22.73	19.50	-19.11	-5.55	-27.89
SAPT2+3	AVDZ	-21.46	25.54	-22.25	-7.46	-25.63
			Diformate			
SAPT0	AVDZ	-7.40	8.57	-3.13	-4.79	-6.74
	AVTZ	-7.55	8.56	-3.19	-5.19	-7.36
SAPT2+3	AVDZ	-7.41	11.76	-3.85	-5.93	-5.44
			Paddlewheel			
SAPT0	AVDZ	-11.94	13.71	-6.24	-5.98	-10.46
	AVTZ	-12.21	13.70	-6.31	-6.48	-11.30
SAPT2+3	AVDZ	-12.20	18.05	-7.57	-7.63	-9.35

Table 12 UHF-SAPT decomposition (kcal/mol) for complexes containing Mn^{2+}.

Method	Basis Set	electrostatics	exchange	induction	dispersion	total
			BDC			
SAPT0	AVDZ	-20.16	18.83	-16.52	-4.31	-22.16
	AVTZ	-20.58	18.85	-16.68	-5.03	-23.45
			FDC			
SAPT0	AVDZ	-19.85	18.63	-16.05	-4.30	-21.56
	AVTZ	-20.25	18.68	-16.11	-5.03	-22.71
			Diformate			
SAPT0	AVDZ	-8.97	10.00	-4.17	-4.73	-7.86
	AVTZ	-9.14	10.04	-4.26	-5.23	-8.59
			Paddlewheel			
SAPT0	AVDZ	-11.48	12.10	-6.20	-4.65	-10.22
	AVTZ	-11.71	12.16	-6.27	-5.17	-10.99

2053–2066.

14 G. Gryn'ova and C. Corminboeuf, *J. Phys. Chem. Lett.*, 2016, **7**, 5198–5204.

15 L. Grajciar, O. Bludský and P. Nachtigall, *J. Phys. Chem. Lett.*, 2010, **1**, 3354–3359.

16 O. Bludský, M. Rubeš, P. Soldán and P. Nachtigall, *J. Chem. Phys.*, 2008, **128**, 114102.

17 M. Rubeš, J. Kysilka, P. Nachtigall and O. Bludský, *Phys. Chem. Chem. Phys.*, 2010, **12**, 6438–6444.

18 S. Grimme, *J. Comput. Chem.*, 2006, **27**, 1787–1799.

19 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.*, 2010, **132**, 154104.

20 S. Grimme, S. Ehrlich and L. Goerigk, *J. Comput. Chem.*, 2011, **32**, 1456–1465.

21 L. Grajciar, A. D. Wiersum, P. L. Llewellyn, J.-S. Chang and P. Nachtigall, *J. Phys. Chem. C*, 2011, **115**, 17925–17933.

22 J. D. Howe, Y. Liu, L. Flores, D. A. Dixon and D. S. Sholl, *J. Chem. Theory Comput.*, 2017, **13**, 1341–1350.

23 A. L. Dzubak, L.-C. Lin, J. Kim, J. A. Swisher, R. Poloni, S. N. Maximoff, B. Smit and L. Gagliardi, *Nat. Chem.*, 2012, **4**, 810–816.

24 J. J. Goings, S. M. Ohlsen, K. M. Blaisdell and D. P. Schofield, *J. Phys. Chem. A*, 2014, **118**, 7411–7417.

25 V. Sladek and I. Tvaroška, *J. Phys. Chem. B*, 2017, **121**, 6148–6162.

26 H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, P. Celani, T. Korona, R. Lindh, A. Mitrushenkov, G. Rauhut, K. R. Shamasundar, T. B. Adler, R. D. Amos, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, E. Goll, C. Hampel, A. Hesselmann, G. Hetzer, T. Hrenar, G. Jansen, C. Köppl, Y. Liu, A. W. Lloyd, R. A. Mata, A. J. May, S. J. McNicholas, W. Meyer, M. E. Mura, A. Nicklass, D. P. O'Neill, P. Palmieri, D. Peng, K. Pflüger, R. Pitzer, M. Reiher, T. Shiozaki, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson and M. Wang, *MOLPRO, version 2012.1, a package of ab initio programs*, 2012, http://www.molpro.net.

27 H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby and M. Schütz, *WIREs Comput Mol Sci*, 2012, **2**, 242–253.

28 R. M. Parrish, L. A. Burns, D. G. A. Smith, A. C. Simmonett, A. E. DePrince, E. G. Hohenstein, U. Bozkaya, A. Y. Sokolov, R. Di Remigio, R. M. Richard, J. F. Gonthier, A. M. James,

Table 13 ROHF-SAPT decomposition (kcal/mol) for complexes containing Mn^{2+}.

Method	Basis Set	electrostatics	exchange	induction	dispersion	total
			BDC			
SAPT0	AVDZ	-20.21	18.84	-16.49	-4.78	-22.64
	AVTZ	-20.61	18.87	-16.55	-5.56	-23.85
			FDC			
SAPT0	AVDZ	-19.91	18.63	-16.01	-4.76	-22.05
	AVTZ	-20.30	18.68	-16.08	-5.54	-23.24
			Diformate			
SAPT0	AVDZ	-8.96	9.92	-4.14	-4.97	-8.14
	AVTZ	-9.12	9.96	-4.22	-5.50	-8.89
			Paddlewheel			
SAPT0	AVDZ	-11.48	12.05	-6.16	-5.00	-10.58
	AVTZ	-11.71	12.10	-6.22	-5.55	-11.38

Table 14 ROHF-SAPT decomposition (kcal/mol) for complexes containing Co^{2+}.

Method	Basis Set	electrostatics	exchange	induction	dispersion	total
			BDC			
SAPT0	AVDZ					
	AVTZ					
			FDC			
SAPT0	AVDZ					
	AVTZ	-18.82	15.39	-15.26	-4.96	-23.65
			Diformate			
SAPT0	AVDZ	-6.85	7.92	-2.61	-4.28	-5.81
	AVTZ	-6.94	7.94	-2.66	-4.67	-6.32
			Paddlewheel			
SAPT0	AVDZ	-7.72	6.57	-3.51	-3.66	-8.32
	AVTZ	-7.83	6.61	-3.53	-3.96	-8.71

Table 15 Components of the PBE0/aDZ dipole moments (a.u.) of the BDC and FDC models with various metals. The BDC models are oriented so the z-axis is the axis of interaction in the dimer. For the FDC models, the XZ plane is the mirror plane of the system. The diformate and paddlewheel models are excluded because they are non-dipolar.

Metal	X	Y	Z
		BDC	
Mn^{2+}	0.00	0.00	11.76
Co^{2+}	0.00	0.00	10.72
Zn^{2+}	0.00	0.00	11.33
		FDC	
Mn^{2+}	2.21	0.00	9.08
Co^{2+}	3.04	0.00	7.50
Zn^{2+}	2.34	0.00	9.15

H. R. McAlexander, A. Kumar, M. Saitow, X. Wang, B. P. Pritchard, P. Verma, H. F. Schaefer, K. Patkowski, R. A. King, E. F. Valeev, F. A. Evangelista, J. M. Turney, T. D. Crawford and C. D. Sherrill, *J. Chem. Theory Comput.*, 2017, **13**, 3185–3197.

29 T. H. Dunning, *J. Chem. Phys.*, 1989, **90**, 1007–1023.

30 R. A. Kendall, T. H. Dunning and R. J. Harrison, *J. Chem. Phys.*, 1992, **96**, 6796–6806.

31 F. Weigend, A. Köhn and C. Hättig, *J. Chem. Phys.*, 2002, **116**, 3175–3183.

32 J. G. Hill and J. A. Platts, *J. Chem. Phys.*, 2008, **128**, 044104.

33 D. H. Bross, J. G. Hill, H.-J. Werner and K. A. Peterson, *J. Chem. Phys.*, 2013, **139**, 094302.

34 F. Weigend, *Phys. Chem. Chem. Phys.*, 2002, **4**, 4285–4291.

35 A. D. Becke, *Phys. Rev. A*, 1988, **38**, 3098–3100.

36 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648–5652.

37 P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, *J. Phys. Chem.*, 1994, **98**, 11623–11627.

38 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865–3868.

39 C. Adamo and V. Barone, *J. Chem. Phys.*, 1999, **110**, 6158–6170.

40 M. Ernzerhof and G. E. Scuseria, *J. Chem. Phys.*, 1999, **110**, 5029–5036.

41 Y. Zhao, N. E. Schultz and D. G. Truhlar, *J. Chem. Theory Comput.*, 2006, **2**, 364–382.

42 Y. Zhao and D. G. Truhlar, *Theor. Chem. Acc.*, 2008, **120**, 215–241.

43 D. G. A. Smith, L. A. Burns, K. Patkowski and C. D. Sherrill, *J. Phys. Chem. Lett.*, 2016, **7**, 2197–2203.

44 P. S. Żuchowski, R. Podeszwa, R. Moszyński, B. Jeziorski and K. Szalewicz, *J. Chem. Phys.*, 2008, **129**, 084101.

45 M. Hapka, P. S. Żuchowski, M. M. Szczęśniak and

G. Chałasiński, *J. Chem. Phys.*, 2012, **137**, 164104.
46 J. F. Gonthier and C. D. Sherrill, *J. Chem. Phys.*, 2016, **145**, 134106.
47 H.-J. Werner and P. J. Knowles, *J. Chem. Phys.*, 1985, **82**, 5053–5063.
48 P. J. Knowles and H.-J. Werner, *Chem. Phys. Lett.*, 1985, **115**, 259–267.
49 S. Boys and F. Bernardi, *Mol. Phys.*, 1970, **19**, 553–566.
50 H.-J. Werner, *Mol. Phys.*, 1996, **89**, 645–661.
51 P. Celani and H.-J. Werner, *J. Chem. Phys.*, 2000, **112**, 5546–5557.
52 C. Angeli, R. Cimiraglia, S. Evangelisti, T. Leininger and J.-P. Malrieu, *J. Chem. Phys.*, 2001, **114**, 10252–10264.
53 C. Angeli, R. Cimiraglia and J.-P. Malrieu, *J. Chem. Phys.*, 2002, **117**, 9138–9153.
54 C. Angeli, M. Pastore and R. Cimiraglia, *Theor. Chem. Acc.*, 2006, **117**, 743–754.
55 H.-J. Werner and P. J. Knowles, *J. Chem. Phys.*, 1988, **89**, 5803–5814.
56 P. J. Knowles and H.-J. Werner, *Chem. Phys. Lett.*, 1988, **145**, 514 – 522.
57 S. R. Langhoff and E. R. Davidson, *Int. J. Quantum Chem.*, 1974, **8**, 61–72.
58 L. Meissner, *Chem. Phys. Lett.*, 1988, **146**, 204 – 210.
59 T. M. Parker, L. A. Burns, R. M. Parrish, A. G. Ryno and C. D. Sherrill, *J. Chem. Phys.*, 2014, **140**, 094106.

Appendix D

Spin splittings from first-order symmetry-adapted perturbation theory without single-exchange approximation

The Journal of Chemical Physics ARTICLE scitation.org/journal/jcp

Spin splittings from first-order symmetry-adapted perturbation theory without single-exchange approximation

Cite as: J. Chem. Phys. 150, 074109 (2019); doi: 10.1063/1.5086079
Submitted: 18 December 2018 • Accepted: 29 January 2019 •
Published Online: 20 February 2019

Jonathan M. Waldrop and Konrad Patkowski

AFFILIATIONS

Department of Chemistry and Biochemistry, Auburn University, Auburn, Alabama 36849, USA

ABSTRACT

The recently proposed spin-flip symmetry-adapted perturbation theory (SF-SAPT) first-order exchange energy [Patkowski *et al.*, J. Chem. Phys. **148**, 164110 (2018)] enables the standard open-shell SAPT approach to treat arbitrary spin states of the weakly interacting complex. Here, we further extend first-order SF-SAPT beyond the single-exchange approximation to a complete treatment of the exchanges of electrons between monomers. This new form of the exchange correction replaces the single-exchange approximation with a more moderate single-spin-flip approximation. The newly developed expressions are applied to a number of small test systems to elucidate the quality of both approximations. They are also applied to the singlet-triplet splittings in pancake bonded dimers. The accuracy of the single-exchange approximation deteriorates at short intermolecular separations, especially for systems with few electrons and for the high-spin state of the complex. In contrast, the single-spin-flip approximation is exact for interactions involving a doublet molecule and remains highly accurate for any number of unpaired electrons. Because the single-exchange approximation affects the high-spin and low-spin states of pancake bonded complexes evenly, the resulting splitting values are of similar accuracy to those produced by the formally more accurate single-spin-flip approximation.

I. INTRODUCTION

The interaction of two open-shell molecules in their high-spin states produces a whole bundle of asymptotically degenerate states corresponding to different values of the total spin quantum number S for the complex. The splittings between these states arise from the resonance tunneling of electrons between the two subsystems. Thus, when the intermolecular interaction is described in terms of symmetry-adapted perturbation theory (SAPT),[1] the splittings can be attributed exclusively to the exchange terms in the interaction: the remaining electrostatic, induction, and dispersion terms are the same for the entire asymptotically degenerate bundle. On the other hand, a uniform description of all spin states is challenging for the more conventional supermolecular approach to interaction energies: while the highest-spin state might often be well described by a single-reference treatment, all the remaining, low-spin states are genuinely multireference. Therefore, in computational studies of open-shell complexes, obtaining sufficiently accurate potential energy surfaces (PESs) for low-spin states is nontrivial. For example, for the interaction of two ground-state ($^3\Sigma_g^-$) O_2 molecules, an accurate PES for the high-spin quintet state could be constructed using restricted coupled-cluster theory with singles, doubles, and perturbative triple excitations (RCCSD(T)).[2] On the other hand, the PESs for the multireference singlet and triplet states of this complex had to be obtained by combining the RCCSD(T) quintet PES with lower-level (complete active space second order perturbation theory, CASPT2, or multireference configuration interaction with single and double excitations, MRCI) estimates of the singlet-quintet and triplet-quintet splittings.[3]

As long as the noninteracting monomers are amenable to a single-reference description (in this work, via spin-restricted high-spin determinants Ψ_A and Ψ_B), the evaluation of SAPT corrections does not require constructing a

multireference wavefunction for the complex. Instead, the perturbation series is built on top of a zeroth-order function $\Psi_0 = \Psi_A\Psi_B$ and the corrections are expressed in terms of single, double, ... excitations out of the single Ψ_0 reference. However, the established open-shell SAPT formulations, based on either spin-restricted[4] on unrestricted[5,6] Hartree-Fock (HF) or Kohn-Sham (KS) determinants, suffer from a different limitation: it is assumed that Ψ_0 is a pure spin state (an eigenfunction of the $\hat{S}^2$ operator). This is true only if the asymptotically degenerate bundle reduces to a single spin state, that is, when either one of the interacting molecules is a singlet or the complex is in the high-spin state, $M_S = \pm S$. Thus, the open-shell SAPT approaches of Refs. 4–6 are only applicable to the high-spin state of the complex. When the interaction energy in a low-spin state is needed, these high-spin approaches can provide the electrostatic, induction, and dispersion corrections, but not the exchange corrections. No estimate of spin splittings can be extracted from these theories.

As the first step towards extending SAPT beyond the high-spin state of the complex, we have recently developed[7] a new first-order exchange correction valid for an arbitrary spin state of two interacting high-spin open-shell molecules, each described by a restricted open-shell HF (ROHF) determinant. The new formalism involves explicitly projecting Ψ_0 onto the subspace corresponding to the desired value of the spin quantum number S. Similarly to nearly all exchange corrections in closed-shell (and high-spin open-shell) SAPT, the evaluation of matrix elements involving the $(N_A + N_B)$-electron antisymmetrizer is greatly simplified by the use of the *single-exchange approximation*, also called the S^2 approximation as it retains terms up to second order in intermolecular overlap integrals.[1] As shown in Ref. 7, the S^2 approximation allows expressing the first-order exchange energy $E^{(10)}_{exch}$ for an arbitrary spin state as a linear combination of two matrix elements: a *diagonal exchange energy* that quantifies the spin-averaged effect and a *spin-flip term* responsible for splittings between multiplets. The coefficients in this linear combination do not depend on a particular system and arise solely from the angular momentum algebra. The name "spin-flip term" reflects the fact that the matrix elements of this term are computed between $\Psi_A\Psi_B$ and a function $\Psi_A^{\downarrow}\Psi_B^{\uparrow}$ where one of the unpaired spins on one monomer has been lowered and one of the unpaired spins on the other monomer has been raised (that is, an intermolecular spin exchange has occurred). Thus, the new formalism, termed spin-flip SAPT (SF-SAPT), bears some similarities to the spin-flip electronic structure theories of Krylov and co-workers:[8,9] in both approaches, a multireference low-spin state is accessed from a single reference configuration. However, the excitations in the methods of Krylov *et al.* alter the total spin of the system (the quantum number M_S is changed). On the other hand, the excitations involved in spin-flip SAPT do not change the value of M_S as a spin raise on one monomer is always accompanied by a spin lowering on the other monomer.

The spin-flip electronic structure formalism has been generalized to multiple spin flips;[10] however, a simple and often adequate approximation[11] relies on using the single-spin-flip approach (accessing the second-highest spin state from the high-spin one) to determine the coupling parameter J_{AB} within the Heisenberg spin Hamiltonian model. Within this scheme, the knowledge of J_{AB} is sufficient to recover the entire bundle of lower-spin states. As demonstrated in Ref. 7, in the case of spin-flip SAPT, the same single-parameter Heisenberg picture is a direct consequence of the single-exchange approximation where the value of J_{AB} results from the single-spin-flip matrix elements (the corresponding elements involving two or more spin flips vanish within the single-exchange approximation).

While it would not be fair to expect quantitative accuracy from such a (conceptually and computationally) simple approximation as first-order perturbation theory, the SF-SAPT approach to $E^{(10)}_{exch}$ has been shown[7] to provide reasonable, qualitatively correct multiplet splittings for a number of representative complexes. The accuracy of first-order SF-SAPT is generally similar to the (much more involved) supermolecular complete active space self-consistent field (CASSCF) calculation and the method does not break down at large separations as is the case for size-inconsistent approaches such as MRCI. However, the accuracy of SF-SAPT splittings in the region of strong intermolecular overlap is limited by two issues: the lack of second- and higher-order exchange effects and the single-exchange approximation. Addressing the first issue involves deriving and implementing the SF-SAPT generalizations of the second-order exchange corrections $E^{(20)}_{exch-ind}$ and $E^{(20)}_{exch-disp}$, which is in progress in our group and will be described in a separate publication. In this work, we focus on addressing the second issue by deriving and implementing an improved expression for $E^{(10)}_{exch}$ in SF-SAPT, in which the single-exchange approximation has been replaced by a much milder *single-spin-flip (1-flip) approximation*.

The full nonapproximated $E^{(10)}_{exch}$ expression in standard closed-shell SAPT has been introduced a long time ago[12] and this expression, unlike those for higher SAPT exchange corrections, is available in most SAPT implementations including the high-spin open-shell ones.[4–6] Much more recently, Schäffer and Jansen derived and implemented nonapproximated expressions for the second-order exchange corrections $E^{(20)}_{exch-ind}$[13] and $E^{(20)}_{exch-disp}$.[14] We will directly adopt Schäffer and Jansen's approach, based on the properties of singly and doubly excited determinants and their cofactors, in our development of the SF-SAPT $E^{(10)}_{exch}$ correction (in fact, as we will see below, the spin-flip excitation is just another kind of double excitation such as the one in the formulas for $E^{(20)}_{exch-disp}$). It should be mentioned that the nonapproximated second-order closed-shell SAPT corrections of Refs. 13 and 14 have recently been implemented within the freely available PSI4NUMPY framework.[15]

It should be stressed that while the single-exchange approximation *implies* the single-spin-flip approximation (the neglect of matrix elements involving double and higher spin

flips), the two approximations are not nearly equivalent. For example, in any interaction involving an ROHF monomer in a doublet state, there is only one unpaired spin on this monomer that can be flipped, so the 1-flip approximation is exact. On the other hand, the single-exchange approximation is not exact for any spin states of the monomers, even closed-shell singlets. It should be noted that the interaction energy in the high-spin state of the complex ($S = S_A + S_B$) can be obtained in two fully equivalent ways: using the high-spin formalism of Refs. 4–6 (which leads to the S, $M_S = S$ configuration) and using the SF-SAPT formalism (which leads to the S, $M_S = S_A - S_B$ configuration degenerate with the previous one). As the first-order exchange energy in the high-spin formalism does not need to involve the S^2 approximation, we will be able to verify the importance of both approximations for the high-spin state of the complex (but not for the low-spin states) by comparison with the high-spin method implemented, e.g., in the PSI4 package.[6,16] In fact, in Ref. 7, the largest errors due to the S^2 approximation in the high-spin complex occurred for the doublet-doublet Li–Li interaction, and this approximation might be far from exact for another doublet-doublet complex tested there, the pancake-bonded phenalenyl (PLY) dimer. Thus, the replacement of the single-exchange approximation by the 1-flip one (which is exact for two interacting doublets) might be expected to significantly improve the short-range splittings in some of the systems studied in Ref. 7 and possibly in pancake bonded complexes in general.[17]

The structure of the rest of this article is as follows: In Sec. II, we develop the formalism and derive formulas for the arbitrary-spin first-order exchange correction in terms of molecular-orbital (MO) integrals. In Sec. III, we recast the MO formulas into the atomic-orbital (AO) basis to facilitate an efficient implementation that does not require integral transformation and can utilize the benefits of density fitting (DF). Sec. IV contains the results of our new methodology for the systems studied in Ref. 7 as well as several larger pancake-bonded complexes. Finally, Sec. V presents conclusions.

II. MOLECULAR ORBITAL FORMALISM

Throughout this paper, the indices i and j denote all occupied spinorbitals of monomers A and B, respectively. Furthermore, the index i is split into inactive (k, corresponding to a doubly occupied orbital) and active (m) spinorbitals of A, and the index j is split into inactive (l) and active (n) spinorbitals of B. The indices r and s are used for arbitrary spinorbitals occupied in the zeroth-order wavefunction regardless of the monomer. We will add an arrow ↑ or ↓ to the spinorbital index whenever an explicit specification of the spin is necessary. Ψ_A and Ψ_B are the ground-state wavefunctions for the individual monomers and are assumed to be ROHF wavefunctions where the unpaired electrons in Ψ_A have α spin and the unpaired electrons of Ψ_B have β spin. The product of these two wavefunctions is considered the zeroth-order dimer wavefunction. As with the previous derivation of the SF-SAPT first-order exchange correction within the single-exchange approximation, the current correction is computed within the symmetrized Rayleigh-Schrödinger (SRS) formalism.[1,18] In the case of low-spin states, the ($N_A + N_B$)-electron antisymmetrizer $\mathcal{A}$ is accompanied by the spin projector P_{SM_S}, which projects the dimer wavefunction onto the subspace corresponding to the spin quantum numbers S and M_S. As such, the SAPT first-order interaction energy for a desired total spin is obtained from the expression

$$E_{\text{int}}^{(10)} = \frac{\langle \Psi_A \Psi_B | V \mathcal{A} P_{SM_S} | \Psi_A \Psi_B \rangle}{\langle \Psi_A \Psi_B | \mathcal{A} P_{SM_S} | \Psi_A \Psi_B \rangle}, \tag{1}$$

where V is the perturbation operator that collects the interactions between the monomers. The spin projector acting on the dimer wavefunction is approximated by its expansion truncated after a single spin flip

$$P_{SM_S} \Psi_A \Psi_B = c_0 \Psi_A \Psi_B + c_1 \Psi_A^{\downarrow} \Psi_B^{\uparrow}, \tag{2}$$

where c_0 and c_1 are the Clebsch-Gordan coefficients $\langle S(S_A - S_B)|S_A S_A S_B\,(-S_B)\rangle$ and $\langle S(S_A - S_B)|S_A(S_A - 1)S_B(-S_B + 1)\rangle$, respectively. The arrow superscripts denote a spin-flipped monomer wavefunction, defined as $\Psi_X^{\downarrow} = (1/\sqrt{2S_X})\hat{S}_- \Psi_X$ or $\Psi_X^{\uparrow} = (1/\sqrt{2S_X})\hat{S}_+ \Psi_X$. $\hat{S}_\pm$ are spin-raising and spin-lowering operators, which act on a wavefunction by applying the one-electron spin-lowering or raising operators to all electrons in the wavefunction. The result of this operation is the sum of the wavefunctions where one of the active electrons (and only the active electrons) has had its spin flipped. The terms in parentheses in the above definitions normalize the functions $\Psi_X^{\downarrow}$ and $\Psi_X^{\uparrow}$ as required in Eq. (2). Due to the assumed spins of the active electrons in Ψ_A and Ψ_B, the spin lowering only makes sense for Ψ_A and spin raising only makes sense for Ψ_B. While the application of the full spin projector produces terms with multiple spin flips,[7] the truncation in Eq. (2) to terms with no more than singly spin-flipped monomers is the essence of the single-spin-flip approximation.

The combination of Eqs. (1) and (2) produces the following modified interaction energy equation:

$$E_{\text{int}}^{(10)} = \frac{\langle \Psi_A \Psi_B | V \mathcal{A} | \Psi_A \Psi_B \rangle + \frac{c_1}{c_0} \langle \Psi_A \Psi_B | V \mathcal{A} | \Psi_A^{\downarrow} \Psi_B^{\uparrow} \rangle}{\langle \Psi_A \Psi_B | \mathcal{A} | \Psi_A \Psi_B \rangle + \frac{c_1}{c_0} \langle \Psi_A \Psi_B | \mathcal{A} | \Psi_A^{\downarrow} \Psi_B^{\uparrow} \rangle}. \tag{3}$$

Thus, in order to compute the first-order exchange energy $E_{\text{exch}}^{(10)}$ within the single-spin-flip approximation, one needs to evaluate the four matrix elements present in Eq. (3) and subtract the electrostatic contribution:

$$E_{\text{exch}}^{(10)} = E_{\text{int}}^{(10)} - E_{\text{elst}}^{(10)} = E_{\text{int}}^{(10)} - \langle \Psi_A \Psi_B | V | \Psi_A \Psi_B \rangle. \tag{4}$$

The leading terms in the numerator and denominator of Eq. (3) are the previously derived components of the complete

SAPT first-order interaction energy[12] and have the following forms:

$$\langle \Psi_A \Psi_B | \mathcal{A} | \Psi_A \Psi_B \rangle = \frac{N_A! N_B!}{N!} \mathcal{S}, \tag{5}$$

$$\langle \Psi_A \Psi_B | V\mathcal{A} | \Psi_A \Psi_B \rangle = \frac{N_A! N_B!}{N!} \Big(W_{AB} \mathcal{S} + \sum_{ir} B_{ir} \mathcal{S}^{ir} + \sum_{jr} A_{jr} \mathcal{S}^{jr} + \frac{1}{2} \sum_{ijrs} \langle ij||rs \rangle \mathcal{S}^{ij,rs} \Big), \tag{6}$$

where $\mathcal{S}$ is the determinant of the overlap matrix of occupied spinorbitals of both monomers, $\mathbf{S}$

$$\mathbf{S} = \begin{bmatrix} \mathbf{1} & \mathbf{S}_{AB} \\ (\mathbf{S}_{AB})^T & \mathbf{1} \end{bmatrix}. \tag{7}$$

In Eq. (6), $\mathcal{S}^{ir}$ is a first cofactor of the determinant $\mathcal{S}$, $\mathcal{S}^{ij,rs}$ is a second cofactor, W_{AB} is the nuclear repulsion between monomers A and B, B_{ir} and A_{jr} are elements of the nuclear attraction matrices for the corresponding monomer, and $\langle ij||rs \rangle$ are antisymmetrized two-electron integrals in the physicists' notation. The first cofactor $\mathcal{S}^{ir}$ of a determinant is obtained by the deletion of row i and column r from the original determinant and multiplying the resulting determinant by $(-1)^{i+r}$. The second cofactor $\mathcal{S}^{ij,rs}$ is obtained by the deletion of two rows i, j and two columns r, s from the original determinant and multiplying the resulting determinant by $(-1)^{i+j+r+s}$. Furthermore, it is antisymmetric with respect to the order of deletions: $\mathcal{S}^{ji,rs} = -\mathcal{S}^{ij,rs}$ and $\mathcal{S}^{ij,sr} = -\mathcal{S}^{ij,rs}$.

The Cramer's rule for the relationship between a determinant, its inverse, and its cofactors implies that

$$D_{ri} = \frac{1}{\mathcal{S}} \mathcal{S}^{ir}, \tag{8}$$

where D_{ri} are the elements of the inverse of the overlap matrix, $\mathbf{D} = \mathbf{S}^{-1}$. It is instructive to examine the structure of matrices $\mathbf{S}$ and $\mathbf{D}$ in more detail. Obviously, $\mathbf{S}$, and thus also $\mathbf{D}$, is block-diagonal with respect to spin. Now, the spin-up and spin-down blocks of $\mathbf{S}$ are not the same: the spin-up block contains overlap between orbital types k, m, and l, and the spin-down block contains overlap between orbital types k, l, and n. Thus, the spin-up and spin-down blocks of the inverse matrix $\mathbf{D}$ are completely distinct even for the common indices such as k: $D^{\alpha}_{kk'} \neq D^{\beta}_{kk'}$. Therefore, in the final orbital formulas for the SF-SAPT expressions, we will explicitly specify the spin block of matrix $\mathbf{D}$ as D^{α}_{rs} or D^{β}_{rs}. Finally, note that the matrix $\mathbf{D}$, as an inverse of a symmetric matrix $\mathbf{S}$, is also symmetric.

According to Eq. (8) and the relationship between the first and second cofactors,[19] the second cofactors are equivalent to

$$\mathcal{S}^{ij,rs} = \mathcal{S}(D_{ri} D_{sj} - D_{si} D_{rj}). \tag{9}$$

Note that the second cofactors are indeed antisymmetric with regard to swapping either i and j or r and s. Using Eqs. (8) and (9), Eq. (6) can be rewritten in the form making explicit use of the D_{ri} matrix elements

$$\langle \Psi_A \Psi_B | V\mathcal{A} | \Psi_A \Psi_B \rangle = \frac{N_A! N_B!}{N!} \mathcal{S} \Big[W_{AB} + \sum_{ir} B_{ir} D_{ri} + \sum_{jr} A_{jr} D_{rj} + \sum_{ijrs} (\langle ij|rs \rangle - \langle ij|sr \rangle) D_{ri} D_{sj} \Big], \tag{10}$$

where we have used the fact that r and s span the same spinorbital space and are therefore interchangeable.

The other terms in the numerator and denominator of Eq. (3) contain the spin-flipped wavefunctions where one active electron from each monomer has had its spin flipped. These terms are a specific subset of the double excitations needed for the computation of the nonapproximated second-order SAPT exchange dispersion correction which was recently derived by Schäffer and Jansen.[14] Adopting the approach of Ref. 14, the second terms in the numerator and denominator of Eq. (3) are rewritten as

$$\langle \Psi_A \Psi_B | \mathcal{A} | \Psi_A^{\downarrow} \Psi_B^{\uparrow} \rangle = \frac{1}{2\sqrt{S_A S_B}} \sum_{mn} \langle \Psi_A \Psi_B | \mathcal{A} | \Psi_{A,m^{\uparrow} \to m^{\downarrow}} \Psi_{B,n^{\downarrow} \to n^{\uparrow}} \rangle = \frac{N_A! N_B!}{N!} \frac{1}{2\sqrt{S_A S_B}} \sum_{mn} \mathcal{S}_{m^{\uparrow} \to m^{\downarrow}, n^{\downarrow} \to n^{\uparrow}} \tag{11}$$

and

$$\begin{aligned} \langle \Psi_A \Psi_B | V\mathcal{A} | \Psi_A^{\downarrow} \Psi_B^{\uparrow} \rangle &= \frac{1}{2\sqrt{S_A S_B}} \sum_{mn} \langle \Psi_A \Psi_B | V\mathcal{A} | \Psi_{A,m^{\uparrow} \to m^{\downarrow}} \Psi_{B,n^{\downarrow} \to n^{\uparrow}} \rangle \\ &= \frac{N_A! N_B!}{N!} \frac{1}{2\sqrt{S_A S_B}} \sum_{mn} \Big(W_{AB} \mathcal{S}_{m^{\uparrow} \to m^{\downarrow}, n^{\downarrow} \to n^{\uparrow}} + \sum_{ir} B_{i\tilde{r}} \mathcal{S}^{ir}_{m^{\uparrow} \to m^{\downarrow}, n^{\downarrow} \to n^{\uparrow}} + \sum_{jr} A_{j\tilde{r}} \mathcal{S}^{jr}_{m^{\uparrow} \to m^{\downarrow}, n^{\downarrow} \to n^{\uparrow}} \\ &\quad + \frac{1}{2} \sum_{ijrs} \langle ij||\tilde{r}\tilde{s} \rangle \mathcal{S}^{ij,rs}_{m^{\uparrow} \to m^{\downarrow}, n^{\downarrow} \to n^{\uparrow}} \Big) \\ &= \frac{N_A! N_B!}{N!} \frac{1}{2\sqrt{S_A S_B}} \sum_{mn} (I_1 + I_2 + I_3 + I_4), \end{aligned} \tag{12}$$

where $\mathcal{S}_{m^{\uparrow} \to m^{\downarrow}, n^{\downarrow} \to n^{\uparrow}}$ is the determinant of the overlap matrix that results from flipping the spin of m and n in the ket, and $\mathcal{S}^{ir}_{m^{\uparrow} \to m^{\downarrow}, n^{\downarrow} \to n^{\uparrow}}$ and $\mathcal{S}^{ij,rs}_{m^{\uparrow} \to m^{\downarrow}, n^{\downarrow} \to n^{\uparrow}}$ are the first and second cofactors of that determinant. The arrows on the indices denote spin up or spin down. The tilded indices $\tilde{r}, \tilde{s}$ in Eq. (12) denote the *contents* of the columns r, s in the spin-flipped determinant $\mathcal{S}_{m^{\uparrow} \to m^{\downarrow}, n^{\downarrow} \to n^{\uparrow}}$: thus, when $r = m^{\uparrow}$, $\tilde{r} = m^{\downarrow}$, and when $r = n^{\downarrow}$, $\tilde{r} = n^{\uparrow}$ (otherwise, $\tilde{r} = r$). Note that this meaning of a tilde over an index is completely different from the notation of Refs. 13 and 14. The four consecutive terms $I_1, \ldots, I_4$ in this equation will be analyzed separately–see below. It is obvious at this point that $\frac{N_A! N_B!}{N!}$ appears in all terms in the top and bottom of Eq. (3) and will be canceled out in the total interaction energy.

It is beneficial at this time to define the relationships between non-excited, singly excited, and doubly excited

determinants and their cofactors. The relationship between a determinant $\mathcal{S}$ and a singly excited determinant $\mathcal{S}_{i\to a}$ is such that

$$\mathcal{S}_{i\to a} = \sum_r S_{ra}\mathcal{S}^{ri} = \mathcal{S}\sum_r D_{ir}S_{ra}, \tag{13}$$

where S_{ra} are elements of the overlap matrix. The first and second cofactors of a singly excited determinant can be expressed in terms of singly excited determinants and the first and second cofactors of the non-excited determinant in the following ways:[13]

$$\mathcal{S}^{rs}_{i\to a} = \begin{cases} \mathcal{S}^{ri} & s = i \\ \frac{1}{\mathcal{S}}\left(\mathcal{S}^{rs}\mathcal{S}_{i\to a} - \mathcal{S}^{ri}\mathcal{S}_{s\to a}\right) & s \neq i, \end{cases} \tag{14}$$

$$\mathcal{S}^{rs,tu}_{i\to a} = \begin{cases} \mathcal{S}^{rs,iu} & t = i \\ \mathcal{S}^{rs,ti} & u = i \\ \frac{1}{\mathcal{S}}\left(\mathcal{S}^{rs,tu}\mathcal{S}_{i\to a} - \mathcal{S}^{rs,iu}\mathcal{S}_{t\to a} - \mathcal{S}^{rs,ti}\mathcal{S}_{u\to a}\right) & i \notin \{t,u\}. \end{cases} \tag{15}$$

The doubly excited determinants can be written in terms of singly excited determinants as

$$\mathcal{S}_{i\to a,j\to b} = \frac{1}{\mathcal{S}}(\mathcal{S}_{i\to a}\mathcal{S}_{j\to b} - \mathcal{S}_{i\to b}\mathcal{S}_{j\to a}). \tag{16}$$

The first and second cofactors of doubly excited determinants can be expressed in terms of the other components as[14]

$$\mathcal{S}^{rs}_{i\to a,j\to b} = \begin{cases} \mathcal{S}^{ri}_{j\to b} & s = i \\ \mathcal{S}^{rj}_{i\to a} & s = j \\ \frac{1}{\mathcal{S}}\left(\mathcal{S}_{i\to a,j\to b}\mathcal{S}^{rs} - \mathcal{S}_{s\to a,j\to b}\mathcal{S}^{ri} - \mathcal{S}_{i\to a,s\to b}\mathcal{S}^{rj}\right) & s \notin \{i,j\}, \end{cases} \tag{17}$$

$$\mathcal{S}^{rs,tu}_{i\to a,j\to b} = \begin{cases} \mathcal{S}^{rs,ij} & t = i, u = j \\ \mathcal{S}^{rs,ji} & t = j, u = i \\ \mathcal{S}^{rs,tj}_{i\to a} & t \neq i, u = j \\ \mathcal{S}^{rs,ti}_{j\to b} & t \neq j, u = i \\ \mathcal{S}^{rs,iu}_{j\to b} & t = i, u \neq j \\ \mathcal{S}^{rs,ju}_{i\to a} & t = j, u \neq i \\ \frac{1}{\mathcal{S}}\left(\mathcal{S}_{i\to a,j\to b}\mathcal{S}^{rs,tu} - \mathcal{S}_{t\to a,j\to b}\mathcal{S}^{rs,iu} - \mathcal{S}_{u\to a,j\to b}\mathcal{S}^{rs,ti}\right. & \\ \left. -\mathcal{S}_{i\to a,t\to b}\mathcal{S}^{rs,ju} - \mathcal{S}_{i\to a,u\to b}\mathcal{S}^{rs,tj} + \mathcal{S}_{t\to a,u\to b}\mathcal{S}^{rs,ij}\right) & i,j \notin \{t,u\}. \end{cases} \tag{18}$$

With the relationships between various determinants and cofactors now defined, it is possible to express the remaining parts of Eqs. (11) and (12) in terms of the elements of the **D** and **S** matrices. Starting with the term from the denominator,

$$\begin{aligned} &\langle \Psi_A\Psi_B|\mathcal{A}|\Psi_A^{\downarrow}\Psi_B^{\uparrow}\rangle \\ &\quad = \frac{N_A!N_B!}{N!}\frac{1}{2\sqrt{S_AS_B}}\sum_{mn}\mathcal{S}_{m^{\uparrow}\to m^{\downarrow},n^{\downarrow}\to n^{\uparrow}} \\ &\quad = \frac{N_A!N_B!}{N!}\frac{1}{2\sqrt{S_AS_B}}\sum_{mn}\left(\frac{1}{\mathcal{S}}(\mathcal{S}_{m^{\uparrow}\to m^{\downarrow}}\mathcal{S}_{n^{\downarrow}\to n^{\uparrow}} - \mathcal{S}_{n^{\downarrow}\to m^{\downarrow}}\mathcal{S}_{m^{\uparrow}\to n^{\uparrow}})\right) \\ &\quad = -\frac{N_A!N_B!}{N!}\frac{1}{2\sqrt{S_AS_B}}\mathcal{S}\sum_{mn}\left(\sum_j D_{n^{\downarrow}j}S_{jm^{\downarrow}}\right)\left(\sum_i D_{m^{\uparrow}i}S_{in^{\uparrow}}\right), \end{aligned} \tag{19}$$

where the cancellation of the $\mathcal{S}_{m^{\uparrow}\to m^{\downarrow}}\mathcal{S}_{n^{\downarrow}\to n^{\uparrow}}$ term is due to the spin-diagonal nature of **S** and **D** [cf. Eq. (13)]. A singly excited determinant with a spin-flipping excitation requires coupling **S** and **D** matrix elements of opposite spin, resulting in a zero spin integral. This property will be used repeatedly to simplify the following equations. The new summations in the last line have been truncated from r since $m^{\downarrow}$ is orthogonal to all occupied spinorbitals on A and $n^{\uparrow}$ is orthogonal to all occupied spinorbitals on B (note that, e.g., the index j covers both the inactive spinorbitals $l^{\uparrow}, l^{\downarrow}$ and active spinorbitals $n^{\downarrow}$). The complete denominator is thus equal to

$$\frac{N_A!N_B!}{N!}\mathcal{S}\left(1 - \frac{c_1}{2c_0\sqrt{S_AS_B}}\sum_{mnij}(D_{n^{\downarrow}j}S_{jm^{\downarrow}}D_{m^{\uparrow}i}S_{in^{\uparrow}})\right). \tag{20}$$

For the numerator term, we will analyze the consecutive contributions I_n to Eq. (12). The first of these terms, I_1, is simply the nuclear repulsion term multiplied by the same doubly excited determinant that appears in the denominator.

It is easy to see that the terms in the numerator that contain the nuclear repulsion are, in fact, equal to the denominator multiplied by the nuclear repulsion. As such, the total interaction energy contains this term exactly once, which is to be expected.

I_2, the term containing the nuclear potential of monomer B, contains a first cofactor of a doubly excited determinant. As the summation over r contains both $m^\uparrow$ and $n^\downarrow$, it has to be broken into three parts accounting for the special cases shown in Eq. (17). The resulting equation is

$$I_2 = \sum_i B_{im^\downarrow} \mathcal{S}^{im^\uparrow}_{n^\downarrow\to n^\uparrow} + \sum_i B_{in^\uparrow} \mathcal{S}^{in^\downarrow}_{m^\uparrow\to m^\downarrow} + \sum_{ir,r\neq(m^\uparrow,n^\downarrow)} B_{i\tilde{r}} \frac{1}{\mathcal{S}} \times \left(\mathcal{S}_{m^\uparrow\to m^\downarrow,n^\downarrow\to n^\uparrow}\mathcal{S}^{ir} - \mathcal{S}_{r\to m^\downarrow,n^\downarrow\to n^\uparrow}\mathcal{S}^{im^\uparrow} - \mathcal{S}_{m^\uparrow\to m^\downarrow,r\to n^\uparrow}\mathcal{S}^{in^\downarrow}\right). \tag{21}$$

We expand the first summation in this term as

$$\sum_i B_{im^\downarrow}\mathcal{S}^{im^\uparrow}_{n^\downarrow\to n^\uparrow} = \sum_i B_{im^\downarrow}\left(\frac{1}{\mathcal{S}}\left(\mathcal{S}^{im^\uparrow}\mathcal{S}_{n^\downarrow\to n^\uparrow} - \mathcal{S}^{in^\downarrow}\mathcal{S}_{m^\uparrow\to n^\uparrow}\right)\right) = -\sum_{ii'} \mathcal{S} B_{im^\downarrow} D_{n^\downarrow i} D_{m^\uparrow i'} S_{i'n^\uparrow}, \tag{22}$$

and the analogous result for the second summation is

$$\sum_i B_{in^\uparrow}\mathcal{S}^{in^\downarrow}_{m^\uparrow\to m^\downarrow} = -\sum_{ij} \mathcal{S} B_{in^\uparrow} D_{m^\uparrow i} D_{n^\downarrow j} S_{jm^\downarrow}. \tag{23}$$

The third summation in I_2 contains the same doubly excited determinant that was previously described, as well as two others. Following the same logic as above to eliminate the singly excited determinants that vanish by spin integration and expand the remaining determinants, and noting that $\tilde{r} = r$ for all terms in the restricted summation, this last term is equal to

$$\sum_{ir,r\neq(m^\uparrow,n^\downarrow)} \mathcal{S} B_{ir}\left(-D_{ri}\sum_{i'j} D_{m^\uparrow i'} S_{i'n^\uparrow} D_{n^\downarrow j} S_{jm^\downarrow} + D_{m^\uparrow i}\sum_{i'j} D_{ri'} S_{i'n^\uparrow} D_{n^\downarrow j} S_{jm^\downarrow} + D_{n^\downarrow i}\sum_{i'j} D_{m^\uparrow i'} S_{i'n^\uparrow} D_{rj} S_{jm^\downarrow}\right). \tag{24}$$

The restriction in the summation in Eq. (24) to r not equal to either $m^\uparrow$ or $n^\downarrow$ can be lifted since the result of the additional $r = m^\uparrow$ and $r = n^\downarrow$ terms is equal to zero in these cases. For example, when $r = m^\uparrow$, the first two terms in parentheses cancel each other and the third term is zero due to spin. A similar result can be obtained for $r = n^\downarrow$.

With these results in hand, the complete I_2 term is

$$I_2 = -\mathcal{S}\Big(\sum_{ii'} B_{im^\downarrow} D_{n^\downarrow i} D_{m^\uparrow i'} S_{i'n^\uparrow} + \sum_{ij} B_{in^\uparrow} D_{m^\uparrow i} D_{n^\downarrow j} S_{jm^\downarrow} + \sum_{ir} B_{ir}\Big(D_{ri}\sum_{i'j} D_{m^\uparrow i'} S_{i'n^\uparrow} D_{n^\downarrow j} S_{jm^\downarrow} - D_{m^\uparrow i}\sum_{i'j} D_{ri'} S_{i'n^\uparrow} D_{n^\downarrow j} S_{jm^\downarrow} - D_{n^\downarrow i}\sum_{i'j} D_{m^\uparrow i'} S_{i'n^\uparrow} D_{rj} S_{jm^\downarrow}\Big)\Big). \tag{25}$$

An analogous derivation for I_3 gives

$$I_3 = -\mathcal{S}\Big(\sum_{ij} A_{jm^\downarrow} D_{n^\downarrow j} D_{m^\uparrow i} S_{in^\uparrow} + \sum_{jj'} A_{jn^\uparrow} D_{m^\uparrow j} D_{n^\downarrow j'} S_{j'm^\downarrow} + \sum_{jr} A_{jr}\Big(D_{rj}\sum_{ij'} D_{m^\uparrow i} S_{in^\uparrow} D_{n^\downarrow j'} S_{j'm^\downarrow} - D_{m^\uparrow j}\sum_{ij'} D_{ri} S_{in^\uparrow} D_{n^\downarrow j'} S_{j'm^\downarrow} - D_{n^\downarrow j}\sum_{ij'} D_{m^\uparrow i} S_{in^\uparrow} D_{rj'} S_{j'm^\downarrow}\Big)\Big). \tag{26}$$

The breakdown of the most complicated I_4 term in Eq. (12) involves each of the cases in Eq. (18)

$$\begin{aligned} I_4 = \frac{1}{2}\Big(&\sum_{ij,r=m^\uparrow,s=n^\downarrow} \langle ij||m^\downarrow n^\uparrow\rangle \mathcal{S}^{ij,m^\uparrow n^\downarrow} + \sum_{ij,r=n^\downarrow,s=m^\uparrow} \langle ij||n^\uparrow m^\downarrow\rangle \mathcal{S}^{ij,n^\downarrow m^\uparrow} \\ &+ \sum_{ij,r\neq m^\uparrow,s=n^\downarrow} \langle ij||rn^\uparrow\rangle \mathcal{S}^{ij,rn^\downarrow}_{m^\uparrow\to m^\downarrow} + \sum_{ij,r\neq n^\downarrow,s=m^\uparrow} \langle ij||rm^\downarrow\rangle \mathcal{S}^{ij,rm^\uparrow}_{n^\downarrow\to n^\uparrow} \\ &+ \sum_{ij,r=m^\uparrow,s\neq n^\downarrow} \langle ij||m^\downarrow s\rangle \mathcal{S}^{ij,m^\uparrow s}_{n^\downarrow\to n^\uparrow} + \sum_{ij,r=n^\downarrow,s\neq m^\uparrow} \langle ij||n^\uparrow s\rangle \mathcal{S}^{ij,n^\downarrow s}_{m^\uparrow\to m^\downarrow} \\ &+ \sum_{ijrs,(r,s)\neq(m^\uparrow,n^\downarrow)} \langle ij||\tilde{r}\tilde{s}\rangle \frac{1}{\mathcal{S}}\Big(\mathcal{S}_{m^\uparrow\to m^\downarrow,n^\downarrow\to n^\uparrow}\mathcal{S}^{ij,rs} - \mathcal{S}_{r\to m^\downarrow,n^\downarrow\to n^\uparrow}\mathcal{S}^{ij,m^\uparrow s} \\ &- \mathcal{S}_{s\to m^\downarrow,n^\downarrow\to n^\uparrow}\mathcal{S}^{ij,rm^\uparrow} - \mathcal{S}_{m^\uparrow\to m^\downarrow,r\to n^\uparrow}\mathcal{S}^{ij,n^\downarrow s} \\ &- \mathcal{S}_{m^\uparrow\to m^\downarrow,s\to n^\uparrow}\mathcal{S}^{ij,rn^\downarrow} + \mathcal{S}_{r\to m^\downarrow,s\to n^\uparrow}\mathcal{S}^{ij,m^\uparrow n^\downarrow}\Big)\Big). \end{aligned} \tag{27}$$

Before tackling these summations, we can take advantage of antisymmetry relations of integrals and cofactors to reduce the number of terms that need to be expanded in the first line of Eq. (27)

$$\langle ij||n^\uparrow m^\downarrow\rangle \mathcal{S}^{ij,n^\downarrow m^\uparrow} = -\langle ij||m^\downarrow n^\uparrow\rangle \mathcal{S}^{ij,n^\downarrow m^\uparrow} = \langle ij||m^\downarrow n^\uparrow\rangle \mathcal{S}^{ij,m^\uparrow n^\downarrow}. \tag{28}$$

Therefore, we need to expand only one of these summations

$$\begin{aligned} \langle ij||m^\downarrow n^\uparrow\rangle \mathcal{S}^{ij,m^\uparrow n^\downarrow} &= \left(\langle ij|m^\downarrow n^\uparrow\rangle - \langle ij|n^\uparrow m^\downarrow\rangle\right)\mathcal{S}\left(D_{m^\uparrow i}D_{n^\downarrow j} - D_{n^\downarrow i}D_{m^\uparrow j}\right) \\ &= -\langle ij|n^\uparrow m^\downarrow\rangle \mathcal{S} D_{m^\uparrow i} D_{n^\downarrow j} - \langle ij|m^\downarrow n^\uparrow\rangle \mathcal{S} D_{n^\downarrow i} D_{m^\uparrow j}, \end{aligned} \tag{29}$$

where the two terms that are omitted in the second line are zero due to spin.

When dealing with the summations in the second and third line of Eq. (27), it is again possible to equate some terms to each other using their antisymmetry. Additionally, r and s can be swapped arbitrarily since they both span the complete occupied space. Therefore,

$$\sum_{r\neq m\uparrow}\langle ij||rn^{\uparrow}\rangle\mathcal{S}^{ij,rn^{\downarrow}}_{m^{\uparrow}\to m^{\downarrow}}=\sum_{s\neq m\uparrow}\langle ij||sn^{\uparrow}\rangle\mathcal{S}^{ij,sn^{\downarrow}}_{m^{\uparrow}\to m^{\downarrow}}=\sum_{s\neq m\uparrow}\langle ij||n^{\uparrow}s\rangle\mathcal{S}^{ij,n^{\downarrow}s}_{m^{\uparrow}\to m^{\downarrow}} \tag{30}$$

and

$$\sum_{r\neq n\downarrow}\langle ij||rm^{\downarrow}\rangle\mathcal{S}^{ij,rm^{\uparrow}}_{n^{\downarrow}\to n^{\uparrow}}=\sum_{s\neq n\downarrow}\langle ij||sm^{\downarrow}\rangle\mathcal{S}^{ij,sm^{\uparrow}}_{n^{\downarrow}\to n^{\uparrow}}=\sum_{s\neq n\downarrow}\langle ij||m^{\downarrow}s\rangle\mathcal{S}^{ij,m^{\uparrow}s}_{n^{\downarrow}\to n^{\uparrow}}. \tag{31}$$

Now, expanding the term in Eq. (30)

$$\begin{aligned}
\langle ij||rn^{\uparrow}\rangle\mathcal{S}^{ij,rn^{\downarrow}}_{m^{\uparrow}\to m^{\downarrow}} &= \frac{\langle ij||rn^{\uparrow}\rangle}{\mathcal{S}}\left(\mathcal{S}_{m^{\uparrow}\to m^{\downarrow}}\mathcal{S}^{ij,rn^{\downarrow}}-\mathcal{S}_{r\to m^{\downarrow}}\mathcal{S}^{ij,m^{\uparrow}n^{\downarrow}}-\mathcal{S}_{n^{\downarrow}\to m^{\downarrow}}\mathcal{S}^{ij,rm^{\uparrow}}\right)\\
&= -\langle ij|rn^{\uparrow}\rangle\mathcal{S}_{r\to m^{\downarrow}}D_{m^{\uparrow}i}D_{n^{\downarrow}j}+\langle ij|n^{\uparrow}r\rangle\mathcal{S}_{r\to m^{\downarrow}}D_{m^{\uparrow}i}D_{n^{\downarrow}j}+\langle ij|rn^{\uparrow}\rangle\mathcal{S}_{r\to m^{\downarrow}}D_{n^{\downarrow}i}D_{m^{\uparrow}j}\\
&\quad-\langle ij|n^{\uparrow}r\rangle\mathcal{S}_{r\to m^{\downarrow}}D_{n^{\downarrow}i}D_{m^{\uparrow}j}-\langle ij|rn^{\uparrow}\rangle\mathcal{S}_{n^{\downarrow}\to m^{\downarrow}}D_{ri}D_{m^{\uparrow}j}+\langle ij|n^{\uparrow}r\rangle\mathcal{S}_{n^{\downarrow}\to m^{\downarrow}}D_{ri}D_{m^{\uparrow}j}\\
&\quad+\langle ij|rn^{\uparrow}\rangle\mathcal{S}_{n^{\downarrow}\to m^{\downarrow}}D_{m^{\uparrow}i}D_{rj}-\langle ij|n^{\uparrow}r\rangle\mathcal{S}_{n^{\downarrow}\to m^{\downarrow}}D_{m^{\uparrow}i}D_{rj}\\
&= \mathcal{S}\Big(\langle ij|n^{\uparrow}r\rangle D_{m^{\uparrow}i}D_{n^{\downarrow}j}\sum_{j'}\left(D_{rj'}S_{j'm^{\downarrow}}\right)+\langle ij|rn^{\uparrow}\rangle D_{n^{\downarrow}i}D_{m^{\uparrow}j}\sum_{j'}\left(D_{rj'}S_{j'm^{\downarrow}}\right)\\
&\quad-\langle ij|rn^{\uparrow}\rangle D_{ri}D_{m^{\uparrow}j}\sum_{j'}\left(D_{n^{\downarrow}j'}S_{j'm^{\downarrow}}\right)+\langle ij|n^{\uparrow}r\rangle D_{ri}D_{m^{\uparrow}j}\sum_{j'}\left(D_{n^{\downarrow}j'}S_{j'm^{\downarrow}}\right)\\
&\quad+\langle ij|rn^{\uparrow}\rangle D_{m^{\uparrow}i}D_{rj}\sum_{j'}\left(D_{n^{\downarrow}j'}S_{j'm^{\downarrow}}\right)-\langle ij|n^{\uparrow}r\rangle D_{m^{\uparrow}i}D_{rj}\sum_{j'}\left(D_{n^{\downarrow}j'}S_{j'm^{\downarrow}}\right)\Big)
\end{aligned} \tag{32}$$

with the terms that are removed vanishing due to spin integration. The analogous treatment of the other term containing single excitations yields

$$\begin{aligned}
\langle ij||rm^{\downarrow}\rangle\mathcal{S}^{ij,rm^{\uparrow}}_{n^{\downarrow}\to n^{\uparrow}} &= \mathcal{S}\Big(\langle ij|m^{\downarrow}r\rangle D_{n^{\downarrow}i}D_{m^{\uparrow}j}\sum_{i'}(D_{ri'}S_{i'n^{\uparrow}})+\langle ij|rm^{\downarrow}\rangle D_{m^{\uparrow}i}D_{n^{\downarrow}j}\sum_{i'}(D_{ri'}S_{i'n^{\uparrow}})-\langle ij|rm^{\downarrow}\rangle D_{ri}D_{n^{\downarrow}j}\sum_{i'}(D_{m^{\uparrow}i'}S_{i'n^{\uparrow}})\\
&\quad+\langle ij|m^{\downarrow}r\rangle D_{ri}D_{n^{\downarrow}j}\sum_{i'}(D_{m^{\uparrow}i'}S_{i'n^{\uparrow}})+\langle ij|rm^{\downarrow}\rangle D_{n^{\downarrow}i}D_{rj}\sum_{i'}(D_{m^{\uparrow}i'}S_{i'n^{\uparrow}})-\langle ij|m^{\downarrow}r\rangle D_{n^{\downarrow}i}D_{rj}\sum_{i'}(D_{m^{\uparrow}i'}S_{i'n^{\uparrow}})\Big).
\end{aligned} \tag{33}$$

The restrictions in Eq. (27) of $r\neq m^{\uparrow}$ and $r\neq n^{\downarrow}$ for the summations involving Eqs. (32) and (33), respectively, can be lifted since it can now be seen that these conditions reduce the given term to zero.

For the last summation of I_4, we begin by noting that $\tilde{r}=r,\tilde{s}=s$ under the restrictions of this summation. Next, we expand the doubly excited determinants. The first of them has been handled above and the others expand as follows:

$$\begin{aligned}
\mathcal{S}_{r\to m^{\downarrow},n^{\downarrow}\to n^{\uparrow}} &= \frac{1}{\mathcal{S}}(\mathcal{S}_{r\to m^{\downarrow}}\mathcal{S}_{n^{\downarrow}\to n^{\uparrow}}-\mathcal{S}_{r\to n^{\uparrow}}\mathcal{S}_{n^{\downarrow}\to m^{\downarrow}})=-\mathcal{S}\sum_{i'j'}(D_{ri'}S_{i'n^{\uparrow}}D_{n^{\downarrow}j'}S_{j'm^{\downarrow}}),\\
\mathcal{S}_{s\to m^{\downarrow},n^{\downarrow}\to n^{\uparrow}} &= \frac{1}{\mathcal{S}}(\mathcal{S}_{s\to m^{\downarrow}}\mathcal{S}_{n^{\downarrow}\to n^{\uparrow}}-\mathcal{S}_{s\to n^{\uparrow}}\mathcal{S}_{n^{\downarrow}\to m^{\downarrow}})=-\mathcal{S}\sum_{i'j'}(D_{si'}S_{i'n^{\uparrow}}D_{n^{\downarrow}j'}S_{j'm^{\downarrow}}),\\
\mathcal{S}_{m^{\uparrow}\to m^{\downarrow},r\to n^{\uparrow}} &= \frac{1}{\mathcal{S}}(\mathcal{S}_{m^{\uparrow}\to m^{\downarrow}}\mathcal{S}_{r\to n^{\uparrow}}-\mathcal{S}_{m^{\uparrow}\to n^{\uparrow}}\mathcal{S}_{r\to m^{\downarrow}})=-\mathcal{S}\sum_{i'j'}(D_{m^{\uparrow}i'}S_{i'n^{\uparrow}}D_{rj'}S_{j'm^{\downarrow}}),\\
\mathcal{S}_{m^{\uparrow}\to m^{\downarrow},s\to n^{\uparrow}} &= \frac{1}{\mathcal{S}}(\mathcal{S}_{m^{\uparrow}\to m^{\downarrow}}\mathcal{S}_{s\to n^{\uparrow}}-\mathcal{S}_{m^{\uparrow}\to n^{\uparrow}}\mathcal{S}_{s\to m^{\downarrow}})=-\mathcal{S}\sum_{i'j'}(D_{m^{\uparrow}i'}S_{i'n^{\uparrow}}D_{sj'}S_{j'm^{\downarrow}}),\\
\mathcal{S}_{r\to m^{\downarrow},s\to n^{\uparrow}} &= \frac{1}{\mathcal{S}}(\mathcal{S}_{r\to m^{\downarrow}}\mathcal{S}_{s\to n^{\uparrow}}-\mathcal{S}_{r\to n^{\uparrow}}\mathcal{S}_{s\to m^{\downarrow}})=\mathcal{S}\Big(\sum_{i'j'}(D_{rj'}S_{j'm^{\downarrow}}D_{si'}S_{i'n^{\uparrow}})-\sum_{i'j'}(D_{ri'}S_{i'n^{\uparrow}}D_{sj'}S_{j'm^{\downarrow}})\Big).
\end{aligned} \tag{34}$$

Now we make the necessary replacements and expand the integrals and second cofactors for the last summation in Eq. (27)

$$\mathcal{S}\Big[- \langle ij|rs\rangle D_{ri}D_{sj}\sum_{i'j'}(D_{m\uparrow i'}S_{i'n\uparrow}D_{n\downarrow j'}S_{j'm\downarrow}) + \langle ij|rs\rangle D_{si}D_{rj}\sum_{i'j'}(D_{m\uparrow i'}S_{i'n\uparrow}D_{n\downarrow j'}S_{j'm\downarrow}) + \langle ij|sr\rangle D_{ri}D_{sj}\sum_{i'j'}(D_{m\uparrow i'}S_{i'n\uparrow}D_{n\downarrow j'}S_{j'm\downarrow})$$
$$- \langle ij|sr\rangle D_{si}D_{rj}\sum_{i'j'}(D_{m\uparrow i'}S_{i'n\uparrow}D_{n\downarrow j'}S_{j'm\downarrow}) + \langle ij|rs\rangle D_{m\uparrow i}D_{sj}\sum_{i'j'}(D_{ri'}S_{i'n\uparrow}D_{n\downarrow j'}S_{j'm\downarrow}) - \langle ij|rs\rangle D_{si}D_{m\uparrow j}\sum_{i'j'}(D_{ri'}S_{i'n\uparrow}D_{n\downarrow j'}S_{j'm\downarrow})$$
$$- \langle ij|sr\rangle D_{m\uparrow i}D_{sj}\sum_{i'j'}(D_{ri'}S_{i'n\uparrow}D_{n\downarrow j'}S_{j'm\downarrow}) + \langle ij|sr\rangle D_{si}D_{m\uparrow j}\sum_{i'j'}(D_{ri'}S_{i'n\uparrow}D_{n\downarrow j'}S_{j'm\downarrow}) + \langle ij|rs\rangle D_{ri}D_{m\uparrow j}\sum_{i'j'}(D_{si'}S_{i'n\uparrow}D_{n\downarrow j'}S_{j'm\downarrow})$$
$$- \langle ij|rs\rangle D_{m\uparrow i}D_{rj}\sum_{i'j'}(D_{si'}S_{i'n\uparrow}D_{n\downarrow j'}S_{j'm\downarrow}) - \langle ij|sr\rangle D_{ri}D_{m\uparrow j}\sum_{i'j'}(D_{si'}S_{i'n\uparrow}D_{n\downarrow j'}S_{j'm\downarrow}) + \langle ij|sr\rangle D_{m\uparrow i}D_{rj}\sum_{i'j'}(D_{si'}S_{i'n\uparrow}D_{n\downarrow j'}S_{j'm\downarrow})$$
$$+ \langle ij|rs\rangle D_{n\downarrow i}D_{sj}\sum_{i'j'}(D_{m\uparrow i'}S_{i'n\uparrow}D_{rj'}S_{j'm\downarrow}) - \langle ij|rs\rangle D_{si}D_{n\downarrow j}\sum_{i'j'}(D_{m\uparrow i'}S_{i'n\uparrow}D_{rj'}S_{j'm\downarrow}) - \langle ij|sr\rangle D_{n\downarrow i}D_{sj}\sum_{i'j'}(D_{m\uparrow i'}S_{i'n\uparrow}D_{rj'}S_{j'm\downarrow})$$
$$+ \langle ij|sr\rangle D_{si}D_{n\downarrow j}\sum_{i'j'}(D_{m\uparrow i'}S_{i'n\uparrow}D_{rj'}S_{j'm\downarrow}) + \langle ij|rs\rangle D_{ri}D_{n\downarrow j}\sum_{i'j'}(D_{m\uparrow i'}S_{i'n\uparrow}D_{sj'}S_{j'm\downarrow}) - \langle ij|rs\rangle D_{n\downarrow i}D_{rj}\sum_{i'j'}(D_{m\uparrow i'}S_{i'n\uparrow}D_{sj'}S_{j'm\downarrow})$$
$$- \langle ij|sr\rangle D_{ri}D_{n\downarrow j}\sum_{i'j'}(D_{m\uparrow i'}S_{i'n\uparrow}D_{sj'}S_{j'm\downarrow}) + \langle ij|sr\rangle D_{n\downarrow i}D_{rj}\sum_{i'j'}(D_{m\uparrow i'}S_{i'n\uparrow}D_{sj'}S_{j'm\downarrow}) + \langle ij|rs\rangle D_{m\uparrow i}D_{n\downarrow j}\sum_{i'j'}(D_{rj'}S_{j'm\downarrow}D_{si'}S_{i'n\uparrow})$$
$$- \langle ij|rs\rangle D_{n\downarrow i}D_{m\uparrow j}\sum_{i'j'}(D_{rj'}S_{j'm\downarrow}D_{si'}S_{i'n\uparrow}) - \langle ij|sr\rangle D_{m\uparrow i}D_{n\downarrow j}\sum_{i'j'}(D_{rj'}S_{j'm\downarrow}D_{si'}S_{i'n\uparrow}) + \langle ij|sr\rangle D_{n\downarrow i}D_{m\uparrow j}\sum_{i'j'}(D_{rj'}S_{j'm\downarrow}D_{si'}S_{i'n\uparrow})$$
$$- \langle ij|rs\rangle D_{m\uparrow i}D_{n\downarrow j}\sum_{i'j'}(D_{ri'}S_{i'n\uparrow}D_{sj'}S_{j'm\downarrow}) + \langle ij|rs\rangle D_{n\downarrow i}D_{m\uparrow j}\sum_{i'j'}(D_{ri'}S_{i'n\uparrow}D_{sj'}S_{j'm\downarrow}) + \langle ij|sr\rangle D_{m\uparrow i}D_{n\downarrow j}\sum_{i'j'}(D_{ri'}S_{i'n\uparrow}D_{sj'}S_{j'm\downarrow})$$
$$- \langle ij|sr\rangle D_{n\downarrow i}D_{m\uparrow j}\sum_{i'j'}(D_{ri'}S_{i'n\uparrow}D_{sj'}S_{j'm\downarrow})\Big]. \tag{35}$$

Inspection of the terms in Eq. (35) shows that four of them vanish upon spin integration. Moreover, upon the summation over r, s, the remaining 24 terms can be collected into 12 equal pairs when we again take advantage of the fact that r and s can be swapped. Again, the restrictions $(r, s) \neq (m^{\uparrow}, n^{\downarrow})$ on the summation can be lifted because the sum of the terms reduces to zero when either r or s is equal to $m^{\uparrow}$ or $n^{\downarrow}$. The total I_4 term can now be rewritten as

$$I_4 = \mathcal{S}\Big(\sum_{ij}\Big[- \langle ij|n^{\uparrow}m^{\downarrow}\rangle D_{m\uparrow i}D_{n\downarrow j} - \langle ij|m^{\downarrow}n^{\uparrow}\rangle D_{n\downarrow i}D_{m\uparrow j}\Big] + \sum_{ijr}\Big[\langle ij|n^{\uparrow}r\rangle D_{m\uparrow i}D_{n\downarrow j}\sum_{j'}(D_{rj'}S_{j'm\downarrow}) + \langle ij|rn^{\uparrow}\rangle D_{n\downarrow i}D_{m\uparrow j}\sum_{j'}(D_{rj'}S_{j'm\downarrow})$$
$$- \langle ij|rn^{\uparrow}\rangle D_{ri}D_{m\uparrow j}\sum_{j'}(D_{n\downarrow j'}S_{j'm\downarrow}) + \langle ij|n^{\uparrow}r\rangle D_{ri}D_{m\uparrow j}\sum_{j'}(D_{n\downarrow j'}S_{j'm\downarrow}) + \langle ij|rn^{\uparrow}\rangle D_{m\uparrow i}D_{rj}\sum_{j'}(D_{n\downarrow j'}S_{j'm\downarrow}) - \langle ij|n^{\uparrow}r\rangle D_{m\uparrow i}D_{rj}\sum_{j'}(D_{n\downarrow j'}S_{j'm\downarrow})$$
$$+ \langle ij|m^{\downarrow}r\rangle D_{n\downarrow i}D_{m\uparrow j}\sum_{i'}(D_{ri'}S_{i'n\uparrow}) + \langle ij|rm^{\downarrow}\rangle D_{m\uparrow i}D_{n\downarrow j}\sum_{i'}(D_{ri'}S_{i'n\uparrow}) - \langle ij|rm^{\downarrow}\rangle D_{ri}D_{n\downarrow j}\sum_{i'}(D_{m\uparrow i'}S_{i'n\uparrow}) + \langle ij|m^{\downarrow}r\rangle D_{ri}D_{n\downarrow j}\sum_{i'}(D_{m\uparrow i'}S_{i'n\uparrow})$$
$$+ \langle ij|rm^{\downarrow}\rangle D_{n\downarrow i}D_{rj}\sum_{i'}(D_{m\uparrow i'}S_{i'n\uparrow}) - \langle ij|m^{\downarrow}r\rangle D_{n\downarrow i}D_{rj}\sum_{i'}(D_{m\uparrow i'}S_{i'n\uparrow})\Big] + \sum_{ijrs}\Big[- \langle ij|rs\rangle D_{ri}D_{sj}\sum_{i'j'}(D_{m\uparrow i'}S_{i'n\uparrow}D_{n\downarrow j'}S_{j'm\downarrow})$$
$$+ \langle ij|rs\rangle D_{si}D_{rj}\sum_{i'j'}(D_{m\uparrow i'}S_{i'n\uparrow}D_{n\downarrow j'}S_{j'm\downarrow}) + \langle ij|rs\rangle D_{ri}D_{m\uparrow j}\sum_{i'j'}(D_{si'}S_{i'n\uparrow}D_{n\downarrow j'}S_{j'm\downarrow}) - \langle ij|rs\rangle D_{m\uparrow i}D_{rj}\sum_{i'j'}(D_{si'}S_{i'n\uparrow}D_{n\downarrow j'}S_{j'm\downarrow})$$
$$- \langle ij|sr\rangle D_{ri}D_{m\uparrow j}\sum_{i'j'}(D_{si'}S_{i'n\uparrow}D_{n\downarrow j'}S_{j'm\downarrow}) + \langle ij|sr\rangle D_{m\uparrow i}D_{rj}\sum_{i'j'}(D_{si'}S_{i'n\uparrow}D_{n\downarrow j'}S_{j'm\downarrow}) + \langle ij|rs\rangle D_{ri}D_{n\downarrow j}\sum_{i'j'}(D_{m\uparrow i'}S_{i'n\uparrow}D_{sj'}S_{j'm\downarrow})$$
$$- \langle ij|rs\rangle D_{n\downarrow i}D_{rj}\sum_{i'j'}(D_{m\uparrow i'}S_{i'n\uparrow}D_{sj'}S_{j'm\downarrow}) - \langle ij|sr\rangle D_{ri}D_{n\downarrow j}\sum_{i'j'}(D_{m\uparrow i'}S_{i'n\uparrow}D_{sj'}S_{j'm\downarrow}) + \langle ij|sr\rangle D_{n\downarrow i}D_{rj}\sum_{i'j'}(D_{m\uparrow i'}S_{i'n\uparrow}D_{sj'}S_{j'm\downarrow})$$
$$- \langle ij|rs\rangle D_{n\downarrow i}D_{m\uparrow j}\sum_{i'j'}(D_{rj'}S_{j'm\downarrow}D_{si'}S_{i'n\uparrow}) - \langle ij|sr\rangle D_{m\uparrow i}D_{n\downarrow j}\sum_{i'j'}(D_{rj'}S_{j'm\downarrow}D_{si'}S_{i'n\uparrow})\Big]\Big), \tag{36}$$

where the common $\mathcal{S}$ has been factored out and the $\frac{1}{2}$ canceled by the factor of 2 from the term pairings.

With all necessary terms in hand, there remain a few minor points to consider. First, it can now be seen that $\mathcal{S}$ appears in all terms in the numerator and denominator of the total energy equation and will cancel out. Second, we introduce the following function:[7]

$$Z(S_A, S_B, S) = \frac{c_1}{2c_0\sqrt{S_A S_B}} = \frac{S(S+1) + 2S_A S_B - S_A(S_A+1) - S_B(S_B+1)}{4S_A S_B}, \quad (37)$$

which replaces both the ratio of the Clebsch-Gordan coefficients and normalization factors from the spin-flipped wavefunctions. Through a proper choice of the factor Z, the desired spin state is obtained.[7] Finally, the equations up to this point remain in the spinorbital form and spin integration has only been accounted for as a means of eliminating terms that vanish. We will now make the spin integration explicit and specify the exact spin blocks of the **D** matrix that give nonzero contributions in the resulting orbital expressions. This step is necessary in preparation for the atomic-orbital equivalents of these expressions (which will be derived in Sec. III), as the latter would otherwise lose any information about the spin combinations that result in nonvanishing contributions. In the resulting expressions below, all sums run over the respective orbitals and not spinorbitals. However, the spin integration results in different ranges of summation for orbitals occupied by spin-up and spin-down electrons. These ranges will be specified by an overbar for spin-up indices and an underbar for the spin-down ones: specifically, $\bar{i}$, $\bar{j}$, and $\bar{r}$ represent the occupied orbitals of A, B, or either monomer, respectively, which can be combined with an α spin function, and $\underline{i}$, $\underline{j}$, and $\underline{r}$ denote the corresponding orbital types that can be combined with a β spin function. In terms of the inactive and active orbitals on both monomers, the summations over $\bar{i}$, $\bar{j}$, $\bar{r}$, $\underline{i}$, $\underline{j}$, and $\underline{r}$ break up into summations over (k, m), (l), (k, l, m), (k), (l, n), and (k, l, n), respectively. In this notation, the orbital equivalents of Eqs. (10), (19), (25), (26), and (36) are

$$\langle \Psi_A \Psi_B | V\mathcal{A} | \Psi_A \Psi_B \rangle = \frac{N_A! N_B!}{N!} \mathcal{S} \Bigg[W_{AB} + \sum_{\bar{i}\bar{r}} B_{ir} D^{\alpha}_{ri} + \sum_{\underline{i}\underline{r}} B_{ir} D^{\beta}_{ri} + \sum_{\bar{j}\bar{r}} A_{jr} D^{\alpha}_{rj} + \sum_{\underline{j}\underline{r}} A_{jr} D^{\beta}_{rj} + \sum_{\bar{i}\bar{j}\bar{r}\bar{s}} \langle ij|rs\rangle D^{\alpha}_{ri} D^{\alpha}_{sj} + \sum_{\bar{i}\bar{j}\underline{r}\underline{s}} \langle ij|rs\rangle D^{\alpha}_{ri} D^{\beta}_{sj} + \sum_{\underline{i}\underline{j}\bar{r}\bar{s}} \langle ij|rs\rangle D^{\beta}_{ri} D^{\alpha}_{sj}$$
$$+ \sum_{\underline{i}\underline{j}\underline{r}\underline{s}} \langle ij|rs\rangle D^{\beta}_{ri} D^{\beta}_{sj} - \sum_{\bar{i}\bar{j}\bar{r}\bar{s}} \langle ij|sr\rangle D^{\alpha}_{ri} D^{\alpha}_{sj} - \sum_{\underline{i}\underline{j}\underline{r}\underline{s}} \langle ij|sr\rangle D^{\beta}_{ri} D^{\beta}_{sj} \Bigg], \quad (38)$$

$$I_1 = -\mathcal{S} W_{AB} \left(\sum_{\underline{j}} D^{\beta}_{nj} S_{jm} \right) \left(\sum_{\bar{i}} D^{\alpha}_{mi} S_{in} \right), \quad (39)$$

$$I_2 = -\mathcal{S} \Bigg(\sum_{\bar{i}\bar{i}'} B_{im} D^{\beta}_{ni} D^{\alpha}_{mi'} S_{i'n} + \sum_{\bar{i}\bar{j}} B_{in} D^{\alpha}_{mi} D^{\beta}_{nj} S_{jm} + \sum_{\bar{i}\bar{r}} B_{ir} D^{\alpha}_{ri} \sum_{\bar{i}'\bar{j}} D^{\alpha}_{mi'} S_{i'n} D^{\beta}_{nj} S_{jm} + \sum_{\underline{i}\underline{r}} B_{ir} D^{\beta}_{ri} \sum_{\bar{i}'\underline{j}} D^{\alpha}_{mi'} S_{i'n} D^{\beta}_{nj} S_{jm}$$
$$- \sum_{\bar{i}\bar{r}} B_{ir} D^{\alpha}_{mi} \sum_{\bar{i}'\underline{j}} D^{\alpha}_{ri'} S_{i'n} D^{\beta}_{nj} S_{jm} - \sum_{\underline{i}\underline{r}} B_{ir} D^{\beta}_{ni} \sum_{\bar{i}'\underline{j}} D^{\alpha}_{mi'} S_{i'n} D^{\beta}_{rj} S_{jm} \Bigg), \quad (40)$$

$$I_3 = -\mathcal{S} \Bigg(\sum_{\bar{i}\bar{j}} A_{jm} D^{\beta}_{nj} D^{\alpha}_{mi} S_{in} + \sum_{\bar{j}\bar{j}'} A_{jn} D^{\alpha}_{mj} D^{\beta}_{nj'} S_{j'm} + \sum_{\bar{j}\bar{r}} A_{jr} D^{\alpha}_{rj} \sum_{\bar{i}\bar{j}'} D^{\alpha}_{mi} S_{in} D^{\beta}_{nj'} S_{j'm} + \sum_{\underline{j}\underline{r}} A_{jr} D^{\beta}_{rj} \sum_{\bar{i}\underline{j}'} D^{\alpha}_{mi} S_{in} D^{\beta}_{nj'} S_{j'm}$$
$$- \sum_{\bar{j}\bar{r}} A_{jr} D^{\alpha}_{mj} \sum_{\bar{i}\underline{j}'} D^{\alpha}_{ri} S_{in} D^{\beta}_{nj'} S_{j'm} - \sum_{\underline{j}\underline{r}} A_{jr} D^{\beta}_{nj} \sum_{\bar{i}\underline{j}'} D^{\alpha}_{mi} S_{in} D^{\beta}_{rj'} S_{j'm} \Bigg), \quad (41)$$

$$
\begin{aligned}
I_4 = \mathcal{S}\Big(&- \sum_{\bar{i}\underline{j}} \langle ij|nm\rangle D^{\alpha}_{mi} D^{\beta}_{nj} - \sum_{\bar{i}\underline{j}} \langle ij|mn\rangle D^{\beta}_{ni} D^{\alpha}_{mj} + \sum_{\bar{i}\underline{j}\bar{r}\underline{j}'} \langle ij|nr\rangle D^{\alpha}_{mi} D^{\beta}_{nj} D^{\beta}_{rj'} S_{j'm} + \sum_{\bar{i}\underline{j}\underline{r}\bar{j}'} \langle ij|rn\rangle D^{\beta}_{ni} D^{\alpha}_{mj} D^{\beta}_{rj'} S_{j'm} - \sum_{\bar{i}\underline{j}\bar{r}\bar{j}'} \langle ij|rn\rangle D^{\alpha}_{ri} D^{\alpha}_{mj} D^{\beta}_{nj'} S_{j'm} \\
&- \sum_{\bar{i}\underline{j}\underline{r}\bar{j}'} \langle ij|rn\rangle D^{\beta}_{ri} D^{\alpha}_{mj} D^{\beta}_{nj'} S_{j'm} + \sum_{\bar{i}\underline{j}\bar{r}\bar{j}'} \langle ij|nr\rangle D^{\alpha}_{ri} D^{\alpha}_{mj} D^{\beta}_{nj'} S_{j'm} + \sum_{\bar{i}\underline{j}\bar{r}\bar{j}'} \langle ij|rn\rangle D^{\alpha}_{mi} D^{\alpha}_{rj} D^{\beta}_{nj'} S_{j'm} - \sum_{\bar{i}\underline{j}\bar{r}\underline{j}'} \langle ij|nr\rangle D^{\alpha}_{mi} D^{\alpha}_{rj} D^{\beta}_{nj'} S_{j'm} \\
&- \sum_{\bar{i}\underline{j}\underline{r}\underline{j}'} \langle ij|nr\rangle D^{\alpha}_{mi} D^{\beta}_{rj} D^{\beta}_{nj'} S_{j'm} + \sum_{\bar{i}\underline{j}\bar{r}\bar{i}'} \langle ij|mr\rangle D^{\beta}_{ni} D^{\alpha}_{mj} D^{\alpha}_{ri'} S_{i'n} + \sum_{\bar{i}\underline{j}\bar{r}\bar{i}'} \langle ij|rm\rangle D^{\alpha}_{mi} D^{\beta}_{nj} D^{\alpha}_{ri'} S_{i'n} - \sum_{\bar{i}\underline{j}\bar{r}\bar{i}'} \langle ij|rm\rangle D^{\alpha}_{ri} D^{\beta}_{nj} D^{\alpha}_{mi'} S_{i'n} \\
&- \sum_{\bar{i}\underline{j}\underline{r}\bar{i}'} \langle ij|rm\rangle D^{\beta}_{ri} D^{\beta}_{nj} D^{\alpha}_{mi'} S_{i'n} + \sum_{\bar{i}\underline{j}\underline{r}\bar{i}'} \langle ij|mr\rangle D^{\beta}_{ri} D^{\beta}_{nj} D^{\alpha}_{mi'} S_{i'n} + \sum_{\bar{i}\underline{j}\underline{r}\bar{i}'} \langle ij|rm\rangle D^{\beta}_{ni} D^{\beta}_{rj} D^{\alpha}_{mi'} S_{i'n} - \sum_{\bar{i}\underline{j}\bar{r}\bar{i}'} \langle ij|mr\rangle D^{\beta}_{ni} D^{\alpha}_{rj} D^{\alpha}_{mi'} S_{i'n} \\
&- \sum_{\bar{i}\underline{j}\underline{r}\bar{i}'} \langle ij|mr\rangle D^{\beta}_{ni} D^{\beta}_{rj} D^{\alpha}_{mi'} S_{i'n} - \sum_{\bar{i}\bar{j}\bar{r}\underline{s}\bar{i}'\underline{j}'} \langle ij|rs\rangle D^{\alpha}_{ri} D^{\alpha}_{sj} D^{\alpha}_{mi'} S_{i'n} D^{\beta}_{nj'} S_{j'm} - \sum_{\bar{i}\underline{j}\bar{r}\underline{s}\bar{i}'\underline{j}'} \langle ij|rs\rangle D^{\alpha}_{ri} D^{\beta}_{sj} D^{\alpha}_{mi'} S_{i'n} D^{\beta}_{nj'} S_{j'm} \\
&- \sum_{\bar{i}\underline{j}\underline{r}\underline{s}\bar{i}'\underline{j}'} \langle ij|rs\rangle D^{\beta}_{ri} D^{\alpha}_{sj} D^{\alpha}_{mi'} S_{i'n} D^{\beta}_{nj'} S_{j'm} - \sum_{\bar{i}\underline{j}\underline{r}\underline{s}\bar{i}'\underline{j}'} \langle ij|rs\rangle D^{\beta}_{ri} D^{\beta}_{sj} D^{\alpha}_{mi'} S_{i'n} D^{\beta}_{nj'} S_{j'm} + \sum_{\bar{i}\bar{j}\underline{r}\underline{s}\bar{i}'\underline{j}'} \langle ij|rs\rangle D^{\alpha}_{si} D^{\alpha}_{rj} D^{\alpha}_{mi'} S_{i'n} D^{\beta}_{nj'} S_{j'm} \\
&+ \sum_{\bar{i}\underline{j}\underline{r}\underline{s}\bar{i}'\underline{j}'} \langle ij|rs\rangle D^{\beta}_{si} D^{\alpha}_{rj} D^{\alpha}_{mi'} S_{i'n} D^{\beta}_{nj'} S_{j'm} + \sum_{\bar{i}\bar{j}\underline{r}\underline{s}\bar{i}'\underline{j}'} \langle ij|rs\rangle D^{\alpha}_{ri} D^{\alpha}_{mj} D^{\alpha}_{si'} S_{i'n} D^{\beta}_{nj'} S_{j'm} + \sum_{\bar{i}\underline{j}\bar{r}\underline{s}\bar{i}'\underline{j}'} \langle ij|rs\rangle D^{\beta}_{ri} D^{\alpha}_{mj} D^{\alpha}_{si'} S_{i'n} D^{\beta}_{nj'} S_{j'm} \\
&- \sum_{\bar{i}\bar{j}\bar{r}\underline{s}\bar{i}'\underline{j}'} \langle ij|rs\rangle D^{\alpha}_{mi} D^{\alpha}_{rj} D^{\alpha}_{si'} S_{i'n} D^{\beta}_{nj'} S_{j'm} - \sum_{\bar{i}\bar{j}\bar{r}\underline{s}\bar{i}'\underline{j}'} \langle ij|sr\rangle D^{\alpha}_{ri} D^{\alpha}_{mj} D^{\alpha}_{si'} S_{i'n} D^{\beta}_{nj'} S_{j'm} + \sum_{\bar{i}\bar{j}\bar{r}\underline{s}\bar{i}'\underline{j}'} \langle ij|sr\rangle D^{\alpha}_{mi} D^{\alpha}_{rj} D^{\alpha}_{si'} S_{i'n} D^{\beta}_{nj'} S_{j'm} \\
&+ \sum_{\bar{i}\underline{j}\underline{r}\bar{s}\bar{i}'\underline{j}'} \langle ij|sr\rangle D^{\alpha}_{mi} D^{\beta}_{rj} D^{\alpha}_{si'} S_{i'n} D^{\beta}_{nj'} S_{j'm} + \sum_{\bar{i}\underline{j}\underline{r}\bar{s}\bar{i}'\underline{j}'} \langle ij|rs\rangle D^{\alpha}_{ri} D^{\beta}_{nj} D^{\alpha}_{mi'} S_{i'n} D^{\beta}_{sj'} S_{j'm} + \sum_{\bar{i}\underline{j}\underline{r}\underline{s}\bar{i}'\underline{j}'} \langle ij|rs\rangle D^{\beta}_{ri} D^{\beta}_{nj} D^{\alpha}_{mi'} S_{i'n} D^{\beta}_{sj'} S_{j'm} \\
&- \sum_{\bar{i}\underline{j}\underline{r}\underline{s}\bar{i}'\underline{j}'} \langle ij|rs\rangle D^{\beta}_{ni} D^{\beta}_{rj} D^{\alpha}_{mi'} S_{i'n} D^{\beta}_{sj'} S_{j'm} - \sum_{\bar{i}\underline{j}\underline{r}\underline{s}\bar{i}'\underline{j}'} \langle ij|sr\rangle D^{\beta}_{ri} D^{\beta}_{nj} D^{\alpha}_{mi'} S_{i'n} D^{\beta}_{sj'} S_{j'm} + \sum_{\bar{i}\bar{j}\underline{r}\underline{s}\bar{i}'\underline{j}'} \langle ij|sr\rangle D^{\beta}_{ni} D^{\alpha}_{rj} D^{\alpha}_{mi'} S_{i'n} D^{\beta}_{sj'} S_{j'm} \\
&+ \sum_{\bar{i}\underline{j}\underline{r}\underline{s}\bar{i}'\underline{j}'} \langle ij|sr\rangle D^{\beta}_{ni} D^{\beta}_{rj} D^{\alpha}_{mi'} S_{i'n} D^{\beta}_{sj'} S_{j'm} - \sum_{\bar{i}\underline{j}\underline{r}\underline{s}\bar{i}'\underline{j}'} \langle ij|rs\rangle D^{\beta}_{ni} D^{\alpha}_{mj} D^{\beta}_{rj'} S_{j'm} D^{\alpha}_{si'} S_{i'n} - \sum_{\bar{i}\underline{j}\underline{r}\underline{s}\bar{i}'\underline{j}'} \langle ij|sr\rangle D^{\alpha}_{mi} D^{\beta}_{nj} D^{\beta}_{rj'} S_{j'm} D^{\alpha}_{si'} S_{i'n} \Big).
\end{aligned}
\tag{42}
$$

III. ATOMIC ORBITAL FORMALISM

We now recast the equations derived in Sec. II from molecular orbitals to atomic orbitals, which provides significant computational benefits as integral transformation is avoided and efficient generalized Coulomb and exchange matrix codes can be utilized.[7,16,20,21] Let us start from the expression for the spin diagonal component of the numerator [Eq. (38)]. In the AO form, the leading nuclear repulsion term stays the same, so we move to the first attractive potential term

$$
\begin{aligned}
\sum_{\bar{i}r} \mathrm{B}_{ir} D^{\alpha}_{ri} &= \sum_{\bar{i}r\mathrm{KL}} \mathrm{C}_{i\mathrm{K}} \mathrm{B}_{\mathrm{KL}} \mathrm{C}_{r\mathrm{L}} D^{\alpha}_{ri} = \sum_{\bar{i}r\mathrm{KL}} \mathrm{B}_{\mathrm{LK}} \mathrm{C}_{r\mathrm{L}} D^{\alpha}_{ri} \mathrm{C}_{i\mathrm{K}} \\
&= \sum_{\mathrm{KL}} \mathrm{B}_{\mathrm{LK}} (\mathbf{D}^{\mathrm{RI}}_{\alpha})_{\mathrm{LK}} = \mathbf{B} \cdot \mathbf{D}^{\mathrm{RI}}_{\alpha},
\end{aligned}
\tag{43}
$$

where $\mathbf{X} \cdot \mathbf{Y} = \sum_{\mathrm{KL}} \mathbf{X}_{\mathrm{KL}} \mathbf{Y}_{\mathrm{KL}}$ is the matrix dot product, the capital letter indices span the AO set, $\mathbf{B}$ is the nuclear potential matrix of monomer B in the AO basis, $\mathrm{C}_{r\mathrm{K}}$ are the SCF coefficients of spinorbital r, and $\mathbf{D}^{\mathrm{RI}}_{\alpha}$, defined by the above equation, is the representation of D^{α}_{ri} in the AO basis. We use capital letters in the superscript of $\mathbf{D}^{\mathrm{RI}}_{\alpha}$ to remind the reader that R, I are no longer the indices of the matrix but merely specify the spinorbital spaces over which the contributions have been summed. The second term involving B_{ir} results in a fully analogous contribution $\mathbf{B} \cdot \mathbf{D}^{\mathrm{RI}}_{\beta}$, and we can now define

$$
\mathbf{D}^{\mathrm{RI}}_{\alpha\beta} = \mathbf{D}^{\mathrm{RI}}_{\alpha} + \mathbf{D}^{\mathrm{RI}}_{\beta},
\tag{44}
$$

as the AO representations of the $\mathbf{D}$ matrices, unlike the MO ones, are all the same size and can be added. Breaking the indices into the inactive and active ones, corresponding to a further restriction of the summations in the definition of $\mathbf{D}^{\mathrm{RI}}_{\alpha}$ and $\mathbf{D}^{\mathrm{RI}}_{\beta}$, yields

$$
\mathbf{D}^{\mathrm{RI}}_{\alpha} = \mathbf{D}^{\mathrm{KK}}_{\alpha} + \mathbf{D}^{\mathrm{LK}}_{\alpha} + \mathbf{D}^{\mathrm{MK}}_{\alpha} + \mathbf{D}^{\mathrm{KM}}_{\alpha} + \mathbf{D}^{\mathrm{LM}}_{\alpha} + \mathbf{D}^{\mathrm{MM}}_{\alpha},
\tag{45}
$$

$$
\mathbf{D}^{\mathrm{RI}}_{\beta} = \mathbf{D}^{\mathrm{KK}}_{\beta} + \mathbf{D}^{\mathrm{LK}}_{\beta} + \mathbf{D}^{\mathrm{NK}}_{\beta}.
\tag{46}
$$

Note that the different $\mathbf{D}$ matrices are in general not symmetric. The derivation of the remaining nuclear attraction terms follows an analogous path, giving

$$
\sum_{\bar{j}r} \mathrm{A}_{jr} D^{\alpha}_{rj} + \sum_{\underline{j}r} \mathrm{A}_{jr} D^{\beta}_{rj} = \mathbf{A} \cdot \mathbf{D}^{\mathrm{RJ}}_{\alpha\beta}.
\tag{47}
$$

The two-electron terms give way to similar transformations of the **D** matrices. Some representative spin-block contributions are

$$\sum_{\bar{i}\bar{j}\bar{r}\bar{s}}\langle ij|rs\rangle D^{\alpha}_{ri}D^{\alpha}_{sj} = \sum_{KLMN}(\mathbf{D}^{RJ}_{\alpha})_{KL}\langle KM|LN\rangle(\mathbf{D}^{RJ}_{\alpha})_{MN} = \mathbf{D}^{RI}_{\alpha}\cdot\mathbf{J}[\mathbf{D}^{RJ}_{\alpha}], \tag{48}$$

$$\sum_{\bar{i}\bar{j}\bar{r}\bar{s}}\langle ij|sr\rangle D^{\alpha}_{ri}D^{\alpha}_{sj} = \sum_{KLMN}(\mathbf{D}^{RJ}_{\alpha})_{KL}\langle KN|ML\rangle(\mathbf{D}^{RJ}_{\alpha})_{MN} = \mathbf{D}^{RI}_{\alpha}\cdot\mathbf{K}[\mathbf{D}^{RJ}_{\alpha}], \tag{49}$$

where $\mathbf{J}[\mathbf{X}]$ and $\mathbf{K}[\mathbf{X}]$ are the generalized Coulomb and exchange matrices, defined as follows:

$$\mathbf{J}[\mathbf{X}]_{KL} = \sum_{MN}\langle KM|LN\rangle\mathbf{X}_{MN} \qquad \mathbf{K}[\mathbf{X}]_{KL} = \sum_{MN}\langle KN|ML\rangle\mathbf{X}_{MN}. \tag{50}$$

The other spin blocks give rise to analogous contributions and the complete spin diagonal term can now be written in the AO form as

$$\langle\Psi_A\Psi_B|V\mathcal{A}|\Psi_A\Psi_B\rangle = \frac{N_A!N_B!}{N!}\mathcal{S}\Big[W_{AB} + \mathbf{B}\cdot\mathbf{D}^{RI}_{\alpha\beta} + \mathbf{A}\cdot\mathbf{D}^{RJ}_{\alpha\beta} + \mathbf{D}^{RI}_{\alpha\beta}\cdot\mathbf{J}[\mathbf{D}^{RJ}_{\alpha\beta}] - \mathbf{D}^{RI}_{\alpha}\cdot\mathbf{K}[\mathbf{D}^{RJ}_{\alpha}] - \mathbf{D}^{RI}_{\beta}\cdot\mathbf{K}[\mathbf{D}^{RJ}_{\beta}]\Big]. \tag{51}$$

Next, we move to the spin-flipped term of the numerator [Eq. (12)]. The I_1 term in Eq. (12) is the nuclear repulsion part for the spin-flip term and is explicitly defined in Eq. (39). This term is transformed into the AO basis as follows:

$$\begin{aligned} I_1 &= -W_{AB}\mathcal{S}\sum_{\bar{i}\bar{j}KLMN} D^{\beta}_{nj}C_{jK}S^{AO}_{KL}C_{mL}D^{\alpha}_{mi}C_{iM}S^{AO}_{MN}C_{nN} \\ &= -W_{AB}\mathcal{S}\sum_{\bar{i}\bar{j}KLMN}(C_{nN}D^{\beta}_{nj}C_{jK}S^{AO}_{KL})\cdot(C_{mL}D^{\alpha}_{mi}C_{iM}S^{AO}_{MN}) \\ &= -W_{AB}\mathcal{S}\sum_{LN}(\mathbf{D}^{NJ}_{\beta}\mathbf{S}^{AO})_{NL}(\mathbf{D}^{MI}_{\alpha}\mathbf{S}^{AO})_{LN} \\ &= -W_{AB}\mathcal{S}[(\mathbf{D}^{NJ}_{\beta}\mathbf{S}^{AO})\cdot(\mathbf{D}^{MI}_{\alpha}\mathbf{S}^{AO})^{T}]. \end{aligned} \tag{52}$$

Note that we used the fact that the orbitals are spin-restricted so that, for example, $C_{m\uparrow L} = C_{m\downarrow L} = C_{mL}$. The term in the last line of Eq. (52) also constitutes the spin-flip part of the denominator after dividing by W_{AB}. Using the same methods as above, the I_2 and I_3 terms of Eq. (12) transform as

$$\begin{aligned} I_2 = -\mathcal{S}\Big(&\mathbf{B}\cdot(\mathbf{D}^{MI}_{\alpha}\mathbf{S}^{AO}\mathbf{D}^{NI}_{\beta}) + \mathbf{B}\cdot(\mathbf{D}^{NJ}_{\beta}\mathbf{S}^{AO}\mathbf{D}^{MI}_{\alpha}) \\ &+ (\mathbf{B}\cdot\mathbf{D}^{RI}_{\alpha\beta})\big((\mathbf{D}^{MI}_{\alpha}\mathbf{S}^{AO}\mathbf{D}^{NJ}_{\beta})\cdot\mathbf{S}^{AO}\big) \\ &- \mathbf{B}\cdot(\mathbf{D}^{RI}_{\alpha}\mathbf{S}^{AO}\mathbf{D}^{NJ}_{\beta}\mathbf{S}^{AO}\mathbf{D}^{MI}_{\alpha}) - \mathbf{B}\cdot(\mathbf{D}^{RJ}_{\beta}\mathbf{S}^{AO}\mathbf{D}^{MI}_{\alpha}\mathbf{S}^{AO}\mathbf{D}^{NI}_{\beta})\Big) \end{aligned} \tag{53}$$

and

$$\begin{aligned} I_3 = -\mathcal{S}\Big(&\mathbf{A}\cdot(\mathbf{D}^{MI}_{\alpha}\mathbf{S}^{AO}\mathbf{D}^{NJ}_{\beta}) + \mathbf{A}\cdot(\mathbf{D}^{NJ}_{\beta}\mathbf{S}^{AO}\mathbf{D}^{MJ}_{\alpha}) \\ &+ (\mathbf{A}\cdot\mathbf{D}^{RJ}_{\alpha\beta})\big((\mathbf{D}^{MI}_{\alpha}\mathbf{S}^{AO}\mathbf{D}^{NJ}_{\beta})\cdot\mathbf{S}^{AO}\big) \\ &- \mathbf{A}\cdot(\mathbf{D}^{RI}_{\alpha}\mathbf{S}^{AO}\mathbf{D}^{NJ}_{\beta}\mathbf{S}^{AO}\mathbf{D}^{MJ}_{\alpha}) - \mathbf{A}\cdot(\mathbf{D}^{RJ}_{\beta}\mathbf{S}^{AO}\mathbf{D}^{MI}_{\alpha}\mathbf{S}^{AO}\mathbf{D}^{NJ}_{\beta})\Big). \end{aligned} \tag{54}$$

Finally, I_4 transforms into the following form:

$$\begin{aligned} I_4 = \mathcal{S}\Big(&-\mathbf{D}^{MI}_{\alpha}\cdot\mathbf{K}[\mathbf{D}^{NJ}_{\beta}]^{T} - \mathbf{D}^{NI}_{\beta}\cdot\mathbf{K}[\mathbf{D}^{MJ}_{\alpha}]^{T} + \mathbf{K}[\mathbf{D}^{NJ}_{\beta}]^{T}\cdot(\mathbf{D}^{RJ}_{\beta}\mathbf{S}^{AO}\mathbf{D}^{MI}_{\alpha}) + \mathbf{K}[\mathbf{D}^{NI}_{\beta}]^{T}\cdot(\mathbf{D}^{RJ}_{\beta}\mathbf{S}^{AO}\mathbf{D}^{MJ}_{\alpha}) - \mathbf{J}[\mathbf{D}^{RI}_{\alpha\beta}]\cdot(\mathbf{D}^{NJ}_{\beta}\mathbf{S}^{AO}\mathbf{D}^{MJ}_{\alpha}) \\ &+ \mathbf{K}[\mathbf{D}^{RI}_{\alpha}]^{T}\cdot(\mathbf{D}^{NJ}_{\beta}\mathbf{S}^{AO}\mathbf{D}^{MJ}_{\alpha}) + \mathbf{K}[\mathbf{D}^{RJ}_{\alpha}]^{T}\cdot(\mathbf{D}^{NJ}_{\beta}\mathbf{S}^{AO}\mathbf{D}^{MI}_{\alpha}) - \mathbf{J}[\mathbf{D}^{RJ}_{\alpha\beta}]\cdot(\mathbf{D}^{NJ}_{\beta}\mathbf{S}^{AO}\mathbf{D}^{MI}_{\alpha}) + \mathbf{K}[\mathbf{D}^{MJ}_{\alpha}]^{T}\cdot(\mathbf{D}^{RI}_{\alpha}\mathbf{S}^{AO}\mathbf{D}^{NI}_{\beta}) + \mathbf{K}[\mathbf{D}^{MI}_{\alpha}]^{T}\cdot(\mathbf{D}^{RI}_{\alpha}\mathbf{S}^{AO}\mathbf{D}^{NJ}_{\beta}) \\ &- \mathbf{J}[\mathbf{D}^{RI}_{\alpha\beta}]\cdot(\mathbf{D}^{MI}_{\alpha}\mathbf{S}^{AO}\mathbf{D}^{NJ}_{\beta}) + \mathbf{K}[\mathbf{D}^{RI}_{\beta}]^{T}\cdot(\mathbf{D}^{MI}_{\alpha}\mathbf{S}^{AO}\mathbf{D}^{NJ}_{\beta}) + \mathbf{K}[\mathbf{D}^{RJ}_{\beta}]^{T}\cdot(\mathbf{D}^{MI}_{\alpha}\mathbf{S}^{AO}\mathbf{D}^{NI}_{\beta}) - \mathbf{J}[\mathbf{D}^{RJ}_{\alpha\beta}]\cdot(\mathbf{D}^{MI}_{\alpha}\mathbf{S}^{AO}\mathbf{D}^{NI}_{\beta}) \\ &- (\mathbf{D}^{RI}_{\alpha\beta}\cdot\mathbf{J}[\mathbf{D}^{RJ}_{\alpha\beta}])((\mathbf{D}^{MI}_{\alpha}\mathbf{S}^{AO}\mathbf{D}^{NJ}_{\beta})\cdot\mathbf{S}^{AO}) + (\mathbf{D}^{RI}_{\alpha}\cdot\mathbf{K}[\mathbf{D}^{RJ}_{\alpha}]^{T})((\mathbf{D}^{MI}_{\alpha}\mathbf{S}^{AO}\mathbf{D}^{NJ}_{\beta})\cdot\mathbf{S}^{AO}) + (\mathbf{D}^{RI}_{\beta}\cdot\mathbf{K}[\mathbf{D}^{RJ}_{\beta}]^{T})((\mathbf{D}^{MI}_{\alpha}\mathbf{S}^{AO}\mathbf{D}^{NJ}_{\beta})\cdot\mathbf{S}^{AO}) \\ &+ \mathbf{J}[\mathbf{D}^{RI}_{\alpha\beta}]\cdot(\mathbf{D}^{RI}_{\alpha}\mathbf{S}^{AO}\mathbf{D}^{NJ}_{\beta}\mathbf{S}^{AO}\mathbf{D}^{MJ}_{\alpha}) - \mathbf{K}[\mathbf{D}^{RJ}_{\alpha}]^{T}\cdot(\mathbf{D}^{RI}_{\alpha}\mathbf{S}^{AO}\mathbf{D}^{NJ}_{\beta}\mathbf{S}^{AO}\mathbf{D}^{MI}_{\alpha}) - \mathbf{K}[\mathbf{D}^{RI}_{\beta}]^{T}\cdot(\mathbf{D}^{RI}_{\alpha}\mathbf{S}^{AO}\mathbf{D}^{NJ}_{\beta}\mathbf{S}^{AO}\mathbf{D}^{MJ}_{\alpha}) \\ &+ \mathbf{J}[\mathbf{D}^{RJ}_{\alpha\beta}]\cdot(\mathbf{D}^{RI}_{\alpha}\mathbf{S}^{AO}\mathbf{D}^{NJ}_{\beta}\mathbf{S}^{AO}\mathbf{D}^{MI}_{\alpha}) + \mathbf{J}[\mathbf{D}^{RI}_{\alpha\beta}]\cdot(\mathbf{D}^{RJ}_{\beta}\mathbf{S}^{AO}\mathbf{D}^{MI}_{\alpha}\mathbf{S}^{AO}\mathbf{D}^{NJ}_{\beta}) - \mathbf{K}[\mathbf{D}^{RJ}_{\beta}]^{T}\cdot(\mathbf{D}^{RJ}_{\beta}\mathbf{S}^{AO}\mathbf{D}^{MI}_{\alpha}\mathbf{S}^{AO}\mathbf{D}^{NI}_{\beta}) - \mathbf{K}[\mathbf{D}^{RI}_{\beta}]^{T}\cdot(\mathbf{D}^{RJ}_{\beta}\mathbf{S}^{AO}\mathbf{D}^{MI}_{\alpha}\mathbf{S}^{AO}\mathbf{D}^{NJ}_{\beta}) \\ &+ \mathbf{J}[\mathbf{D}^{RJ}_{\alpha\beta}]\cdot(\mathbf{D}^{RJ}_{\beta}\mathbf{S}^{AO}\mathbf{D}^{MI}_{\alpha}\mathbf{S}^{AO}\mathbf{D}^{NI}_{\beta}) - (\mathbf{D}^{RJ}_{\beta}\mathbf{S}^{AO}\mathbf{D}^{MJ}_{\alpha})\cdot\mathbf{K}[\mathbf{D}^{RI}_{\alpha}\mathbf{S}^{AO}\mathbf{D}^{NI}_{\beta}]^{T} - (\mathbf{D}^{RJ}_{\beta}\mathbf{S}^{AO}\mathbf{D}^{MI}_{\alpha})\cdot\mathbf{K}[\mathbf{D}^{RI}_{\alpha}\mathbf{S}^{AO}\mathbf{D}^{NJ}_{\beta}]^{T}\Big). \end{aligned} \tag{55}$$

Note that the final formula (55) can be written in many equivalent ways due to the "Hermitian-like" symmetry of the Coulomb and exchange matrices

$$\mathbf{X}\cdot\mathbf{J}[\mathbf{Y}] = \mathbf{Y}\cdot\mathbf{J}[\mathbf{X}] \qquad \mathbf{X}\cdot\mathbf{K}[\mathbf{Y}] = \mathbf{Y}\cdot\mathbf{K}[\mathbf{X}]. \tag{56}$$

The form adopted in Eq. (55) minimizes the number of Coulomb and exchange matrix evaluations necessary.

IV. RESULTS

Both the MO and AO formulas for the newly developed $E^{(10)}_{exch}$ correction were implemented using PSI4[16] and the PSI4NUMPY framework.[15] The agreement between the high-spin SF-SAPT results and the conventional ROHF-based SAPT in PSI4[6] was verified for the test systems where the single-spin-flip approximation is exact, e.g., the systems containing Li. For other systems, the 1-flip values are compared with the exact

high-spin $E^{(10)}_{exch}$ to determine the adequacy of the proposed approximation.

The aug-cc-pVTZ basis set[22–24] was used for the Li···Li, Li···N, N···N, and Mn···Mn complexes. The results for the pancake bonded systems were obtained in the aug-cc-pVDZ basis set.[22,23] To allow for comparison with the previous results, the O_2···O_2 calculations were performed in the ANO-VTZ basis set[25,26] and the Li···H calculations were performed in the basis set of Ref. 27. When the density fitting (DF) approximation was utilized, the def2-QZVPP/JKFIT[28] sets were used for Li and Mn atoms as well as for O_2 ···O_2. All other atoms used the aug-cc-pVXZ/JKFIT[29] sets with X being the same as for the orbital basis set. Throughout the discussion below, we will refer to the exchange splitting between the highest and lowest spin states of the complexes, which is defined as

$$\Delta E^{(10)}_{exch} = E^{(10)}_{exch}(S = S_A + S_B) - E^{(10)}_{exch}(S = |S_A - S_B|). \quad (57)$$

For the Mn···Mn complex, difficulties converging the ROHF iterations in PSI4 led us to use MOLPRO[30] for the monomer calculations. The ROHF orbital energies and vectors were read from the MOLPRO output and passed into PSI4NUMPY. It should also be noted that in the initial SF-SAPT work (Ref. 7), the Mn···Mn results were computed from incorrectly converged monomer ROHF wavefunctions. Therefore, the corrected SF-SAPT(S^2) results for this system will be presented below.

A. Diatomics and the O_2 dimer

The first of the smaller test complexes that we consider is Li···H, where only one pair of electrons can be exchanged. As already mentioned, the single-spin-flip approximation is exact whenever at least one monomer in the complex is a doublet. The values of $E^{(10)}_{exch}$ calculated in both approximations for the singlet and triplet states of the Li···H complex are provided in Table I, along with the $\Delta E^{(10)}_{exch}$ value and the full configuration interaction (FCI) based SAPT results of Ref. 27. While the 1-flip results agree remarkably well with the FCI-based values, confirming that the effects of intramolecular correlation are minuscule for this system as observed earlier,[7] there is a noticeable discrepancy between these values and the S^2 results. While it may come as a surprise, the S^2 approximation in the commonly used form is not exact even in this case. During the derivation of the $E^{(10)}_{exch}$ correction in the S^2 approximation,[31] the following intermediate formula is reached:

$$\langle V\rangle\langle\mathcal{P}\rangle + E^{(10)}_{exch} + E^{(10)}_{exch}\langle\mathcal{P}\rangle = \langle V\mathcal{P}\rangle, \quad (58)$$

where $\mathcal{P}$ is the single-exchange operator. A related formula is obtained during the derivation of the $E^{(10)}_{exch}(S^2)$ correction in SF-SAPT [Eq. (9) of Ref. 7]. At that point, the $E^{(10)}_{exch}\langle\mathcal{P}\rangle$ term on the lhs of Eq. (58) is normally neglected because it is at least of the order S^4 [S^2 in $E^{(10)}_{exch}$ and S^2 in $\langle\mathcal{P}\rangle$]. However, this term is not zero even for systems such as Li···H where only a single electron exchange is possible. We have verified that the differences between the S^2 and nonapproximated results for Li···H originate solely from the removal of the $E^{(10)}_{exch}\langle\mathcal{P}\rangle$ term.

For the Li dimer and the Li···N complex, the single-spin-flip $E^{(10)}_{exch}$ result is exact due to the doublet nature of a lithium atom, meaning there is only one spin to be flipped. Figure 1 shows how severely the $E^{(10)}_{exch}$ values for the lithium dimer are affected by the S^2 approximation. The deviation of the S^2 results from exact $E^{(10)}_{exch}$ is already visible at the van der Waals minimum of the triplet state (7.9 bohrs), and is on its way to catastrophic failure at the chemical minimum of the singlet state (5.0 bohrs). At around 4.2 bohrs, the splitting predicted by $E^{(10)}_{exch}(S^2)$ changes sign, leading to an unphysical energetic ordering of the two states. It can also be seen that the S^2 value for the singlet deviates from its complete $E^{(10)}_{exch}$ counterpart at a slower rate than for the triplet. Similar results are seen for the Li···N complex in Fig. 2. The two methods give equivalent results near the 10.2 bohr van der Waals minimum for the quintet state of this system

TABLE I. Singlet and triplet $E^{(10)}_{exch}$ values and the $\Delta E^{(10)}_{exch}$ splitting (cm^{-1}) in the Li···H complex as a function of interatomic distance (bohr) from FCI-based SAPT[27] and SF-SAPT with and without the single-exchange approximation.

	$E^{(10)}_{exch}(S=0)$			$E^{(10)}_{exch}(S=1)$			$\Delta E^{(10)}_{exch}$		
R	FCI	S^2	1-flip	FCI	S^2	1-flip	FCI	S^2	1-flip
6.00	−1434.15	−1484.93	−1434.71	1562.92	1509.18	1564.05	2997.07	2994.11	2998.76
7.00	−602.08	−611.89	−602.71	624.05	615.34	624.87	1226.13	1227.23	1227.58
8.00	−230.99	−232.81	−231.39	234.31	233.31	234.76	465.30	466.12	466.15
9.00	−83.09	−83.48	−83.28	83.54	83.55	83.75	166.63	167.03	167.03
10.00	−28.39	−28.49	−28.47	28.44	28.50	28.53	56.83	56.99	56.99
11.00	−9.24	−9.27	−9.27	9.25	9.28	9.28	18.49	18.55	18.55
11.50	−5.18	−5.19	−5.19	5.18	5.19	5.19	10.35	10.39	10.39
12.00	−2.86	−2.87	−2.87	2.86	2.87	2.87	5.72	5.74	5.74
12.50	−1.56	−1.56	−1.56	1.55	1.56	1.56	3.11	3.12	3.12
13.00	−0.83	−0.84	−0.84	0.83	0.84	0.84	1.66	1.67	1.67
14.00	−0.23	−0.23	−0.23	0.23	0.23	0.23	0.45	0.45	0.45

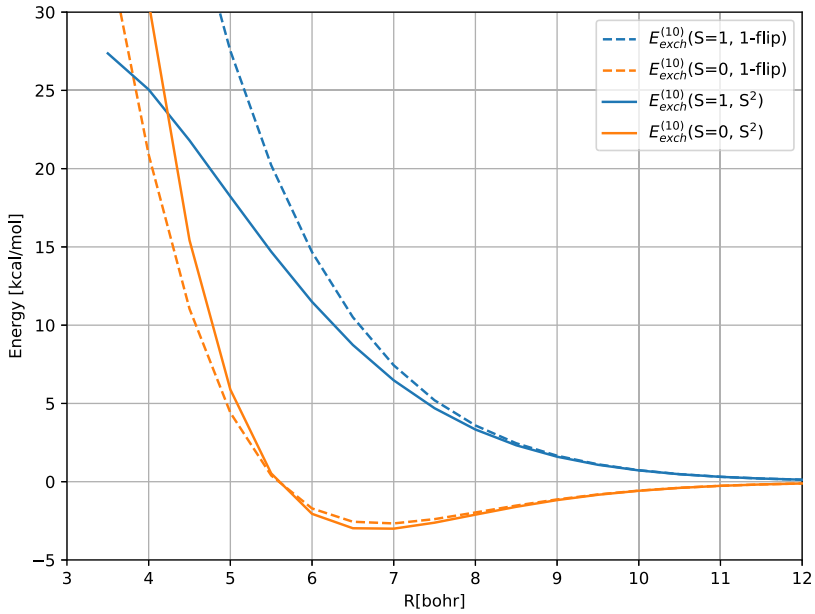


FIG. 1. Singlet and triplet $E^{(10)}_{\text{exch}}$ values (kcal/mol) in the lithium dimer as a function of interatomic distance (bohr) from SF-SAPT with and without the single-exchange approximation.

(excluded from the figure) and the S^2 results for the triplet also deviate slowly from the complete exchange results. At the chemical minimum of the triplet state (3.5 bohrs), the S^2 approximation diverges considerably from the complete results, recovering 92% of the exchange energy for the triplet, 88% for the quintet, and 58% of $\Delta E^{(10)}_{\text{exch}}$.

For the Li dimer, we also investigated the significance of the $E^{(10)}_{\text{exch}}\langle\mathcal{P}\rangle$ term, which can easily be incorporated into the S^2 approximation by solving Eq. (58) for $E^{(10)}_{\text{exch}}$. Even though the S^2 approximation does not become exact for this system when the $E^{(10)}_{\text{exch}}\langle\mathcal{P}\rangle$ term is taken into account (one is still missing the effects of double and triple electron exchanges), the inclusion of this term resulted in a much improved agreement with the complete exchange results. Therefore, the inclusion of $E^{(10)}_{\text{exch}}\langle\mathcal{P}\rangle$ while neglecting all other multiple exchanges was tested on a number of other small complexes. Unfortunately, such a treatment was found to drastically overcorrect for the effects missing in $E^{(10)}_{\text{exch}}(S^2)$ so that the resulting values were

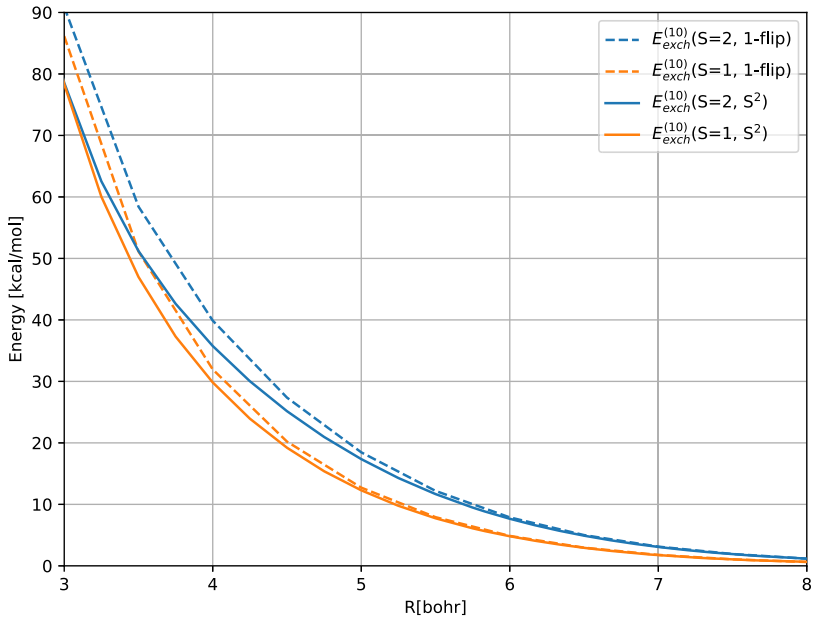


FIG. 2. Triplet and quintet $E^{(10)}_{\text{exch}}$ values (kcal/mol) of Li$\cdots$N as a function of interatomic distance (bohr) from SF-SAPT with and without the single-exchange approximation.

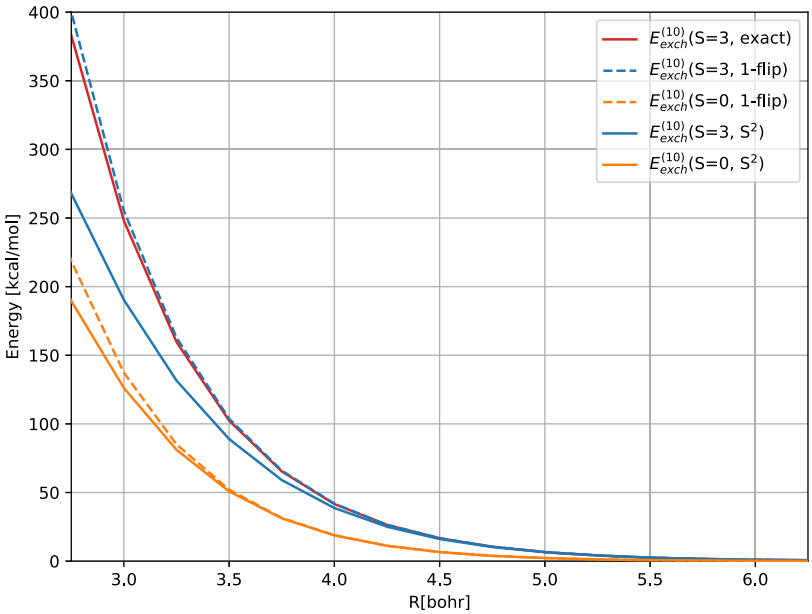


FIG. 3: $E^{(10)}_{exch}$ values (kcal/mol) of the N···N complex as a function of interatomic distance (bohr) from exact high-spin SAPT and SF-SAPT with and without the single-exchange approximation.

in worse agreement with the complete exchange results than the S^2 data. We suspect that the good performance of the $E^{(10)}_{exch}\langle\mathcal{P}\rangle$ inclusion for Li···Li is merely a consequence of each monomer having only one valence electron.

The N···N complex is the first of our test systems where the exact and single-spin-flip $E^{(10)}_{exch}$ values are distinct. Both of these results can be calculated for the high-spin state [the exact one using the standard high-spin SAPT(ROHF) implementation[4,6]]. This provides an opportunity to gauge the relative accuracy of the 1-flip and S^2 approximations, as shown in Fig. 3. The 1-flip treatment is a much milder approximation that slightly overestimates the exact $E^{(10)}_{exch}$ value, as opposed to the single-exchange approximation which underestimates it. Figure 3 also shows the low-spin exchange energies, where again the singlet state is less affected by the S^2 approximation than the highest spin state. In this way, the difference in the splittings predicted by the two approximations is primarily driven by the high-spin state error. The range presented in Fig. 3 falls between the septet van der Waals minimum of 7.2 bohrs and the singlet chemical minimum of 2.1 bohrs.

Table II provides the SF-SAPT $E^{(10)}_{exch}$ values for the manganese dimer. For the undecaplet exchange energies, the single-spin-flip results maintain perfect agreement with the exact high-spin $E^{(10)}_{exch}$ throughout nearly the entire range presented here, while the S^2 results begin to deviate at a relatively long-range distance of 9 bohrs. That said, the energy splittings from the two approximations are in close agreement due to the more even deviation of the S^2 values from the 1-flip ones for both the highest and lowest spin states of this system.

The last system that we consider in this section is the first bimolecular complex, the O_2 dimer. This system is interesting due to the large effect that geometry has on the splitting between the singlet and quintet states. Figure 4 shows the $\Delta E^{(10)}_{exch}$ values for the four representative geometries at several center-of-mass distances. The spin splitting is much larger in the L (linear) configuration and much smaller in the X configuration. For the H, T, and X structures, the 1-flip and S^2 splittings are virtually indistinguishable, while the two approximations slightly deviate from each other at short range for the L configuration. We report in the supplementary material that the 1-flip results show great agreement with the exact high-spin $E^{(10)}_{exch}$ values for all four geometries. The very good recovery of the singlet-quintet splitting by the S^2 approximation, shown in Fig. 4, stems from a cancellation of beyond-S^2 effects between the two spin states of the complex.

TABLE II. Singlet and undecaplet $E^{(10)}_{exch}$ values and $\Delta E^{(10)}_{exch}$ splittings (kcal/mol) for the manganese dimer as a function of interatomic distance (bohr). The S^2 and exact $E^{(10)}_{exch}$ values are different than in Ref. 7, as the ROHF convergence issues of the former calculations have been corrected.

	$E^{(10)}_{exch}$(S = 0)		$E^{(10)}_{exch}$(S = 5)			$\Delta E^{(10)}_{exch}$	
R	S^2	1-flip	S^2	1-flip	Exact	S^2	1-flip
4.50	115.32	157.35	117.51	159.52	159.51	2.19	2.17
5.00	79.10	99.44	79.89	100.23	100.23	0.79	0.79
5.50	52.96	62.39	53.25	62.69	62.69	0.29	0.30
6.00	34.62	38.84	34.73	38.95	38.95	0.11	0.11
6.50	22.16	23.98	22.20	24.02	24.02	0.04	0.04
7.00	13.92	14.68	13.94	14.70	14.70	0.02	0.02
7.50	8.60	8.91	8.61	8.92	8.92	0.01	0.01
8.00	5.25	5.37	5.25	5.37	5.37	0.00	0.00
9.00	1.88	1.90	1.88	1.90	1.90	0.00	0.00
10.00	0.65	0.65	0.65	0.65	0.65	0.00	0.00
11.00	0.22	0.22	0.22	0.22	0.22	0.00	0.00
12.00	0.07	0.07	0.07	0.07	0.07	0.00	0.00

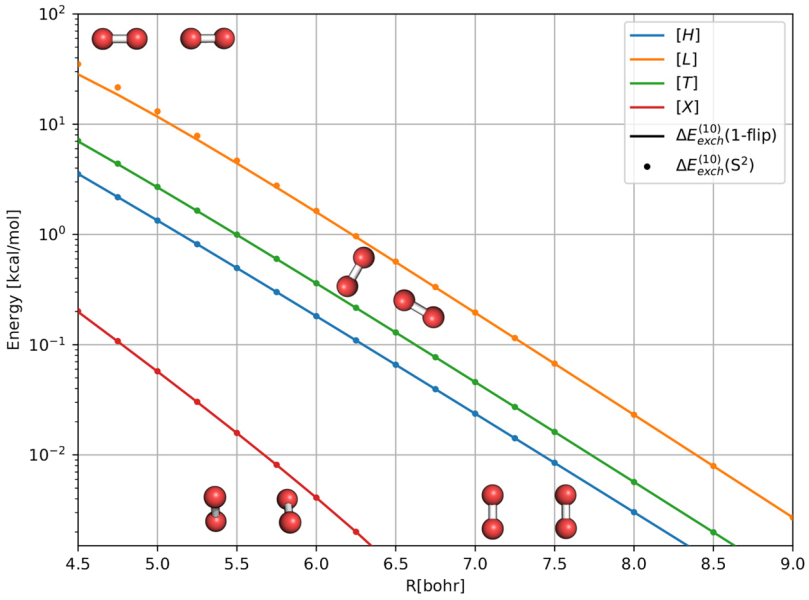


FIG. 4. The $\Delta E^{(10)}_{exch}$ approximation to the singlet-quintet splitting (kcal/mol) for different geometries of the O_2 dimer (as illustrated) as a function of the center-of-mass distance (bohr) from SF-SAPT with the single-exchange approximation and the single-spin-flip approximation.

B. Pancake bonded systems

An interesting possible application of SF-SAPT, already initiated in Ref. 7, is the spin-state splittings of pancake bonded dimers.[17] These systems are composed of radicals with highly delocalized singly occupied orbitals which interact in a fashion that is intermediate between covalent and noncovalent bonding. The pancake bonded singlet minimum is separated by a relatively small gap from the van der Waals-bonded triplet state. SF-SAPT is one of the simplest approaches to investigate these splittings, and the single-spin-flip approximation is exact for this doublet-doublet interaction. We have selected a number of pancake-bonded complexes to test our method, illustrating the favorable performance of our AO implementation with density-fitted generalized Coulomb and exchange matrices [Eq. (50)] from PSI4.[16] Figure 5 shows the monomers selected for our investigation. These include the phenalenyl (PLY) radical,[32] one of the prototypical examples of pancake bonding systems, as well as four of its derivatives[33] and the larger trioxotriangulene (TOT) radical.[34] The orientations of these systems in the homodimer prefer a maximum overlap of the delocalized singly occupied orbitals, i.e., the monomers stack directly on top of one another with their atoms lining up. With the central carbons of the monomers aligned, these dimers can have a staggered or eclipsed conformation as illustrated in Fig. 5. Both conformations of the TOT dimer are considered, but only the staggered conformation is used for PLY and its derivatives.

The spin state splitting in the PLY dimer was analyzed previously within the S^2 formulation of first-order SF-SAPT, together with the (much more demanding) supermolecular complete active space self consistent field theory (CASSCF).[7] Figure 6 presents these results along with the new nonapproximated $\Delta E^{(10)}_{exch}$ ones, the multireference averaged quadratic coupled cluster (MR-AQCC)[35] benchmark from Ref. 32, and the density functional theory (DFT) results obtained with the M05-2X functional.[36,37] Similar to the manganese dimer case, the splittings produced by the two versions of SF-SAPT do not differ greatly within this range. Some deviation can be observed at the shortest ranges in Fig. 6, but it is not drastic compared to the differences between various methods presented. The effect of the S^2 approximation appears to be relatively more consistent between the two spin states than observed for smaller systems.

The results for the homodimers of the PLY derivatives, shown in Fig. 7, further indicate that the S^2 and 1-flip $\Delta E^{(10)}_{exch}$ values agree well even at shorter distances. This observation would seem to support the idea that the removal of the S^2 approximation is less important to the improvement of the SF-SAPT splittings than the inclusion of higher-order exchange terms (which is in progress in our group). For comparison, we look at the M05-2X splitting values from Ref. 33. This functional was chosen as the most suitable for pancake bonded systems based on previous benchmarking.[37] The tri(tert-butyl)phenalenyl (TBPLY) dimer has the smallest splitting value at the minimum, where the SF-SAPT values for both approximations agree with each other and represent about two-thirds of the DFT result. For the other derivatives, the SF-SAPT results are about half the magnitude of the DFT results at the minimum. Also, for the triaminophenalenyl (TAPLY), trifluorophenalenyl (TFPLY), and trimethylphenalenyl (TMPLY) dimers, SF-SAPT shows a different ordering of the splitting values. The DFT results show the order of the splittings as TAPLY > TFPLY > TMPLY, while both SF-SAPT methods provide the order TMPLY > TAPLY > TFPLY for the M05-2X-optimized structures.

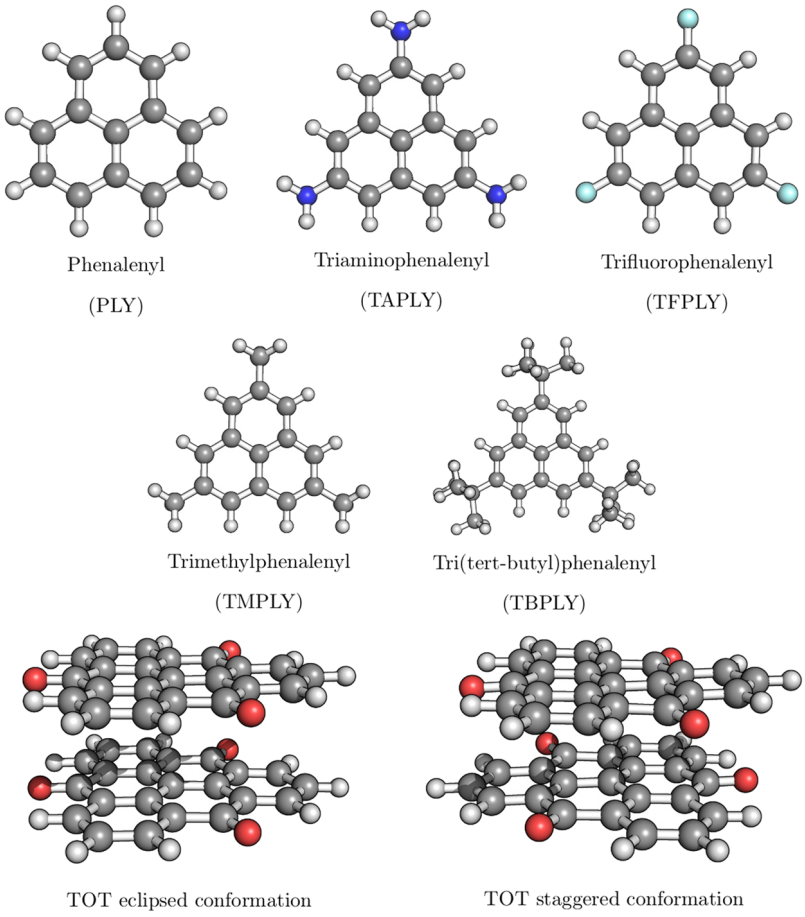


FIG. 5. Pancake-bonded radicals considered in this work. The staggered geometry was used for all dimers, and the eclipsed geometry was also considered for the trioxotriangulene (TOT) dimer.

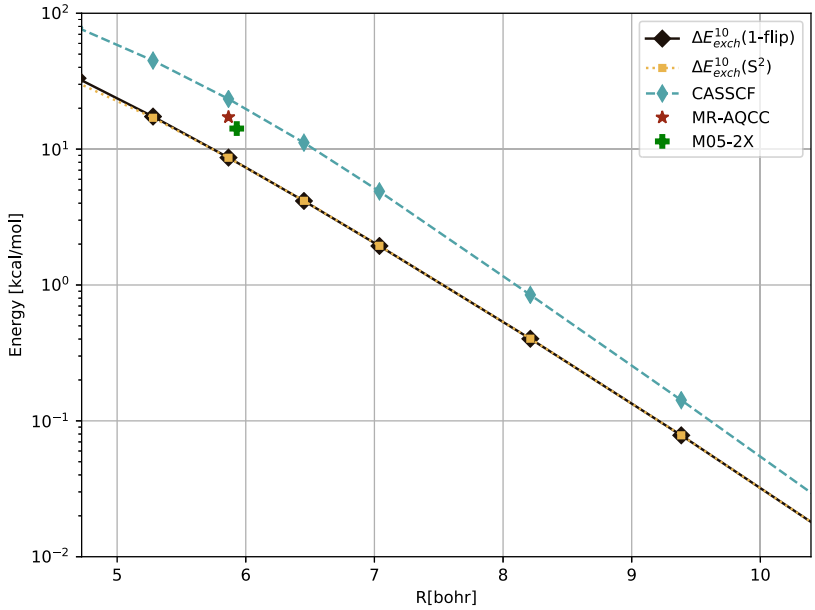


FIG. 6. Comparison of the $\Delta E^{(10)}_{exch}$ singlet-triplet splitting values (kcal/mol) for the staggered phenalenyl dimer as a function of center-of-mass distance (bohr) with literature results.[7,32,37]

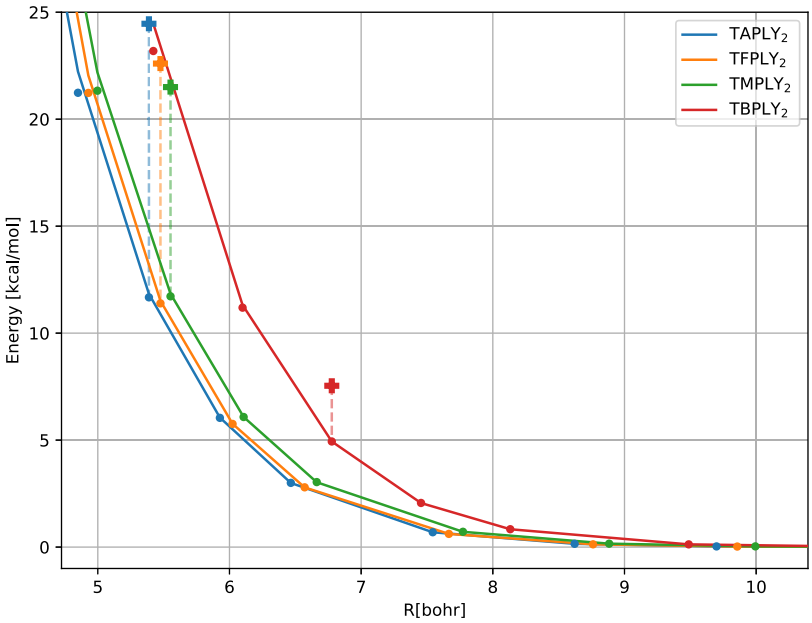


FIG. 7. Singlet-triplet splitting $\Delta E_{exch}^{(10)}$ estimates (kcal/mol) for the phenalenyl derivative dimers as functions of interplanar distance (bohr) compared to the literature M05-2X results.[33] The lines and points are, respectively, the SF-SAPT results with and without the single-exchange approximation. The plus marks are the M05-2X data.

Finally, we look at the trioxotriangulene dimer in its staggered and eclipsed conformations in Table III. The SF-SAPT methods agree with the DFT results on the ordering of the splittings of the two conformations, showing a larger splitting for the staggered one. Again, the SF-SAPT values are between half and two-thirds of the M05-2X values.[34] The S^2 results for both $\Delta E_{exch}^{(10)}$ and the individual $E_{exch}^{(10)}$ values are quite accurate in the presented range for this system compared to the nonapproximate 1-flip data.

TABLE III. Singlet and triplet $E_{exch}^{(10)}$ values and $\Delta E_{exch}^{(10)}$ splittings (kcal/mol) for the eclipsed and staggered geometries of the trioxotriangulene dimer as a function of interatomic distance expressed in units of the M05-2X optimized minimum separation ($R_0^{Eclipsed}$ = 6.659 bohrs, $E_{splitting}^{M05-2X,Eclipsed}$ = 3.33 kcal/mol, $R_0^{Staggered}$ = 5.962 bohrs, $E_{splitting}^{M05-2X,Staggered}$ = 10.67 kcal/mol).[34]

	$E_{exch}^{(10)}(S=0)$		$E_{exch}^{(10)}(S=1)$		$\Delta E_{exch}^{(10)}$	
R/R_0	S^2	1-flip	S^2	1-flip	S^2	1-flip
			Eclipsed			
0.80	225.43	228.41	238.31	241.22	12.88	12.81
0.90	79.48	79.74	84.67	84.89	5.19	5.15
1.00	27.11	27.11	29.05	29.05	1.94	1.93
1.10	9.00	9.00	9.70	9.69	0.70	0.69
1.20	2.92	2.92	3.16	3.16	0.24	0.24
1.40	0.28	0.28	0.31	0.31	0.03	0.03
1.60	0.02	0.02	0.02	0.02	0.00	0.00
			Staggered			
0.80	417.36	426.73	443.09	453.31	25.74	26.58
0.90	167.05	168.34	179.42	180.70	12.37	12.35
1.00	64.48	64.60	69.86	69.94	5.38	5.34
1.10	24.17	24.17	26.36	26.35	2.19	2.18
1.20	8.85	8.84	9.70	9.69	0.85	0.85
1.40	1.10	1.10	1.22	1.22	0.12	0.12
1.60	0.12	0.12	0.13	0.13	0.01	0.01

V. SUMMARY

We derived a new formula for the first-order spin-flip SAPT exchange energy,[7] valid for an arbitrary spin state of a weakly interacting complex, replacing the conventional single-exchange (S^2) approximation by the much milder single-spin-flip approximation. The new SF-SAPT correction, with the noninteracting monomers described by their ROHF determinants, involves terms similar to those appearing in the expressions for the complete (non-S^2) first-order exchange and second-order exchange-dispersion energies.[12,14] In this way, the spin flips are treated as a subset of the double excitations found in second-order dispersion and exchange-dispersion corrections. The resulting equations were implemented in both their molecular orbital and atomic orbital forms, where the latter allows for the application of this method to much larger systems.

The newly enhanced first-order SF-SAPT approach was applied to the same selection of diatomic and small molecular test systems as in Ref. 7. The S^2 approximation is not exact even for the Li···H system where only a single electron exchange is possible, as this approximation neglects a term that is a product of two single exchanges. This approximation is also particularly poor for the lithium dimer, leading to an unphysical crossing of the singlet and triplet curves at short range. In contrast, the single-spin-flip approximation was demonstrated to be much milder than

the S^2 one. The 1-flip treatment is formally exact for any dimer with at least one doublet monomer and deviates from the exact $E^{(10)}_{exch}$ much slower than the S^2 variant in other cases. Another observation made possible by the new development is that the S^2 approximation is generally better for the low-spin states of small complexes than for the high-spin state. As the size of the system increases, at least for the complexes considered here the effect of the S^2 approximation on the high and low-spin states becomes comparable.

The new $E^{(10)}_{exch}$ formulation was further applied to the determination of the singlet-triplet splittings for the phenalenyl radical dimer and other pancake bonded systems. These calculations are feasible, thanks to the recasting of the MO formulas into their AO form, which allows us to take advantage of PSI4's efficient tools for producing density-fitted generalized Coulomb and exchange matrices.[16] The resulting implementation is only somewhat more expensive than the $E^{(10)}_{exch}(S^2)$ one of Ref. 7, exhibiting the same N^4 scaling with the basis set size with a somewhat larger prefactor. In fact, the computation time is dominated by the construction of the Coulomb and exchange matrices, and the number of such matrices is 8 for the S^2 correction and 11 for the 1-flip one. Despite the formal exactness of the 1-flip treatment, it does not provide especially different first-order splittings compared to the S^2 approximation. This is due to a more even effect of the S^2 approximation on the singlet and triplet states of these systems. The SF-SAPT splittings were compared to a selection of literature values, in particular, the M05-2X results from Refs. 33, 34, and 37. In comparison to the DFT results, the first-order SF-SAPT treatment underestimates the splittings and shows a different ordering of the PLY derivatives. While these splittings are generally difficult to calculate and the literature results are far from benchmark quality, we cannot expect a high quantitative accuracy from a simple first-order perturbation theory that also neglects intramolecular electron correlation. The work to extend the new SF-SAPT formalism to arbitrary-spin second-order exchange induction and exchange dispersion energies is in progress in our group.

SUPPLEMENTARY MATERIAL

See the supplementary material for the numerical data presented in all figures and the Cartesian coordinates for the complexes considered in this work.

ACKNOWLEDGMENTS

This work was supported by the U.S. National Science Foundation CAREER Award No. CHE-1351978. We thank Bogumił Jeziorski and Piotr Żuchowski for stimulating discussions.

REFERENCES

[1] B. Jeziorski, R. Moszyński, and K. Szalewicz, Chem. Rev. **94**, 1887 (1994).
[2] M. Bartolomei, E. Carmona-Novillo, M. I. Hernández, J. Campos-Martínez, and R. Hernández-Lamoneda, J. Chem. Phys. **128**, 214304 (2008).
[3] M. Bartolomei, E. Carmona-Novillo, M. I. Hernández, J. Campos-Martínez, and R. Hernández-Lamoneda, J. Chem. Phys. **133**, 124311 (2010).
[4] P. S. Żuchowski, R. Podeszwa, R. Moszyński, B. Jeziorski, and K. Szalewicz, J. Chem. Phys. **129**, 084101 (2008).
[5] M. Hapka, P. S. Żuchowski, M. M. Szczęśniak, and G. Chałasiński, J. Chem. Phys. **137**, 164104 (2012).
[6] J. F. Gonthier and C. D. Sherrill, J. Chem. Phys. **145**, 134106 (2016).
[7] K. Patkowski, P. S. Żuchowski, and D. G. A. Smith, J. Chem. Phys. **148**, 164110 (2018).
[8] A. I. Krylov, Chem. Phys. Lett. **338**, 375 (2001).
[9] A. I. Krylov, Annu. Rev. Phys. Chem. **59**, 433 (2008).
[10] P. M. Zimmerman, F. Bell, M. Goldey, A. T. Bell, and M. Head-Gordon, J. Chem. Phys. **137**, 164110 (2012).
[11] N. J. Mayhall and M. Head-Gordon, J. Phys. Chem. Lett. **6**, 1982 (2015).
[12] B. Jeziorski, M. Bulski, and L. Piela, Int. J. Quantum Chem. **10**, 281 (1976).
[13] R. Schäffer and G. Jansen, Theor. Chem. Acc. **131**, 1235 (2012).
[14] R. Schäffer and G. Jansen, Mol. Phys. **111**, 2570 (2013).
[15] D. G. A. Smith, L. A. Burns, D. A. Sirianni, D. R. Nascimento, A. Kumar, A. M. James, J. B. Schriber, T. Zhang, B. Zhang, A. S. Abbott *et al.*, J. Chem. Theory Comput. **14**, 3504 (2018).
[16] R. M. Parrish, L. A. Burns, D. G. A. Smith, A. C. Simmonett, A. E. DePrince III, E. G. Hohenstein, U. Bozkaya, A. Y. Sokolov, R. Di Remigio, R. M. Richard *et al.*, J. Chem. Theory Comput. **13**, 3185 (2017).
[17] M. Kertesz, Chem. Eur. J. **25**, 400 (2019).
[18] B. Jeziorski, G. Chałasiński, and K. Szalewicz, Int. J. Quantum Chem. **14**, 271 (1978).
[19] R. Vein and P. Dale, *Determinants and Their Applications in Mathematical Physics* (Springer, New York, 1999).
[20] K. U. Lao and J. M. Herbert, J. Chem. Theory Comput. **14**, 2955 (2018).
[21] R. M. Parrish, K. C. Thompson, and T. J. Martínez, J. Chem. Theory Comput. **14**, 1737 (2018).
[22] T. H. Dunning, Jr., J. Chem. Phys. **90**, 1007 (1989).
[23] R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, J. Chem. Phys. **96**, 6796 (1992).
[24] N. B. Balabanov and K. A. Peterson, J. Chem. Phys. **123**, 064107 (2005).
[25] P.-O. Widmark, P.-Å. Malmqvist, and B. O. Roos, Theor. Chem. Acc. **77**, 291 (1990).
[26] M. Bartolomei, M. I. Hernández, J. Campos-Martínez, E. Carmona-Novillo, and R. Hernández-Lamoneda, Phys. Chem. Chem. Phys. **10**, 5374 (2008).
[27] K. Patkowski, T. Korona, and B. Jeziorski, J. Chem. Phys. **115**, 1137 (2001).
[28] F. Weigend, J. Comput. Chem. **29**, 167 (2007).
[29] F. Weigend, A. Köhn, and C. Hättig, J. Chem. Phys. **116**, 3175 (2002).
[30] H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, and M. Schütz, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **2**, 242 (2012).
[31] R. Moszyński, B. Jeziorski, S. Rybak, K. Szalewicz, and H. L. Williams, J. Chem. Phys. **100**, 5080 (1994).
[32] Z. Cui, H. Lischka, H. Z. Beneberu, and M. Kertesz, J. Am. Chem. Soc. **136**, 5539 (2014).
[33] Z. Mou, T. Kubo, and M. Kertesz, Chem. Eur. J. **21**, 18230 (2015).
[34] Z. Mou and M. Kertesz, Angew. Chem., Int. Ed. **56**, 10188 (2017).
[35] P. G. Szalay and R. J. Bartlett, Chem. Phys. Lett. **214**, 481 (1993).
[36] Y. Zhao and D. G. Truhlar, J. Phys. Chem. A **110**, 5121 (2006).
[37] Z. Mou, Y.-H. Tian, and M. Kertesz, Phys. Chem. Chem. Phys. **19**, 24761 (2017).