

The Improvement of Intramolecular Symmetry-Adapted Perturbation Theory

by

Du Luu

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
December 13, 2025

Keywords: intramolecular interactions, molecular fragmentation, range separation

Copyright 2025 by Du Luu

Approved by

Konrad Patkowski, Chair, Professor of Chemistry and Biochemistry
Vincent Ortiz, Professor Emeritus of Chemistry and Biochemistry
Evangelos Miliordos, James E. Land Associate Professor of Chemistry and Biochemistry
Rashad Karimov, S.D. and Karen H. Worley Associate Professor of Chemistry and
Biochemistry
Dmitry Glotov, Associate Professor, Department of Mathematics and Statistics

Abstract

Symmetry-adapted perturbation theory (SAPT) is a popular and versatile tool to compute and decompose noncovalent interaction energies between molecules. The intramolecular SAPT (ISAPT) variant provides a similar energy decomposition between two nonbonded fragments of the same molecule, covalently connected by a third fragment. The presence of artificial dipole moments at the interfragment boundary, as the atoms of A and B directly connected to C are missing electrons on one of their hybrid orbitals, displays several issues for many fragmentation patterns (that is, specific assignments of atoms to the A/B/C subsystems), including an artificially repulsive electrostatic energy (even when the fragments are hydrogen-bonded) and very large and mutually cancelling induction and exchange-induction terms. The improved ISAPT(SIAO1) approach reassigns one electron on a singly occupied link hybrid orbital from C to each of A/B, providing reasonable values of all ISAPT corrections for all fragmentation patterns, and a fast and systematic basis set convergence. An alternative approach to study the SAPT/ISAPT interaction energy is established by singling out the noncovalent interaction using the long-ranged part of the Coulomb potential based on the Gaussian or error-function range separation. The long-ranged Coulomb potential is evaluated with either the entire intermolecular/interfragment interaction or only its attractive terms. The energy corrections from range-separated SAPT/ISAPT are in reasonable agreement with complete SAPT/ISAPT data. The best consistency is attained for the error-function separation applied to all interaction terms, both attractive and repulsive. The range separation appears to be a potentially promising technique towards a fragmentation-free decomposition of intramolecular nonbonded energy.

Acknowledgments

The graduate program has been a meaningful and important milestone in my academic journey. I joined the computational research group with great excitement and deep curiosity about both the theory and practice of quantum chemistry. For this, I have several people to thank.

First and foremost, I would like to express my deepest gratitude to my research advisor, Professor Konrad Patkowski. His interesting guidance and tireless dedication have taught me to be more resilient, to think critically, and to work efficiently. His sharp insight has encouraged me to think outside the box. My research accomplishments would not have been possible without his support. Professor Patkowski has made my graduate experience an unforgettable and enriching academic journey.

I would also like to thank all the professors whose courses I took during my graduate studies. I am sincerely grateful for the opportunity to learn so many new and fascinating topics in quantum chemistry. Their teachings provided numerous insights that have greatly contributed to my research.

I am deeply thankful to my elder brother, Trieu Luu, for his unwavering support—financially, mentally, and emotionally. I consider myself incredibly fortunate to have someone who has helped shape my personal and professional growth. He has always cared for me in countless ways and taught me the true value of work-life balance. Through thick and thin, I can always count on him to brighten my day.

Most importantly, I would like to thank my parents, Minh Luu and Ngoc Lam. I would not be where I am today without their support, sacrifices, and unconditional love. Words cannot fully express how grateful I am for everything they have done for me.

Finally, I gratefully acknowledge that this research was supported by funding from the National Science Foundation (NSF).

Table of Contents

Abstract	ii
Acknowledgments	iii
1 Introduction - Motivation	1
2 Introduction - Methods	2
2.1 Hartree-Fock theory	3
2.2 Symmetry-Adapted Perturbation Theory	4
2.3 Intramolecular Symmetry-Adapted Perturbation Theory	8
2.3.1 Splitting of Intrinsic Atomic Orbitals	10
2.3.2 Range Separation of the Interaction Potential in SAPT and ISAPT methods	11
2.4 Intrinsic Bond Orbital and Intrinsic Atomic Orbital Localizations	13
2.5 McMurchie-Davidson recursion for Gaussian integrals	14
3 Overcoming Artificial Multipoles in Intramolecular Symmetry-Adapted Perturbation Theory	16
4 Range Separation of the Interaction Potential in Intermolecular and Intramolecular Symmetry-Adapted Perturbation Theory	42
5 The Application of Intramolecular-SAPT to Z/E-Azobenzene Isomerization	67
6 Summary and Outlook	77
References	80

List of Figures

3.1	ISAPT energy components in 2,4-pentanediol (P242 fragmentation pattern) computed using the aXZ bases (X=D, T, Q) and three link assignments: (original) C, SAO1, and SIAO1.	18
3.2	Comparison of the ISAPT0 energy components between n-heptane (C73, colored bars) and its branched isomer, 2,4-dimethylpentane (C7B24, empty bars) computed with the (original) C, SAO1, and SIAO1 variants using the aXZ bases (X=D, T, Q).	19
3.3	ISAPT exchange energy $E_{\text{exch}}^{(10)}$ from parallel (PAR), perpendicular (PERP), and average (AVG) spin couplings in models P244, P142, C73, and C7B24. The values are computed with SAO1 (empty bars) and SIAO1 (colored bars) in the aXZ bases, X=D,T,Q.	19
4.1	The SAPT0(RSEP/erf) interaction energy components in the ammonia-chlorine monofluoride complex computed with various values of ω (bohr ⁻¹) in the juncc-pVDZ basis set. The intermolecular center-of-mass distance R is given relative to the global minimum value $R_{\text{min}} = 5.569$ bohr.	44
4.2	The SAPT0(RSEP-eN/erf) interaction energy components in the ammonia-chlorine monofluoride complex computed with various values of ω (bohr ⁻¹) in the juncc-pVDZ basis set. The intermolecular center-of-mass distance R is given relative to the global minimum value $R_{\text{min}} = 5.569$ bohr.	45
4.3	The components of the range-separated ISAPT/SIAO1 interaction energy of the 1,4-pentanediol molecule (fragmented as pictured) computed with different values of η (bohr ⁻² , colored bars) and ω (bohr ⁻¹ , empty bars) in the aDZ basis set. The energy results are compared to those of the traditional ISAPT/SIAO1 method.	46
5.1	The geometric structures of the cis- all-meta alkyl substituted azobenzene molecules. The alkyl groups include methyl (met), ethyl (eth), isopropyl (ipr), and tert-butyl (tbu).	71
5.2	The geometric structures of the trans- all-meta alkyl substituted azobenzene molecules. The alkyl groups include methyl (met), ethyl (eth), isopropyl (ipr), and tert-butyl (tbu).	72

List of Tables

3.1	Length of HF dipole moments (in a.u.) of fragments <i>A</i> and <i>B</i> computed with the ISAPT variants of (original) C, SAO1, and SIAO1 on bases aDZ, aTZ, and aQZ.	18
5.1	The ISAPT/SIAO1 decomposition energies (in kcal/mol) of 10 models of the all-meta alkyl substituted azobenzene molecule. The alkyl substituents include Me, Et, iPr, and tBu. The values of δHF are taken from the ISAPT/C linker assignment calculations for the total net energy. [39] The calculations are computed using the aug-cc-pVDZ basis set.	73
5.2	The (original) ISAPT/C decomposition energies (in kcal/mol) of 10 models of the all-meta alkyl substituted azobenzene molecule. The alkyl substituents include Me, Et, iPr, and tBu. The calculations are computed using the aug-cc-pVDZ basis set.	74
5.3	The DLPNO-CCSD(T)/aug-cc-pVDZ rotational energy (in kcal/mol) for the H-substituted intermediate geometries of the cis and trans azobenzene isomers. The electron correlation energy $E(\text{L-MP2})$ extracted from the DLPNO-CCSD(T) calculation serves for the reference to the MP2 correlation energy.	75
5.4	The relative isomerization energy (in kcal/mol) of 10 azobenzene models evaluated by the sum of the ISAPT/SIAO1 conformational energy and the predicted DLPNO-CCSD(T)/aug-cc-pVDZ deformational energy in the intermediate systems of X-Ph-N=N-H and X-Ph-H. The calculations are computed in the aug-cc-pVDZ basis set.	76

Chapter 1

Introduction - Motivation

The nature of weak noncovalent interactions between separate molecules and between non-bonded fragments of the same molecule is critical for applications in many areas of chemistry, physics, materials science, and biology. In the exploration of noncovalent interactions, Symmetry-Adapted Perturbation Theory (SAPT) [30, 32, 65, 66] has emerged as one of the powerful frameworks that satisfies both qualitative and quantitative insight into the energy contributions. The SAPT method decomposes the total interaction energy into fundamental contributions such as electrostatics, exchange, induction, and dispersion. While SAPT is typically applied to interactions between distinct molecules, its extension to intramolecular interactions, known as Intramolecular SAPT (ISAPT) [49], offers a rigorous, first-principles approach for analyzing noncovalent forces between fragments in a single molecular entity. The computational study of intramolecular interactions is important for accurately describing conformational preferences, intramolecular hydrogen bonding, and the folding behavior of biomolecules. Although the ISAPT methodology is capable of providing physical insight into the interaction energy, some of the predicted intramolecular energy contributions are unphysical in some fragmentation models. [49] In addition, the reliance on fragmentation in the ISAPT method gives rise to the challenge of quantifying intramolecular interactions in an entire molecule. To enhance the performance of the ISAPT method, the issue that causes the unphysical energy values has been resolved [39], and the range separation functions serve as a potential computational tool to program an ISAPT algorithm without requiring fragmentation. [40]

Chapter 2

Introduction - Methods

In quantum chemistry, the electronic structure theory describes the interaction and behavior of electrons to predict the material properties and to understand the nature of chemical bonding. [25, 26, 54, 58] For the atomic system of a single electron, the exact ground state energy can be computed by solving the static time-independent Schrodinger equation. The challenge comes with solving the exact energy of the many-electrons system because of the quantity and motion of the nuclei and electrons.

$$\hat{H}\Psi = E\Psi \quad (2.1)$$

$$\hat{H} = \frac{-\hbar^2}{2} \sum_{A=1}^N \frac{\nabla_A^2}{M_A} - \frac{\hbar^2}{2m_e} \sum_{i=1}^n \nabla_i^2 + \frac{e^2}{4\pi\epsilon_0} \sum_{A<B}^N \frac{Z_A Z_B}{r_{AB}} + \frac{e^2}{4\pi\epsilon_0} \sum_{i<j}^n \frac{1}{r_{ij}} - \frac{e^2}{4\pi\epsilon_0} \sum_A^N \sum_i^n \frac{Z_A}{r_{Ai}} \quad (2.2)$$

where the Hamiltonian operator $\hat{H}$ represents the total energy E of a system, including the potential and kinetic energies of all particles, described by a wavefunction Ψ . The first two terms of the Hamiltonian operator represent the kinetic energy of the N nuclei and n electrons, where M_A and m_e are the mass of the nucleus A and the electron respectively. The next two terms are the potential energies that describe the repulsion of two nuclei A, B at distance r_{AB} , where Z_A and Z_B are the corresponding nuclear charge of A and B , and of two electrons ij at distance r_{ij} . The final term is the electrostatic potential energy of the nuclei and electrons at distance r_{Ai} .

To simplify the Hamiltonian operator, the Born-Oppenheimer approximation is applied to interpret the molecular dynamic. The approximation neglects the motion of the nuclei based

on the fact that the nuclei are much more massive than the electrons, generating the electronic-Hamiltonian with respect to the motion and interaction between the electrons and the static nuclei in a many-body system.

2.1 Hartree-Fock theory

The Hartree-Fock (HF) method is a simplest approximate theory and a central starting point to describe the N-body clamped-nuclei Schrodinger equation. The HF method determines a set of spin orbitals, presented in the wavefunction Ψ , where all electrons can be best approximated by a single Slater-determinant. [16, 17, 23, 61, 62]The following matrix represents the Slater-determinant of N-electron $x_1 \dots x_N$ constructed by the spin orbitals $\chi_1 \dots \chi_N$:

$$\Psi = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(x_1) & \chi_2(x_1) & \dots & \chi_N(x_1) \\ \chi_1(x_2) & \chi_2(x_2) & \dots & \chi_N(x_2) \\ \vdots & & & \\ \chi_1(x_N) & \chi_2(x_N) & \dots & \chi_N(x_N) \end{vmatrix} \quad (2.3)$$

The Slater-determinant satisfies the anti symmetry of the wavefunction, as a consequence to the Pauli-exclusive principle for fermions, by changing sign in the exchange of two electrons.

$$\Psi(x_1, x_2) = -\Psi(x_2, x_1) \quad (2.4)$$

The HF method assumes that each electron interacts with the mean field created by all other electrons in a closed-shell system and provides a mean of approximating wavefunction, occupied orbitals, and their energy eigenvalues. The electronic-Hamiltonian can be rewritten as a combination of one-electron operator $\hat{h}$ and the potential energy operator.

$$\hat{h} = \frac{-1}{2} \nabla_i^2 - \sum_A \frac{Z_A}{r_{iA}} \quad (2.5)$$

$$v(i, j) = \frac{1}{r_{ij}} \quad (2.6)$$

where the $\hat{h}$ operator involves the kinetic energy of the electron i and the interaction of electron i with each nucleus. The potential operator accounts for the interaction of a pair of electrons ij .

In order to make the HF equation numerically feasible, Roothaan developed a matrix version using linear algebra to effectively calculate the HF orbitals and their eigen-energies. [55]

$$FC = SC\epsilon \quad (2.7)$$

where F is the Fock matrix, C is the matrix of coefficients of molecular orbitals, S is the overlap matrix of atomic orbitals basis, and ϵ is the diagonal matrix of orbital energies.

The basis set is introduced to produce a self-consistent field (SCF) of the Fock matrix. Molecular orbitals are expressed as a linear combination of a set of basis functions $\tilde{\chi}$ with some coefficients C .

$$\tilde{\Phi}_i = \sum_{\mu} C_{\mu i} \tilde{\chi}_{\mu} \quad (2.8)$$

The Fock matrix depends on the generated orbitals and the SCF cycles minimize the total energy within a given basis function. The SCF protocol is performed by iterations until the smallest value of energy is achieved, when it converges with the proposed threshold value.

2.2 Symmetry-Adapted Perturbation Theory

In the many-body approximations, Symmetry-Adapted Perturbation Theory (SAPT) is one of the common methodology to describe the non-covalent intermolecular interaction using the perturbation technique. [19, 30, 32, 51, 65, 66] The SAPT0 is the base level of SAPT in the

perturbational series and is utilized in the thesis. [28, 29, 48] Typically, the interaction energy between two isolated monomers can be calculated by a subtraction from the total energy of the dimer, namely the supermolecular method.

$$E_{int} = E_{AB} - E_A - E_B \quad (2.9)$$

where E_{AB} , E_A , and E_B are the approximate energies that can be computed separately using any method in computational chemistry.

The supermolecular method has to fulfill the error cancellation, particularly needs to be size consistent. For each term E_{AB} , E_A , and E_B , the approximation to the exact ground state energy has to be as small as possible, and E_{AB} approaches $E_A + E_B$ when the intersystem separation increases.

The formalism of SAPT works under the Rayleigh-Schrodinger (RS) expansion that directly computes the interaction energy. In addition, the valuable insight into the components of the intermolecular interaction energy, corresponding to the physical contribution of distinct forces, is described as the perturbative expressions.

In the RS perturbational method, the clamped-nuclei Hamiltonian operator of a complex AB is written as:

$$H = H_A + H_B + V \quad (2.10)$$

where H_A and H_B are treated as the unperturbed zeroth-order Hamiltonian of two isolated monomers A and B . V is the perturbational operator that represents the intermolecular interaction of dimer AB .

In cases of the N-body molecules, the calculation of the exact eigenfunction of the Hamiltonian is not practical. The simpler formulation of the monomer Hamiltonian is defined by the sum of the operators of the Fock matrix and the intramonomer electron correlation as

$H_X = F_X + W_X$, where $X = A, B$. Consequently, in SAPT methodology, the clamped-nuclei Hamiltonian operator is expressed as below, where λ and ζ are the perturbation expansion coefficients of operators V and W respectively.

$$H = F_A + F_B + \lambda V + \zeta(W_A + W_B) \quad (2.11)$$

At the level of SAPT0, the energy corrections are terminated at the second order of the RS perturbation theory as the following expression:

$$E_{RS}^{(1)} = \langle \Psi_A^{(0)} \Psi_B^{(0)} | V | \Psi_A^{(0)} \Psi_B^{(0)} \rangle \quad (2.12)$$

$$E_{RS}^{(2)} = \sum_{i \neq 0} \frac{|\langle \Psi_A^{(0)} \Psi_B^{(0)} | V | \Psi_A^{(i)} \Psi_B^{(0)} \rangle|^2}{E_A^{(0)} - E_A^{(i)}} + \sum_{j \neq 0} \frac{|\langle \Psi_A^{(0)} \Psi_B^{(0)} | V | \Psi_A^{(0)} \Psi_B^{(j)} \rangle|^2}{E_B^{(0)} - E_B^{(j)}} + \sum_{i,j \neq 0} \frac{|\langle \Psi_A^{(0)} \Psi_B^{(0)} | V | \Psi_A^{(i)} \Psi_B^{(j)} \rangle|^2}{E_A^{(0)} + E_B^{(0)} - E_A^{(i)} - E_B^{(j)}} \quad (2.13)$$

where $\Psi_A^{(0)} \cdot \Psi_B^{(0)}$ are the product of the unperturbed wavefunction of monomers A and B . $E_A^{(0)}$ and $E_B^{(0)}$ are the corresponding zeroth-order energies. $\Psi_A^{(i)}$ and $\Psi_B^{(j)}$ are the excited eigenfunctions of monomers in the polarization expansion. $E_A^{(i)}$ and $E_B^{(j)}$ are the excited eigenvalues of operators H_A and H_B respectively.

In the family of the energy corrections in SAPT0, the $E_{RS}^{(10)}$ represents the first-order electrostatic energy $E_{elst}^{(10)}$. The first two terms of the $E_{RS}^{(20)}$ describe the second-order induction energy $E_{ind,B \rightarrow A}^{(20)}$ and $E_{ind,A \rightarrow B}^{(20)}$, and the last term of the $E_{RS}^{(20)}$ accounts for the second-order dispersion energy $E_{disp}^{(20)}$. The exchange corrections are generated by performing the symmetry adaptation onto the operator of the perturbation interaction $\hat{V}$. The antisymmetrizer operator $\hat{A}$ enforces the product of the unperturbed wavefunction for monomers A and B to become antisymmetric with respect to electron exchange between A and B .

$$\hat{\mathcal{A}} = \frac{N_A!N_B!}{(N_A + N_B)!} \hat{\mathcal{A}}_{\mathcal{N}_A} \hat{\mathcal{A}}_{\mathcal{N}_B} (1 + \hat{\mathcal{P}}) \quad (2.14)$$

where the operator $\hat{\mathcal{P}} = \sum_{i \in A} \sum_{j \in B} \hat{\mathcal{P}}_{ij}$ collects all single transpositions of electrons between A and B .

In SAPT calculations, the above truncated antisymmetrizer operator $\hat{\mathcal{A}}$, namely the single exchange or S^2 -approximation, is more practical and less costly in the evaluation of the exchange corrections. An important exception comes from the formation of the first-order exchange energy $E_{exch}^{(10)}$ at the SAPT0 level of theory, where the non-approximated energy correction using the full $(N_A + N_B)$ -electron antisymmetrizer $\hat{\mathcal{A}}$ and the S^2 -approximation are introduced.

The following is the expression of the SAPT0 energy decomposition:

$$E_{int}^{SAPT0} = E_{elst}^{(10)} + E_{exch}^{(10)} + E_{ind,resp}^{(20)} + E_{exch-ind,resp}^{(20)} + E_{disp}^{(20)} + E_{exch-disp}^{(20)} + \delta_{HF}^{(2)} \quad (2.15)$$

In the denotation of the SAPT energy term, the digits in the superscript correspond to the orders of the perturbations of the intermolecular interaction $\hat{V}$ and the intramolecular electron correlation $\hat{W}$. In SAPT0, all energy corrections neglect the interaction of the later perturbation.

In a complex AB , the electrostatic energy describes the Coulombic attraction between the charge distribution of isolated monomers A and B . At the short range of the intermolecular separation, the electrostatic energy accounts for the effect of the charge penetration, resulting from the overlap of the electron densities of monomers. The exchange energy represents the repulsive force at the short-range interaction, following the Pauli-exclusive principle. The induction energy is the response of the polarization of one monomer due to the electric field

generated by the interacting monomer. The dispersion energy is the interaction of the simultaneous excitation, or the correlation of the fluctuation, of two electron densities. The exchange-counterparts of the induction and dispersion energies adjust the overpolarization at the short-range interaction. The $\delta E_{HF}^{(2)}$ term is an approximation for the third and higher-order induction and exchange-induction contributions, as the result from the supermolecular HF calculation:

$$\delta E_{HF}^{(2)} = E_{int}^{HF} - E_{elst}^{(10)} - E_{exch}^{(10)} - E_{ind,resp}^{(20)} - E_{exch-ind,resp}^{(20)} \quad (2.16)$$

2.3 Intramolecular Symmetry-Adapted Perturbation Theory

In order to evaluate intramolecular interactions, the available methodologies [21,49,50] rely on fragmentation. There are a few options to define and partition the interaction energy between fragments A and B in the presence of central linker C . The most straightforward method is to remove fragment C and cap the dangling bonds of fragments A and B with hydrogen atoms. The approximation of the interaction energy between fragments A and B can be evaluated using any available computational method, e.g. the standard SAPT variants to decompose the intermolecular interactions. [45,46] However, the hydrogen-replacement approach requires altering the molecular system and, if the linking fragment is small, artificial means to avoid an overly pronounced repulsion between the capping hydrogens. Gonthier et al. introduced two alternative fragmentations that preserve the geometric structure by localizing the occupied orbitals on fragments A , B , and C and defining the zeroth-order wavefunctions. [21,49,50]

The first method [21] involves constructing a non-Hermitian zeroth-order Hamiltonian in which the A - B interactions are removed following the Chemical Hamiltonian approach introduced by Mayer. [42] These removed interactions are then brought back as the perturbation that gives rise to the noncovalent intramolecular interaction. [50] The need for a non-Hermitian, biorthogonal formalism makes the theory somewhat more complicated, and the resulting energy decomposition the polarization corrections cannot be separated from their exchange counterparts (for example, the entire first-order energy is a single term), but the second-order energy

includes an explicit delocalization (charge transfer) term in addition to the induction and dispersion ones.

The second method is an extension of SAPT, known as the Intramolecular Symmetry-Adapted Perturbation Theory (ISAPT), which provides a simulation to compute the energy contribution of the intramolecular non-covalent interaction occurring within a molecule. [49] Similarly to the SAPT calculation, the ISAPT interaction energy comprises the contributions of first-order electrostatic and exchange, and second-order induction and dispersion. The main features of the ISAPT method rely on the molecular fragmentation, orbital localization, and wavefunction embedding. The goal is to determine the non-interacting zeroth-order HF wavefunctions of two fragments, in a molecule, that contain the functional groups. The computational framework of the SAPT0 methodology can then be adapted to analyze the interaction energy between those two fragments.

A crucial criteria of the ISAPT calculation is that the molecule of interest can be partitioned into three subsystems $A - C - B$ by cutting off two single σ bonds; where A and B are two fragments whose interaction energy calculation will be involved, and C is the fragment that represents the central linker of fragments A and B . The identification of the fragments A , B , and C is required in the ISAPT calculation input.

In a molecule, the standard HF-SCF calculation is a starting point to obtain the zeroth-order HF wavefunction and occupied orbitals of the complete chemical system. The HF occupied orbitals are then localized and assigned to the fragments A , B , and C . The intrinsic bond orbital (IBO), in Section 2.4, is utilized in the localization of the HF molecular occupied orbitals. [34] The HF wavefunctions of fragments A and B are obtained by performing the HF calculation for the substituted occupied orbitals of fragments A and B in the potential fields of the linker C , respectively. The occupied orbitals of the fragments A and B of the HF embedded wavefunctions are then involved in the energy calculation following the SAPT0 protocol.

An important drawback of the ISAPT method arises from the artificial dipole moment that presents on the boundary of the interfragments A and B . [46, 49] In the molecular fragmentation, the electrons of the linking IBOs of the $A - C$ and $B - C$ bonds are assigned to the fragment C . The linking atoms on the boundary of the fragments A and B are thus missing

one proton and one electron in the hybrid orbital in the direction of the linking IBOs, leading to the incomplete electronic densities for those two fragments. In some fragmentation models of the polar systems that feature the hydrogen bonding, the issue results in the unphysical quantification of the interaction energy components, including the strong repulsion of the first-order electrostatic energy and the overestimation of the second-order induction and exchange induction energies.

2.3.1 Splitting of Intrinsic Atomic Orbitals

In order to overcome the issue of the artificial dipole moment in the ISAPT calculations, in the thesis, the method of the Splitting of Intrinsic Atomic Orbitals (SIAO) is introduced by performing the reassignment of the electrons in the link IBOs of the bonds $A - C$ and $B - C$, where the electrons in the link IBOs are equally assigned between A/B and C . [39] The hybrid orbitals χ_x and χ_y , serving as the atomic orbitals in the linking IBOs, of the linking atoms at the interfragments A and B , respectively, are identified by projecting the link IBOs of the bonds $A - C$ and $B - C$ onto the occupied IAOs space of the corresponding fragments. The missing electrons on the linking atoms at the interfragment boundaries of fragments A and B are placed into the corresponding hybrid orbitals χ_x and χ_y . The augmented densities of the fragments A and B , in the presence of the reassigned hybrid orbitals χ_x and χ_y , are generated by the HF calculations of the modified embedding potentials of the fragments A and B . The occupied space of the fragment A is reoptimized in an embedding excluding χ_y and the occupied space of the fragment B is reoptimized in an embedding excluding χ_x . The link orbitals χ_x and χ_y are enforced orthogonalization to the augmented space of the corresponding fragments A and B .

In the evaluation of the reassigned link spin-orbital χ_x/χ_y , the formulations of the first- and second-order exchange energy corrections are regenerated in cases of the spin-up and spin-down electrons. Ideally, the $\chi_x - \chi_y$ interaction would be analyzed as “spin-free” as possible, imposing a random relative orientation of the two spins. The singly occupied spin-orbital χ_x/χ_y can be expressed in the state of half spin-up $|\uparrow\rangle$ and half spin-down $|\downarrow\rangle$ that gives rise to two

possible spin couplings between two spin-orbitals χ_x and χ_y , including the parallel $\langle \psi_x^{\parallel} | \psi_y^{\parallel} \rangle$ and perpendicular $\langle \psi_x^{\perp} | \psi_y^{\perp} \rangle$ spins in the following expressions:

$$\psi_x^{\parallel} = \frac{1}{\sqrt{2}} \chi_x (|\uparrow\rangle + |\downarrow\rangle) \quad (2.17)$$

$$\psi_y^{\parallel} = \frac{1}{\sqrt{2}} \chi_y (|\uparrow\rangle + |\downarrow\rangle) \quad (2.18)$$

$$\psi_x^{\perp} = \frac{1}{\sqrt{2}} \chi_x (|\uparrow\rangle - |\downarrow\rangle) \quad (2.19)$$

$$\psi_y^{\perp} = \frac{1}{\sqrt{2}} \chi_y (|\uparrow\rangle - |\downarrow\rangle) \quad (2.20)$$

$$\langle \psi_x^{\parallel} | \psi_y^{\parallel} \rangle = \langle \chi_x | \chi_y \rangle \quad (2.21)$$

$$\langle \psi_x^{\perp} | \psi_y^{\perp} \rangle = 0 \quad (2.22)$$

In the augmented occupied IAOs spaces of the fragments A and B , the single electron transposition accounts for the pairwise interactions of two sets of occupied orbitals that involve χ_x and the IAOs space of fragment A , and χ_y and the IAOs space of fragment B . In the case of the perpendicular spin, the contribution of the interaction of χ_x and χ_y is zero. Consequently, the exchange corrections in the S^2 -approximation are the average of the exchange energies in terms of the parallel and perpendicular spins.

2.3.2 Range Separation of the Interaction Potential in SAPT and ISAPT methods

The recent developments of the ISAPT methods rely on the scheme of molecular fragmentation, in which the occupied orbitals of the fragment linker C are excluded from the energy calculations. [39, 49] In addition, the interaction energy within the fragment A/B is neglected in the ISAPT methods. In the motivation for the full description of the non-covalent interaction

energy in a molecule, the range separation of the interaction potential is one of the possible methods to construct an ISAPT algorithm that is independent of fragmentation. In range separation methods, in the thesis, the Gaussian and error functions are used to separate the perturbation operator $\hat{V}$ into parts of long- and short-range potentials. [1, 9, 31, 40, 47, 52, 67]

The range separation is a versatile computational tool that is commonly applied in density functionals incorporating long-range Hartree–Fock (HF) exchange to avoid some common issues such as excessive electron delocalization and the resulting overstabilization of charge-transfer states. [31, 38, 67] They are also useful in the context of SAPT(DFT) to avoid an explicit asymptotic correction. [22, 35, 69] On the other hand, density functionals with short-range HF exchange only improve the computational efficiency for periodic systems. [27] Another application of the range separation technique occurs in the regularized SAPT method, addressing the challenges associated with the divergence of the perturbation series [2, 52, 53] as well as allowing a decomposition of induction energy into polarization and charge-transfer contributions. [47]

$$\frac{1}{r} = v_l(r) + v_s(r) \quad (2.23)$$

$$v_l^{gau}(r) = \frac{1 - e^{-\eta r^2}}{r} \quad (2.24)$$

$$v_s^{gau}(r) = \frac{e^{-\eta r^2}}{r} \quad (2.25)$$

$$v_l^{erf}(r) = \frac{erf(\omega r)}{r} \quad (2.26)$$

$$v_s^{erf}(r) = \frac{1 - erf(\omega r)}{r} \quad (2.27)$$

where $\frac{1}{r}$ is the Coulomb potential of the perturbation operator $\hat{V}$ that can be rewritten as the combination of the long-range $v_l(r)$ and short-range $v_s(r)$ potentials. $v_l^{gau}(r)$ and $v_s^{gau}(r)$ are the Gaussian function splitting with the separation parameter η . $v_l^{erf}(r)$ and $v_s^{erf}(r)$ are the error function splitting with the separation parameter ω .

In the SAPT and ISAPT/SIAO1 energy calculations, the long-ranged Coulomb potentials, including the attractive one-electron and repulsive two-electron integrals, are used to evaluate the asymptotic behavior and the consistency of the interaction energy terms when the separation parameter increases. In SAPT energy corrections, the long-ranged potential is more important in the description of the intermolecular interaction between two monomers. The range-separated one-electron integrals are computed in the McMurchie-Davidson recursion scheme, in Section 2.5, in terms of the Gaussian and error functions. [43]

2.4 Intrinsic Bond Orbital and Intrinsic Atomic Orbital Localizations

In the space of a molecular orbital, the intrinsic atomic orbital (IAO) describes the polarized AO centered on a given atom. [34]. The polarized AOs basis contains a set of the minimal and valence AOs that exactly express the occupied MOs of the HF-SCF wavefunction. The minimal basis set is a set of basis functions that incorporates only the occupied orbitals in the electronic configuration of an atom.

The intrinsic bond orbital (IBO) is constructed as a linear combination of the intrinsic hybrid orbitals (IHOs), which are the IAOs pointing at the bond direction.

$$L = \sum_A^{atoms} \sum_i^{occ} [n_A(i)]^4 \quad (2.28)$$

where n_A is the number of electrons of the occupied orbitals located on atom A .

2.5 McMurchie-Davidson recursion for Gaussian integrals

In the McMurchie-Davidson recursion, the one-electron integral relies on expanding the products of two Cartesian basis functions into one-center Hermite Gaussians. [43,44,73] The s-type nuclear attraction integral is a starting point for the angular momentum recursion relation.

The Cartesian Gaussian function is an atomic Gaussian-type-orbital (GTO) basis function centered on the atomic nucleus.

$$G_{ijk,\eta} = x^i y^j z^k e^{-\eta r^2} \quad (2.29)$$

where r is the length of the position vector of an electron pointing from the nucleus. η is the exponential parameter that determines the size of the orbital. $L = i + j + k$ is the angular momentum of the atomic orbital.

In the molecular space, the two-center s-type Gaussian function at two positions of the nuclei is expressed as the product of two s-type GTOs. The generated one-center Cartesian Gaussian function at the position $\mathbf{P} = (P_x, P_y, P_z)$ is expressed as follows:

$$\chi_{\mathbf{P}}(\mathbf{r}) = \chi_{\mathbf{A}}(\mathbf{r})\chi_{\mathbf{B}}(\mathbf{r}) = e^{\frac{-\alpha\beta}{\alpha+\beta}(\mathbf{A}-\mathbf{B})^2} e^{-(\alpha+\beta)(\mathbf{r}-\mathbf{P})^2} \quad (2.30)$$

$$\mathbf{P} = \frac{\alpha\mathbf{A} + \beta\mathbf{B}}{\alpha + \beta} \quad (2.31)$$

where α and β are the exponents of the GTOs centered on the nuclei $\mathbf{A}$ and $\mathbf{B}$, respectively.

The one-center Hermite Gaussians in the expansion of two-center product GTOs at the center $\mathbf{P}$ with exponent $p = \alpha + \beta$ has the following expression:

$$\Lambda_{tuv}(\mathbf{r}, p, \mathbf{P}) = \left(\frac{\partial}{\partial P_x}\right)^t \left(\frac{\partial}{\partial P_y}\right)^u \left(\frac{\partial}{\partial P_z}\right)^v \exp(-p(\mathbf{r} - \mathbf{P})^2) \quad (2.32)$$

where t, u, v are the order of Hermite Gaussian derivatives and the sum $t + u + v$ represents the angular momentum of the product of two Cartesian Gaussian basis functions.

The nuclear-attraction integral of the Coulomb potential with the location of each atomic nucleus $\mathbf{C} = (Cx, Cy, Cz)$ for a Hermite Gaussian is expressed as follows:

$$V_{tuv}^{Coul} = \int \Lambda_{tuv}(\mathbf{r}, p, \mathbf{P}) \frac{1}{|\mathbf{r} - \mathbf{C}|} d\mathbf{r} = \frac{2\pi}{p} \left(\frac{\partial}{\partial P_x} \right)^t \left(\frac{\partial}{\partial P_y} \right)^u \left(\frac{\partial}{\partial P_z} \right)^v (-2p)^n \mathbf{F}_n(pR_{\mathbf{PC}}^2) \quad (2.33)$$

where $R_{\mathbf{PC}}$ is the length of the position vector $\vec{\mathbf{PC}}$. $\mathbf{F}_n(x)$ is the Boys function of order n that is defined in Eq. 2.34. The s-type integral has the zero order of the Boys function and zero values of (t, u, v) .

$$\mathbf{F}_n(x) = \int_0^1 u^{2n} \exp(-xu^2) du \quad (2.34)$$

The nuclear-attraction s-type integrals using the potentials of the short-range Gaussian and long-range error functions with the range separation parameter η for a Hermite Gaussian are derived as follows:

$$V_{000}^{Gau} = \frac{2\pi}{p + \eta} \exp\left(\frac{-p\eta}{p + \eta} R_{\mathbf{PC}}^2\right) \mathbf{F}_0\left(\frac{p^2}{p + \eta} R_{\mathbf{PC}}^2\right) \quad (2.35)$$

$$V_{000}^{Erf} = \frac{2\pi}{p} \left(\frac{\eta^2}{p + \eta^2} \right)^{1/2} \mathbf{F}_0\left(\frac{p\eta^2}{p + \eta^2} R_{\mathbf{PC}}^2\right) \quad (2.36)$$

The angular recursion formulas over t, u, v of the nuclear attraction integrals are written in Ref. [40] in Eqs. 21-23 for Gaussian function and in Eqs. 29-31 for error function.

Chapter 3

Overcoming Artificial Multipoles in Intramolecular Symmetry-Adapted Perturbation Theory

This chapter was published as Luu, D.; Patkowski, K. *JPCA* (2023).

The ISAPT is a computational method that provides the decomposition of the noncovalent intramolecular interactions within a molecule, including electrostatic, exchange, induction, and dispersion effects. The development of the ISAPT methodology depends on the strategy of molecular fragmentation, orbitals assignment, and wavefunction embedding. [49] If a molecule can be partitioned into three species $A - C - B$, where $A - C$ and $B - C$ are covalently single bonds, the interaction energy between A and B in the presence of C can be quantified using the SAPT procedure. The HF wavefunctions for A and B are obtained by solving the HF equations for each fragment embedded in the HF potential of linker C . The product of the A and B wavefunctions obtained in this way can then be used as the zeroth-order wavefunction for a standard SAPT expansion. In the default ISAPT protocol, the doubly occupied IBOs describing the interfragment bonds are treated as part of the linker C that the spurious dipoles are created at the interfragment boundaries of A and B . For example, the fragmentation is performed for the propane molecule to calculate the noncovalent interaction between two CH_3 -groups connected by the $-\text{CH}_2-$ linker. Then, each of the fragments A and B features a central carbon atom, with its nuclear charge reduced by one to maintain zero net charges on A and B , and three hydrogen atoms located along the directions of three sp^3 hybrid orbitals. The fourth hybrid orbital in the tetrahedral symmetry of the linking carbon atom is missing, such a fragment has a substantial dipole moment, and the interaction between these spurious dipoles on A and B dominate the SAPT electrostatic contribution. The electrostatic energy is observed in strong repulsion in the molecular models that favor the hydrogen bond interaction. A solution is proposed to overcome the issue of the spurious dipoles by reassigning the doubly occupied

interfragment IBOs, together with the corresponding +1 nuclear charges, to fragments A and B or equally shared between A/B and C .

The improved ISAPT variants involve SAOn and SIAOn, where $n = 0, 1, 2$ indicating the iteration of the HF-SCF relaxation on the modified embedding potential of the fragments A and B . [39] The two algorithms differ from the space of projection for the identification of link IHOs. In the SAOn variants, the occupied space is spanned by the AO basis functions centered on fragments A/B , while in the SIAOn variants, the occupied space is spanned by the intrinsic atomic orbitals [34] centered on fragments A/B . The improved ISAPT variants are implemented in the Psi4 software package. [63] The one iteration toward self-consistency, as in the SAO1 and SIAO1 methods, is entirely sufficient for the evaluation of the interaction energy. Both ISAPT(SAO1) and ISAPT(SIAO1) resolve the issue with artificial dipole moments on the linking atoms. The ISAPT(SAO1) has a somewhat larger reduction in the magnitude of the fragment dipole moments in systems with entirely nonpolar fragments, e.g. Table 3.1, see Table S160 in the Supporting Information of Ref. [39] for other systems.

In the ISAPT(SAO1) and ISAPT(SIAO1) calculations, all energy contributions exhibit physically meaningful values for all fragmentation schemes, e.g. Fig. 3.1 for the P244 model of the 2,4-pentanediol molecule and Fig. 3.2 for the C73 and C7B24 models of the linear and branch heptane molecules, see the Supporting Information of Ref. [39] for all fragmentation models. In particular, the electrostatic energy is corrected into strong attraction for intramolecular hydrogen-bonded systems (Fig. 3.1) and weak attraction (due to charge penetration effect) for nonpolar fragments (Fig. 3.2). The induction and exchange-induction energies are not nearly as large as in the original ISAPT(C) variant, which makes their sum more meaningful. The first-order exchange energy typically becomes less repulsive, and the dispersion and exchange-dispersion contributions are in good agreement with the original ISAPT(C) method. ISAPT(SIAO1) is the best variant for calculating and interpreting intramolecular interactions that shows remarkable consistency between basis sets due to a much smaller overlap between the link IHOs, e.g. Figs. 3.1-3.3. In the spin couplings of the first-order exchange correction,

the parallel and perpendicular values from the ISAPT(SIAO1) are highly consistent for all systems studied in Fig. 3.3. The average of the parallel and perpendicular spin-coupling values is chosen as the total exchange contribution.

Table 3.1: Length of HF dipole moments (in a.u.) of fragments *A* and *B* computed with the ISAPT variants of (original) C, SAO1, and SIAO1 on bases aDZ, aTZ, and aQZ.

Conf	Basis	ABC	Fragment A			Fragment B		
			C	SAO1	SIAO1	C	SAO1	SIAO1
C7B24	aDZ	0.045	0.799	0.091	0.385	0.799	0.091	0.385
	aTZ	0.044	0.800	0.178	0.384	0.800	0.177	0.384
	aQZ	0.044	0.800	0.100	0.384	0.800	0.100	0.384

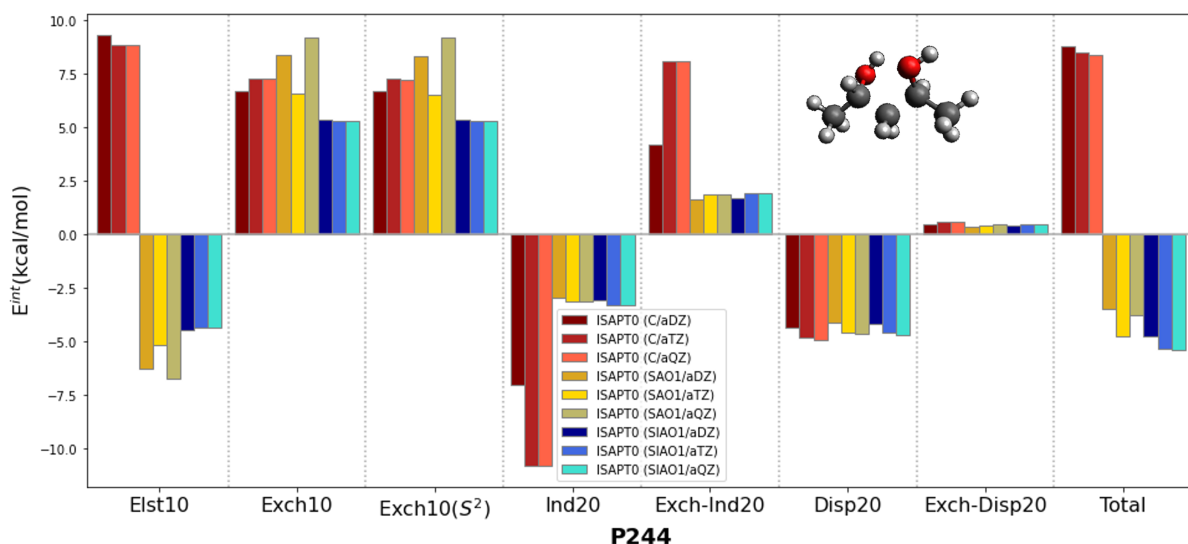


Figure 3.1: ISAPT energy components in 2,4-pentandiol (P242 fragmentation pattern) computed using the aXZ bases (X=D, T, Q) and three link assignments: (original) C, SAO1, and SIAO1.

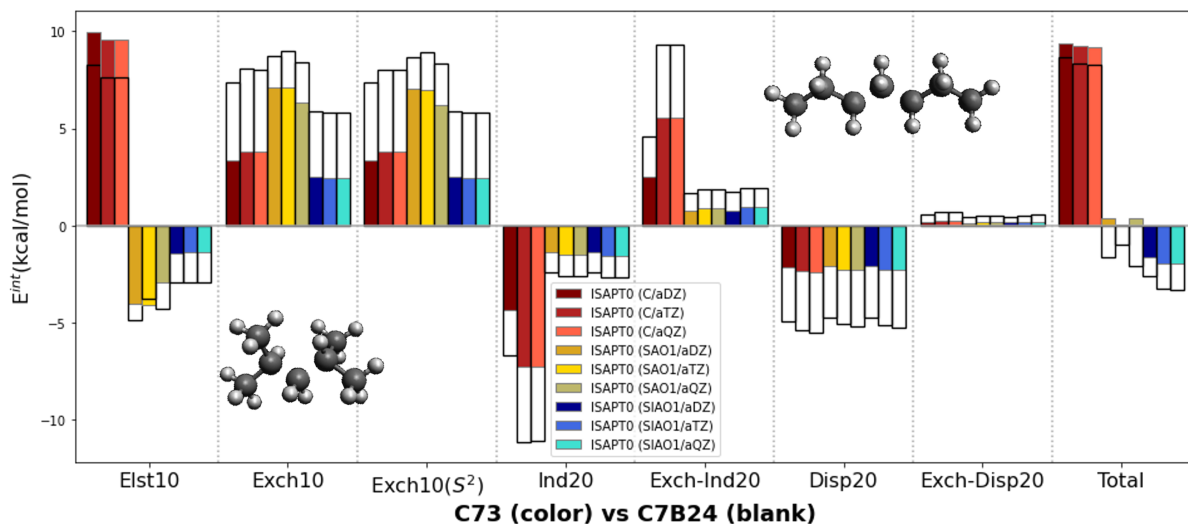


Figure 3.2: Comparison of the ISAPT0 energy components between n-heptane (C73, colored bars) and its branched isomer, 2,4-dimethylpentane (C7B24, empty bars) computed with the (original) C, SAO1, and SIAO1 variants using the aXZ bases (X=D, T, Q).

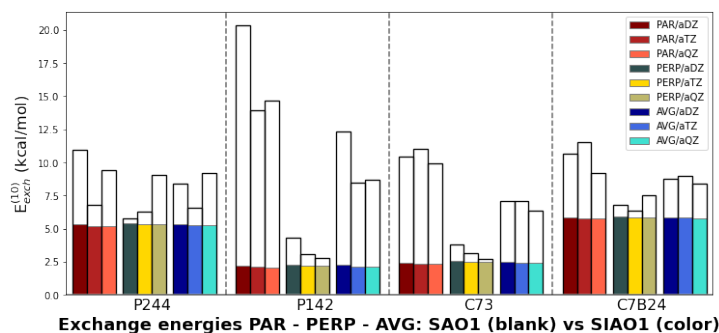


Figure 3.3: ISAPT exchange energy $E_{\text{exch}}^{(10)}$ from parallel (PAR), perpendicular (PERP), and average (AVG) spin couplings in models P244, P142, C73, and C7B24. The values are computed with SAO1 (empty bars) and SIAO1 (colored bars) in the aXZ bases, X=D,T,Q.

Overcoming Artificial Multipoles in Intramolecular Symmetry-Adapted Perturbation Theory

Published as part of *The Journal of Physical Chemistry virtual special issue "MQM 2022: The 10th Triennial Conference on Molecular Quantum Mechanics"*.

Du Luu and Konrad Patkowski*



Cite This: *J. Phys. Chem. A* 2023, 127, 356–377



Read Online

ACCESS |



Metrics & More



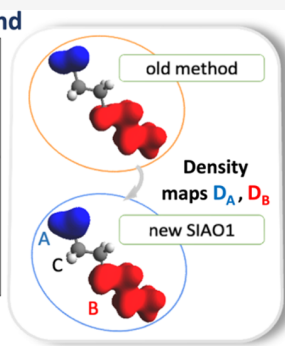
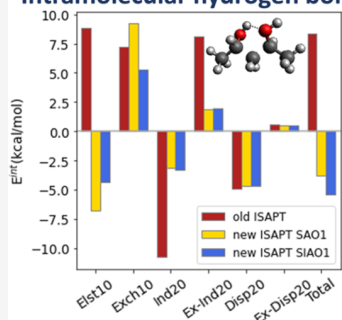
Article Recommendations



Supporting Information

ABSTRACT: Intramolecular symmetry-adapted perturbation theory (ISAPT) is a method to compute and decompose the noncovalent interaction energy between two molecular fragments A and B covalently connected via a linker C. However, the existing ISAPT algorithm displays several issues for many fragmentation patterns (that is, specific assignments of atoms to the A/B/C subsystems), including an artificially repulsive electrostatic energy (even when the fragments are hydrogen-bonded) and very large and mutually cancelling induction and exchange-induction terms. We attribute those issues to the presence of artificial dipole moments at the interfragment boundary, as the atoms of A and B directly connected to C are missing electrons on one of their hybrid orbitals. Therefore, we propose several new partitioning algorithms which reassign one electron, on a singly occupied link hybrid orbital, from C to each of A/B. Once the contributions from these link orbitals are added to fragment density matrices, the computation of ISAPT electrostatic, induction, and dispersion energies proceeds exactly as normal, and the exchange energy expressions need only minor modifications. Among the link partitioning algorithms introduced, the so-called ISAPT(SIAO1) approach (in which the link orbital is obtained by a projection onto the intrinsic atomic orbitals (IAOs) of a given fragment followed by orthogonalization to this fragment's occupied space) leads to reasonable values of all ISAPT corrections for all fragmentation patterns, and exhibits a fast and systematic basis set convergence. This improvement is made possible by a significant reduction in magnitude (even though not a complete elimination) of the unphysical dipole moments at the interfragment boundaries. We demonstrate the utility of the improved ISAPT partitioning by examining intramolecular interactions in several pentanediol isomers, examples of linear and branched alkanes, and the open and closed conformations of a family of N-arylimide molecular torsion balances.

Intramolecular hydrogen bond



INTRODUCTION

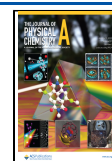
Weak noncovalent interactions are ubiquitous in nature as they occur both between separate molecules and between nonbonded fragments of the same molecule. As a result, the knowledge and understanding of these interactions is indispensable in many areas of chemistry, physics, materials science, and biology. When the energy of a weak interaction is computed using quantum chemistry, it is highly desirable to use a computational protocol that is both quantitative and conceptual, that is, it provides an accurate overall value and its meaningful decomposition into individual terms that can be separately interpreted and understood. There are many approaches that satisfy one of these conditions but only very few methods can do both; among the latter, symmetry-adapted perturbation theory (SAPT)^{1,2} is a particularly attractive choice thanks to its many robust variants and efficient computer implementations. The interaction energy in SAPT

is decomposed into four major contributions: electrostatics (Coulomb interaction of unperturbed molecular charge densities, including charge penetration), exchange (the short-range repulsion stemming from the Pauli exclusion principle), induction (the polarization of one molecule by the interacting partner), and dispersion (the correlation between instantaneous charge density fluctuations). For example, at the lowest (qualitatively accurate) SAPT0 level of theory, the interaction energy is approximated as a sum of the following corrections, grouped together to account for the four major contributions:

Received: September 9, 2022

Revised: November 29, 2022

Published: December 23, 2022



ACS Publications

© 2022 American Chemical Society

356

<https://doi.org/10.1021/acs.jpca.2c06465>
J. Phys. Chem. A 2023, 127, 356–377

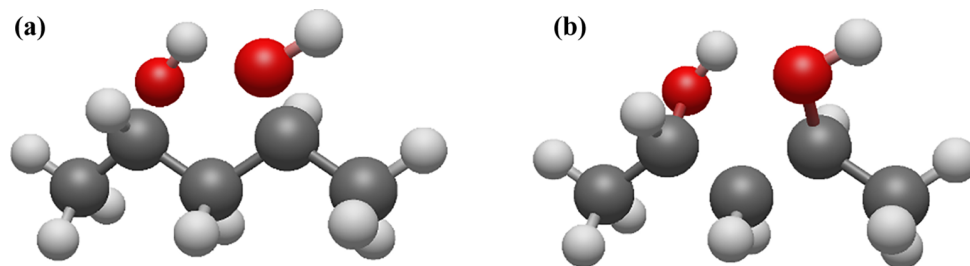


Figure 1. Two different partitionings of the 2,4-pentanediol molecule into hydrogen bonded fragments **A** and **B** covalently connected through a linking fragment **C**. A missing bond signifies a fragment boundary.

$$E_{\text{int}}^{\text{SAPT0}} = (E_{\text{elst}}^{(10)}) + (E_{\text{exch}}^{(10)}) + (E_{\text{ind,resp}}^{(20)} + E_{\text{exch-ind,resp}}^{(20)} + \delta_{\text{HF}}) + (E_{\text{disp}}^{(20)} + E_{\text{exch-disp}}^{(20)}) \quad (1)$$

In eq 1 and below, the consecutive superscripts denote orders of perturbation theory with respect to intermolecular interaction and intramolecular correlation (thus, there is no intramolecular correlation in SAPT0). Furthermore, the additional subscript “resp” denotes the relaxation (response) of each molecule’s Hartree-Fock (HF) orbitals to the electric field of the interacting partner, and the δ_{HF} term brings in higher-order induction and exchange-induction effects contained in the supermolecular HF interaction energy. SAPT has been employed to compute and interpret intermolecular interactions for numerous complexes of theoretical and experimental interest; many of these applications, together with major advances in the SAPT methodology, have been summarized in the recent reviews.^{3–7}

While there are many options to decompose the intermolecular interaction energy, and several ways to do the same thing for a covalent bond,^{8–10} the physical decomposition of intramolecular noncovalent interactions is much less understood. While it is not at present possible to quantify the *entire* nonbonded interaction involving all molecular fragments, if one can identify two nonbonded fragments **A** and **B** connected to each other through a linker fragment **C**, there are a few options to define and partition the interaction energy between **A** and **B** in the presence of **C**. On one hand, one could remove **C** and cap the dangling bonds of **A** and **B** with hydrogen atoms. As the interaction in question is now intermolecular, one can use any of the standard SAPT variants to evaluate its contributions.^{11,12} However, the fragmentation approach requires altering the molecular system and, if the linking fragment is small, artificial means to avoid an overly pronounced repulsion between the capping hydrogens. Two alternatives that do not require cutting and capping have been introduced by Gonthier and coworkers.^{13–15} The first one involves localizing the occupied orbitals on the fragments **A**, **B**, and **C** and constructing a non-Hermitian zeroth-order Hamiltonian¹³ in which the **A–B** interactions are removed following the Chemical Hamiltonian approach introduced by Mayer.¹⁶ These removed interactions are then brought back as the perturbation that gives rise to the noncovalent intramolecular interaction.¹⁵ The need for a non-Hermitian, biorthogonal formalism makes the theory somewhat more complicated, and the resulting energy decomposition is a little different than in standard SAPT: the polarization corrections cannot be separated from their exchange counterparts (for example, the entire first-order energy is a single term), but the

second-order energy includes an explicit delocalization (charge transfer) term in addition to the induction and dispersion ones.

The second approach¹⁴ leads to the same set of corrections as intermolecular SAPT (at its lowest, SAPT0 level). In this method, referred to as ISAPT, the occupied orbitals are localized according to the intrinsic bond orbital (IBO) scheme.¹⁷ Subsequently, the HF wavefunctions for **A** and **B** are obtained by solving the HF equations for each fragment embedded in the HF potential of linker **C** (that is, the **A–C** and **B–C** interactions are included but the **A–B** one is not). The product of the **A** and **B** wavefunctions obtained in this way can then be used as the zeroth-order wavefunction for a standard SAPT expansion.

In any variant, the treatment of the interfragment boundaries is critical to the performance of the method. In the default ISAPT protocol, the doubly occupied IBOs describing the interfragment bonds are treated as part of the linker **C** (it is assumed that (**A,C**) and (**B,C**) are each connected by a single bond only). To maintain zero net charges on **A** and **B** (as long as the entire molecule is electrically neutral), one proton worth of the charge on the atom from **A** (**B**) that takes part in the interfragment bond is reassigned to **C**. Thus, no spurious charges are created at the interfragment boundaries (obviously, the presence of such charges would badly skew the SAPT electrostatic energy). However, the same cannot be said about spurious dipole moments, even when the interfragment bonds are nonpolar. Imagine, for illustration, that the fragmentation is performed for the propane molecule to calculate the noncovalent interaction between two CH_3- groups connected by the $-\text{CH}_2-$ linker. Then, each of the fragments **A** and **B** features a central carbon atom, with its nuclear charge reduced by one, and three hydrogen atoms located along the directions of three sp^3 hybrid orbitals. However, as the fourth hybrid orbital that would complete the tetrahedral symmetry is missing, such a fragment has a substantial dipole moment, and the interaction between these spurious dipoles on **A** and **B** might dominate the SAPT electrostatic contribution. If the doubly occupied interfragment IBOs, together with the corresponding +1 nuclear charges, are instead reassigned to fragments **A** and **B** or equally shared between **A/B** and **C**,¹⁴ the spurious dipole moments remain, and an additional problem appears when the linker is small and the resulting charge densities of **A** and **B** are too close together. For instance, in our propane example, such a reassignment is not acceptable because it would lead to assigning a +1 nuclear charge on the same central carbon atom to both **A** and **B**.

While the process of fragmentation and the definition of intramolecular SAPT corrections is not unique as it does not lead to any experimental observables, a useful approach should lead to ISAPT results that make physical sense. For example,

an intramolecular hydrogen bond should show up as a favorable interaction both in terms of the total SAPT energy and its electrostatic component, similar to an intermolecular hydrogen bond. This is not always the case for any of the variants (link orbital assignments) of the original ISAPT method of ref 14. A good example is the 2,4-pentanediol molecule that involves an intramolecular hydrogen bond. In the fragmentation scheme of Figure 1(a), which was extensively studied in ref 14, the ISAPT results make sense and resemble the SAPT decomposition for the water dimer; in particular, the electrostatic energy is attractive (negative). However, in the fragmentation scheme of Figure 1(b), the electrostatic energy is strongly positive, which does not make physical sense as the addition of nonpolar hydrocarbon chains to A and B cannot offset the favorable dipole–dipole interaction between the –OH groups. We believe that the reason for this unphysical behavior is the artificial dipole moments resulting from the interfragment boundaries cutting through the (nonpolar) C–C bonds.

The problem of unphysical multipoles at the interfragment boundary was already mentioned in ref 14, and it was further explored by Meitei and Hesselmann in ref 12. These authors, in the specific context of highly sterically crowded hydrocarbons such as the all-meta *tert*-butyl derivative of hexaphenylethane,¹⁸ proposed a solution to replace the embedding in the linking fragment (in this case, two central carbon atoms) by an embedding in a set of capping hydrogen atoms at very short C–H bond lengths to reduce the artificial dipoles. In most cases, this approach indeed led to a sensible, somewhat attractive electrostatic energy; however, such a correction is very system specific as it cannot be expected that the embedding in the capping hydrogens can replace an embedding in *any* linker fragment C. Therefore, the issue of unphysical dipoles appearing at the ISAPT interfragment boundary remains unsolved, limiting the applications of the method to highly favorable cases and/or cases where the errors cancel (which might be expected when investigating differences between similar systems).

In this work, we propose several algorithms to repartition the link bonds between fragments to reduce the unphysical fragment dipoles while preserving the embedding in the actual, unaltered chemical system. In a sense, we follow a direction stated, but not pursued, by Meitei and Hesselmann,¹² and explore an alternative assignment of the linking electron pair—the one where one electron is ascribed to each of the two adjacent fragments. This leads to several new ISAPT variants, differing in the way in which the linking IBO is deconstructed into a pair of intrinsic hybrid orbitals (IHOs) on the bonded atoms.¹⁷ It should be stressed that our goal is not the ultimate, rigorous definition of ISAPT corrections (we are fully aware that such fragment–fragment contributions are not measurable quantities), but a practical and robust ISAPT algorithm whose results make physical sense for a wide variety of molecular systems and fragmentation patterns. In particular, we expect a practical ISAPT variant

- (a) to substantially reduce the magnitude of the fragment dipole moments when a nonpolar system is partitioned into nonpolar fragments,
- (b) to give attractive electrostatic energies for intramolecular hydrogen-bonded fragments, and slightly attractive electrostatic energies between nonpolar fragments (due to charge penetration^{19,20}),

- (c) to closely follow the results of standard intermolecular SAPT at large A–B separations,
- (d) to exhibit stable behavior as the one-electron basis set is enlarged.

In the remainder of this article, we develop and test our new ISAPT link bond partitionings with a goal of satisfying the conditions (a)–(d) above. First, in the **Methodology** section, we describe the new algorithms. Next, the new partitionings are applied to a set of representative intramolecular interactions and compared with the original link assignment of ref 14. The final **section** contains conclusions.

METHODOLOGY

Definitions. In the following discussion, the capital letters $K, L, M, \dots$ will denote atomic orbital (AO) basis functions (the AO basis always covers the entire molecule). The fragments A and B interact with each other only noncovalently, but are covalently connected via the linker C. We will assume that A and C are connected by a single bond, and likewise for B and C; let ϕ_{AC} and ϕ_{BC} denote the connecting IBOs.¹⁷ The initial set of IBOs is obtained by localizing the HF wavefunction for the entire molecule, and its orthogonal subspaces containing IBOs for the respective fragments will be denoted by $\mathcal{B}_A^0, \mathcal{B}_B^0$, and $\mathcal{B}_C$. Subsequently, the occupied orbitals of A and B are reconstructed with the A–B interaction switched off but the A–C and B–C ones fully present. Specifically, the HF equations for A and B embedded in the frozen HF wavefunction of C are solved at this point. Afterwards, the entire one-electron space $\mathcal{B}$ (the space spanned by the AOs) includes occupied subspaces $\mathcal{B}_A, \mathcal{B}_B, \mathcal{B}_C$ for the respective fragments, with both $\mathcal{B}_A$ and $\mathcal{B}_B$ orthogonal to $\mathcal{B}_C$ (however, $\mathcal{B}_A$ is not orthogonal to $\mathcal{B}_B$; note also that the subspace for the linker C has not changed in the last step). The virtual orbitals for fragments A and B live in the orthogonal complements to the spaces $\mathcal{B}_A \oplus \mathcal{B}_C$ and $\mathcal{B}_B \oplus \mathcal{B}_C$, respectively. Thus, the occupied orbitals on C are projected out from both the occupied and virtual orbitals on A/B.

In the original ISAPT method¹⁴ with the default link assignment, the doubly occupied orbitals ϕ_{AC} and ϕ_{BC} are assigned to linker C together with a single proton from the nucleus of A/B that participates in the covalent bond (the remaining protons are still assigned to A/B, so that this nucleus has an effective charge reduced by one within that fragment). Thus, if the original molecule is electrically neutral, all fragments are neutral too. Unfortunately, the same cannot be said about the dipole moments on the individual fragments. The connecting atom on A or B, besides missing one proton, is missing an electron on one of its hybrid orbitals, the one responsible for the covalent bond with C. Thus, even for a nonpolar environment such as an sp^3 carbon in a hydrocarbon molecule, an artificial dipole moment is created on the fragment A/B because one hybrid orbital, with its bonding electron, has been removed. We believe that this artificial dipole moment is the cause of often unphysical interfragment electrostatic energies in ISAPT.

Link Orbital Reassignment. The IBOs ϕ_{AC} and ϕ_{BC} can be viewed as linear combinations of “intrinsic hybrid orbitals” (IHOs) on the atoms connected by the linking bond—combinations of intrinsic atomic orbitals (IAOs)¹⁷ resulting in a hybrid orbital pointing in the bond direction. We propose to partially undo the IAO $\rightarrow$ IBO transformation and identify IHOs χ_x and χ_y that, in some sense, constitute the A-fragment

and B-fragment parts of ϕ_{AC} and ϕ_{BC} , respectively. More than one algorithm of determining these IHOs is possible and we will present two specific choices below. Normally, we do assume that χ_x and χ_y , like the whole link orbital space where they come from, are orthogonal to the occupied spaces $\mathcal{B}_A$ and $\mathcal{B}_B$, respectively. Note that this orthogonalization will have to be explicitly enforced as χ_x and χ_y are not contained in the subspace $\mathcal{B}_C$; we will also investigate what happens if one forgoes this orthogonalization. In this process of bond orbital decomposition, we associate one electron with each of χ_x and χ_y and, for the purpose of updating density matrices, we will assume that half of this electron comes in with spin up and half with spin down. In our new link assignments, the single electron on χ_x is reassigned from C to A together with the corresponding +1 nuclear charge on the A atom connected to C. Thus, the assignment of nuclear charges is now very simple—the full charges of all nuclei in A belong to A—but the assignment of electronic charges becomes more complicated. An analogous reassignment takes place between C and B—the net result is that A and B gain one electron each relative to the default ISAPT link assignment. Our hope is that, by the addition of electrons on χ_x and χ_y , the fragment charges around the linking atoms become more spherically symmetrical, and any unphysical dipole moments at the fragment boundary should be substantially diminished.

First-Order ISAPT Terms. At this point, the redefinition of the ISAPT first-order electrostatic energy is straightforward. The nuclear potentials of fragments A and B are modified to account for the entire charge $+Ze^-$, not $+(Z-1)e^-$, on the linking atom. The HF-level fragment density matrices D^A and D^B are supplemented by the contributions of the (singly occupied) orbitals χ_x and χ_y , respectively. The updated nuclear potentials and density matrices are then used in standard SAPT expressions to determine the ISAPT $E_{\text{elst}}^{(10)}$ term between fragments A and B. Note that while the addition of linking electrons to A and B is reflected in the updated density matrices, their removal from C must be reflected in the embedding. For example, the embedded HF calculation for A still includes the A–C linking pair, with one electron now in A and one still in C. However (as explained in detail later), the treatment of the more distant interfragment boundary B–C has been altered as one linking electron has been removed from C and thus from the embedding potential.

The $E_{\text{exch}}^{(10)}$ correction, within its customary S^2 approximation,¹ requires, in addition to the full density matrix, its partitioning into spin density matrices. It is natural to assume that the supplementary contributions of χ_x and χ_y are divided equally between the spin-up and spin-down density matrix (half spin-up and half spin-down). However, this assumption does not determine the exchange energy uniquely, as the interactions between the electrons on χ_x and χ_y depend on the mutual alignment of their spins. In the two limiting cases that both lead to equal spin-up and spin-down density matrices, the link spinorbitals ψ_x and ψ_y can have parallel spins,

$$\psi_x^{\parallel} = \frac{1}{\sqrt{2}}\chi_x(|\uparrow\rangle + |\downarrow\rangle) \quad \psi_y^{\parallel} = \frac{1}{\sqrt{2}}\chi_y(|\uparrow\rangle + |\downarrow\rangle) \quad (2)$$

or perpendicular spins,

$$\psi_x^{\perp} = \frac{1}{\sqrt{2}}\chi_x(|\uparrow\rangle + |\downarrow\rangle) \quad \psi_y^{\perp} = \frac{1}{\sqrt{2}}\chi_y(|\uparrow\rangle - |\downarrow\rangle) \quad (3)$$

Note that the parallel case corresponds to $\langle\psi_x^{\parallel}|\psi_y^{\parallel}\rangle = \langle\chi_x|\chi_y\rangle$ and includes one electron worth of exchange interaction between the link spinorbitals. The perpendicular case leads to $\langle\psi_x^{\perp}|\psi_y^{\perp}\rangle = 0$; thus, no exchange interaction between the link spinorbitals exists. Both cases introduce an interdependence between the spins of the A–C and B–C link electrons which can be viewed as unphysical—the A–C and B–C partitionings should be independent of each other, leading to a random mutual orientation of spins. To model this random spin coupling, we propose to adopt as the final exchange energy the average of the parallel and perpendicular approaches:

$$E_{\text{exch}}^{(10)}(S^2) = \frac{1}{2}(E_{\text{exch}}^{(10)\parallel}(S^2) + E_{\text{exch}}^{(10)\perp}(S^2)) \quad (4)$$

Following the standard derivation of MO-based exchange corrections within the density-matrix formalism²¹ (valid in both dimer and monomer basis sets) and recasting the resulting formulas to the AO basis,^{22–24} we obtain the following formulas for the parallel-spin and perpendicular-spin variants:

$$\begin{aligned} E_{\text{exch}}^{(10)\parallel/\perp}(S^2) = & -2\mathbf{D}^A \cdot \mathbf{K}^B \mp \frac{1}{2}\mathbf{D}^X \cdot \mathbf{K}^Y - 2\mathbf{V}^A \cdot (\mathbf{D}^A \mathbf{S}^A \mathbf{O} \mathbf{D}^B) \\ & \mp \frac{1}{2}\mathbf{V}^A \cdot (\mathbf{D}^X \mathbf{S}^A \mathbf{O} \mathbf{D}^Y) - 4\mathbf{J}^A \cdot (\mathbf{D}^A \mathbf{S}^A \mathbf{O} \mathbf{D}^B) \mp \mathbf{J}^A \cdot (\mathbf{D}^X \mathbf{S}^A \mathbf{O} \mathbf{D}^Y) \\ & + 2\mathbf{K}^A \cdot (\mathbf{D}^A \mathbf{S}^A \mathbf{O} \mathbf{D}^B) + \frac{1}{2}\mathbf{K}^X \cdot (\mathbf{D}^X \mathbf{S}^A \mathbf{O} \mathbf{D}^B) \pm \frac{1}{2}\mathbf{K}^X \cdot (\mathbf{D}^A \mathbf{S}^A \mathbf{O} \mathbf{D}^Y) \\ & \pm \frac{1}{2}\mathbf{K}^A \cdot (\mathbf{D}^X \mathbf{S}^A \mathbf{O} \mathbf{D}^Y) - 2\mathbf{V}^B \cdot (\mathbf{D}^B \mathbf{S}^B \mathbf{O} \mathbf{D}^A) \mp \frac{1}{2}\mathbf{V}^B \cdot (\mathbf{D}^Y \mathbf{S}^B \mathbf{O} \mathbf{D}^X) \\ & - 4\mathbf{J}^B \cdot (\mathbf{D}^B \mathbf{S}^B \mathbf{O} \mathbf{D}^A) \mp \mathbf{J}^B \cdot (\mathbf{D}^Y \mathbf{S}^B \mathbf{O} \mathbf{D}^X) + 2\mathbf{K}^B \cdot (\mathbf{D}^B \mathbf{S}^B \mathbf{O} \mathbf{D}^A) \\ & + \frac{1}{2}\mathbf{K}^Y \cdot (\mathbf{D}^Y \mathbf{S}^B \mathbf{O} \mathbf{D}^A) \pm \frac{1}{2}\mathbf{K}^Y \cdot (\mathbf{D}^B \mathbf{S}^B \mathbf{O} \mathbf{D}^X) \pm \frac{1}{2}\mathbf{K}^B \cdot (\mathbf{D}^Y \mathbf{S}^B \mathbf{O} \mathbf{D}^X) \\ & + 2\mathbf{V}^A \cdot (\mathbf{D}^B \mathbf{S}^A \mathbf{O} \mathbf{D}^A \mathbf{S}^A \mathbf{O} \mathbf{D}^B) + \frac{1}{2}\mathbf{V}^A \cdot (\mathbf{D}^Y \mathbf{S}^A \mathbf{O} \mathbf{D}^A \mathbf{S}^A \mathbf{O} \mathbf{D}^Y) \\ & \pm \frac{1}{2}\mathbf{V}^A \cdot (\mathbf{D}^Y \mathbf{S}^A \mathbf{O} \mathbf{D}^X \mathbf{S}^A \mathbf{O} \mathbf{D}^B) \pm \frac{1}{2}\mathbf{V}^A \cdot (\mathbf{D}^B \mathbf{S}^A \mathbf{O} \mathbf{D}^X \mathbf{S}^A \mathbf{O} \mathbf{D}^Y) \\ & + 4\mathbf{J}^A \cdot (\mathbf{D}^B \mathbf{S}^A \mathbf{O} \mathbf{D}^A \mathbf{S}^A \mathbf{O} \mathbf{D}^B) + \mathbf{J}^A \cdot (\mathbf{D}^Y \mathbf{S}^A \mathbf{O} \mathbf{D}^A \mathbf{S}^A \mathbf{O} \mathbf{D}^Y) \\ & \pm \mathbf{J}^A \cdot (\mathbf{D}^Y \mathbf{S}^A \mathbf{O} \mathbf{D}^X \mathbf{S}^A \mathbf{O} \mathbf{D}^B) \pm \mathbf{J}^A \cdot (\mathbf{D}^B \mathbf{S}^A \mathbf{O} \mathbf{D}^X \mathbf{S}^A \mathbf{O} \mathbf{D}^Y) \\ & + 2\mathbf{V}^B \cdot (\mathbf{D}^A \mathbf{S}^A \mathbf{O} \mathbf{D}^B \mathbf{S}^A \mathbf{O} \mathbf{D}^A) + \frac{1}{2}\mathbf{V}^B \cdot (\mathbf{D}^X \mathbf{S}^A \mathbf{O} \mathbf{D}^B \mathbf{S}^A \mathbf{O} \mathbf{D}^X) \\ & \pm \frac{1}{2}\mathbf{V}^B \cdot (\mathbf{D}^X \mathbf{S}^A \mathbf{O} \mathbf{D}^Y \mathbf{S}^A \mathbf{O} \mathbf{D}^A) \pm \frac{1}{2}\mathbf{V}^B \cdot (\mathbf{D}^A \mathbf{S}^A \mathbf{O} \mathbf{D}^Y \mathbf{S}^A \mathbf{O} \mathbf{D}^X) \\ & + 4\mathbf{J}^B \cdot (\mathbf{D}^A \mathbf{S}^A \mathbf{O} \mathbf{D}^B \mathbf{S}^A \mathbf{O} \mathbf{D}^A) + \mathbf{J}^B \cdot (\mathbf{D}^X \mathbf{S}^A \mathbf{O} \mathbf{D}^B \mathbf{S}^A \mathbf{O} \mathbf{D}^X) \\ & \pm \mathbf{J}^B \cdot (\mathbf{D}^X \mathbf{S}^A \mathbf{O} \mathbf{D}^Y \mathbf{S}^A \mathbf{O} \mathbf{D}^A) \pm \mathbf{J}^B \cdot (\mathbf{D}^A \mathbf{S}^A \mathbf{O} \mathbf{D}^Y \mathbf{S}^A \mathbf{O} \mathbf{D}^X) \\ & - 2(\mathbf{D}^A \mathbf{S}^A \mathbf{O} \mathbf{D}^B) \cdot \mathbf{K}^\dagger[\mathbf{D}^A \mathbf{S}^A \mathbf{O} \mathbf{D}^B] - \frac{1}{2}(\mathbf{D}^X \mathbf{S}^A \mathbf{O} \mathbf{D}^B) \cdot \mathbf{K}^\dagger[\mathbf{D}^X \mathbf{S}^A \mathbf{O} \mathbf{D}^B] \\ & - \frac{1}{2}(\mathbf{D}^A \mathbf{S}^A \mathbf{O} \mathbf{D}^Y) \cdot \mathbf{K}^\dagger[\mathbf{D}^A \mathbf{S}^A \mathbf{O} \mathbf{D}^Y] \mp \frac{1}{2}(\mathbf{D}^A \mathbf{S}^A \mathbf{O} \mathbf{D}^Y) \cdot \mathbf{K}^\dagger[\mathbf{D}^X \mathbf{S}^A \mathbf{O} \mathbf{D}^B] \\ & \mp \frac{1}{2}(\mathbf{D}^A \mathbf{S}^A \mathbf{O} \mathbf{D}^B) \cdot \mathbf{K}^\dagger[\mathbf{D}^X \mathbf{S}^A \mathbf{O} \mathbf{D}^Y] \mp \frac{1}{2}(\mathbf{D}^X \mathbf{S}^A \mathbf{O} \mathbf{D}^Y) \cdot \mathbf{K}^\dagger[\mathbf{D}^A \mathbf{S}^A \mathbf{O} \mathbf{D}^B] \\ & \mp \frac{1}{2}(\mathbf{D}^X \mathbf{S}^A \mathbf{O} \mathbf{D}^B) \cdot \mathbf{K}^\dagger[\mathbf{D}^A \mathbf{S}^A \mathbf{O} \mathbf{D}^Y] - \frac{1}{8}(\mathbf{D}^X \mathbf{S}^A \mathbf{O} \mathbf{D}^Y) \cdot \mathbf{K}^\dagger[\mathbf{D}^X \mathbf{S}^A \mathbf{O} \mathbf{D}^Y] \end{aligned} \quad (5)$$

In the above equation, the upper plus/minus signs correspond to $E_{\text{exch}}^{(10)\parallel}(S^2)$ and the lower signs to $E_{\text{exch}}^{(10)\perp}(S^2)$. Furthermore, the AO-basis density matrices are computed from the LCAO MO coefficients C_{iK} in the usual way:

$$\begin{aligned}
(\mathbf{D}^A)_{KL} &= \sum_i C_{iK} C_{iL} + \frac{1}{2} C_{xK} C_{xL} \\
(\mathbf{D}^B)_{KL} &= \sum_j C_{jK} C_{jL} + \frac{1}{2} C_{yK} C_{yL} \\
(\mathbf{D}^X)_{KL} &= C_{xK} C_{xL} \\
(\mathbf{D}^Y)_{KL} &= C_{yK} C_{yL}
\end{aligned} \quad (6)$$

where the indices i and j run over occupied orbitals of **A** and **B**, respectively, and indices x, y correspond to the link orbitals χ_x, χ_y (the factors $\frac{1}{2}$ result from the link orbital containing only half an electron of a given spin). Note that $\mathbf{D}^A$ and $\mathbf{D}^B$ already contain the link-orbital contributions $\mathbf{D}^X$ and $\mathbf{D}^Y$, respectively, but the numerical factors multiplying the link-orbital terms need to be adjusted by the presence of the additional terms involving $\mathbf{D}^X$ and $\mathbf{D}^Y$. The elements of the commonly used generalized Coulomb and exchange matrices are given by

$$\mathbf{J}[\mathbf{X}]_{KL} = \sum_{MN} (KLIMN) \mathbf{X}_{MN} \quad \mathbf{K}[\mathbf{X}]_{KL} = \sum_{MN} (KMILN) \mathbf{X}_{MN} \quad (7)$$

the shorthand notations $\mathbf{J}^Z \equiv \mathbf{J}[\mathbf{D}^Z]$ and $\mathbf{K}^Z \equiv \mathbf{K}[\mathbf{D}^Z]$ ($Z = A, B, X, Y$) indicate the regular (not generalized) Coulomb and exchange matrices, $\mathbf{V}^A/\mathbf{V}^B$ are matrices containing the nuclear attraction integrals for an appropriate fragment, $\mathbf{S}^{AO}$ is the AO-basis overlap matrix ($\mathbf{S}^{AO})_{KL} = \langle K|L \rangle$, and the dot signifies an inner product of matrices:

$$\mathbf{X} \cdot \mathbf{Y} = \sum_{KL} \mathbf{X}_{KL} \mathbf{Y}_{KL} \quad (8)$$

The case of the full nonapproximated $E_{\text{exch}}^{(10)}$ is more complicated and will be presented in detail in [Appendix A](#).

Second-Order ISAPT Terms. A more subtle issue is defining a suitable induction energy $E_{\text{ind}}^{(20)}$ (following ref 25, the capitalized correction name denotes a sum of pure induction and exchange-induction terms) and, in a coherent manner, the δ_{HF} correction that captures higher-order induction and exchange-induction effects beyond that specific form of $E_{\text{ind}}^{(20)}$. For these purposes, we note that the single electron on χ_x (χ_y) does contribute (via the electron density) to the electrostatic potential of a given fragment that polarizes the other one. However, the orbitals χ_x (χ_y) themselves are assumed to be frozen (unpolarizable). In other words, we consider the response (polarization) of the occupied orbitals in $\mathcal{B}_A$ under the influence of the electrostatic potential resulting from the nuclei of **B** and the occupied orbitals in $\mathcal{B}_B \oplus (1/2)\{\chi_y\}$ and vice versa, where “(1/2)” reminds that the orbital χ_y is singly occupied while the orbitals spanning $\mathcal{B}_B$ are doubly occupied. A technical issue is the precise choice of the virtual space for the response of fragment-A orbitals. We employed for this virtual space the orthogonal complement of $\mathcal{B}_A \oplus \mathcal{B}_C$ like in the default ISAPT variant, disregarding the fact that the IHO χ_x , carved out of $\mathcal{B}_C$, is not precisely contained in $\mathcal{B}_C$ and has a nonzero component in the unoccupied space.

The coupled perturbed Hartree-Fock (CPHF) coefficients C_a^i/C_b^i for each fragment and the pure induction correction $E_{\text{ind,resp}}^{(20)}$ can now be computed using standard SAPT0 formulas (note that a and b denote virtual orbitals of **A** and **B**, respectively). The exchange-induction correction $E_{\text{exch-ind,resp}}^{(20)}$ computed in ISAPT within the S^2 approximation, again depends on whether a parallel (eq 2) or perpendicular (eq 3) spin coupling of link electrons is assumed, and we will again opt for the average of the two as our final correction. The

resulting AO-based formula for both couplings, derived via the density matrix formalism,²¹ reads

$$\begin{aligned}
E_{\text{exch-ind,resp}}^{(20) \parallel / \perp} (A \leftarrow B, S^2) &= \mathbf{C}^A \cdot [-2\mathbf{K}^B - 4\mathbf{J}^B \mathbf{D}^B \mathbf{S}^{AO} - 2\mathbf{V}^B \mathbf{D}^B \mathbf{S}^{AO} \\
&+ 2\mathbf{K}^B \mathbf{D}^B \mathbf{S}^{AO} + \frac{1}{2} \mathbf{K}^Y \mathbf{D}^Y \mathbf{S}^{AO} - 4\mathbf{J}[\mathbf{D}^A \mathbf{S}^{AO} \mathbf{D}^B] \mp \mathbf{J}[\mathbf{D}^X \mathbf{S}^{AO} \mathbf{D}^Y] \\
&+ 2\mathbf{S}^{AO} \mathbf{D}^B \mathbf{K}^A \pm \frac{1}{2} \mathbf{S}^{AO} \mathbf{D}^Y \mathbf{K}^X + 2\mathbf{K}[\mathbf{D}^A \mathbf{S}^{AO} \mathbf{D}^B] \pm \frac{1}{2} \mathbf{K}[\mathbf{D}^X \mathbf{S}^{AO} \mathbf{D}^Y] \\
&- 4\mathbf{S}^{AO} \mathbf{D}^B \mathbf{J}^A - 2\mathbf{S}^{AO} \mathbf{D}^B \mathbf{V}^A + 4\mathbf{S}^{AO} \mathbf{D}^B \mathbf{S}^{AO} \mathbf{D}^A \mathbf{J}^B + \mathbf{S}^{AO} \mathbf{D}^Y \mathbf{S}^{AO} \mathbf{D}^X \mathbf{J}^B \\
&+ 2\mathbf{S}^{AO} \mathbf{D}^B \mathbf{S}^{AO} \mathbf{D}^A \mathbf{V}^B \pm \frac{1}{2} \mathbf{S}^{AO} \mathbf{D}^Y \mathbf{S}^{AO} \mathbf{D}^X \mathbf{V}^B + 4\mathbf{J}^B \mathbf{D}^A \mathbf{S}^{AO} \mathbf{D}^B \mathbf{S}^{AO} \\
&\pm \mathbf{J}^B \mathbf{D}^X \mathbf{S}^{AO} \mathbf{D}^Y \mathbf{S}^{AO} + 2\mathbf{V}^B \mathbf{D}^A \mathbf{S}^{AO} \mathbf{D}^B \mathbf{S}^{AO} \pm \frac{1}{2} \mathbf{V}^B \mathbf{D}^X \mathbf{S}^{AO} \mathbf{D}^Y \mathbf{S}^{AO} \\
&+ 4\mathbf{J}[\mathbf{D}^B \mathbf{S}^{AO} \mathbf{D}^A \mathbf{S}^{AO} \mathbf{D}^B] \pm \mathbf{J}[\mathbf{D}^Y \mathbf{S}^{AO} \mathbf{D}^X \mathbf{S}^{AO} \mathbf{D}^Y] \pm \mathbf{J}[\mathbf{D}^B \mathbf{S}^{AO} \mathbf{D}^X \mathbf{S}^{AO} \mathbf{D}^Y] \\
&+ \mathbf{J}[\mathbf{D}^Y \mathbf{S}^{AO} \mathbf{D}^A \mathbf{S}^{AO} \mathbf{D}^Y] - 2\mathbf{S}^{AO} \mathbf{D}^B \mathbf{K}[\mathbf{D}^B \mathbf{S}^{AO} \mathbf{D}^A] \\
&\mp \frac{1}{2} \mathbf{S}^{AO} \mathbf{D}^B \mathbf{K}[\mathbf{D}^Y \mathbf{S}^{AO} \mathbf{D}^X] \mp \frac{1}{2} \mathbf{S}^{AO} \mathbf{D}^Y \mathbf{K}[\mathbf{D}^B \mathbf{S}^{AO} \mathbf{D}^X] \\
&- \frac{1}{2} \mathbf{S}^{AO} \mathbf{D}^Y \mathbf{K}[\mathbf{D}^Y \mathbf{S}^{AO} \mathbf{D}^A] - 2\mathbf{K}[\mathbf{D}^A \mathbf{S}^{AO} \mathbf{D}^B] \mathbf{D}^B \mathbf{S}^{AO} \\
&\mp \frac{1}{2} \mathbf{K}[\mathbf{D}^X \mathbf{S}^{AO} \mathbf{D}^B] \mathbf{D}^Y \mathbf{S}^{AO} \mp \frac{1}{2} \mathbf{K}[\mathbf{D}^X \mathbf{S}^{AO} \mathbf{D}^Y] \mathbf{D}^B \mathbf{S}^{AO} \\
&- \frac{1}{2} \mathbf{K}[\mathbf{D}^A \mathbf{S}^{AO} \mathbf{D}^Y] \mathbf{D}^Y \mathbf{S}^{AO} + 4\mathbf{S}^{AO} \mathbf{D}^B \mathbf{J}^A \mathbf{D}^B \mathbf{S}^{AO} + \mathbf{S}^{AO} \mathbf{D}^Y \mathbf{J}^A \mathbf{D}^Y \mathbf{S}^{AO} \\
&+ 2\mathbf{S}^{AO} \mathbf{D}^B \mathbf{V}^A \mathbf{D}^B \mathbf{S}^{AO} + \frac{1}{2} \mathbf{S}^{AO} \mathbf{D}^Y \mathbf{V}^A \mathbf{D}^Y \mathbf{S}^{AO}]
\end{aligned} \quad (9)$$

where $\mathbf{C}^A$ is the matrix of fragment-A CPHF coefficients C_a^i backtransformed to the AO basis using the appropriate LCAO MO coefficients C_{iK} and C_{aL} .²⁶

$$\mathbf{C}_{KL}^A = C_a^i C_{iK} C_{aL} \quad (10)$$

An analogous formula for the other exchange induction term $E_{\text{exch-ind,resp}}^{(20) \parallel / \perp} (B \leftarrow A, S^2)$ (where, this time, fragment **A** polarizes fragment **B**) is obtained from eq 9 by an exchange of all symbols pertaining to **A** by the corresponding ones of **B** and vice versa.

The adjustment to the treatment of second-order exchange-dispersion energy resulting from our link bond reassignment is quite analogous to exchange-induction energy (note that the $E_{\text{disp}}^{(20)}$ term needs no adjustment as excitations from the link orbital are not considered). The resulting AO-based formulas for $E_{\text{exch-disp}}^{(20) \parallel / \perp} (S^2)$ are presented in the [Supporting Information](#). Notably, all additional terms resulting from the link-electron spin coupling cancel out when averaging the parallel and perpendicular cases, so the spin-averaged value of $E_{\text{exch-disp}}^{(20)} (S^2)$ is given by the standard SAPT0 formula (see, e.g., ref 26 for its AO form) without any modifications.

Hartree-Fock Delta Term. Unfortunately, our proposed definition of interfragment induction energy does not lend itself to an infinite-order generalization that can be used to define δ_{HF} . Recall that, in standard intermolecular SAPT, δ_{HF} is defined as the difference between the supermolecular HF interaction energy $E_{\text{int}}^{\text{HF}}$ and its SAPT approximation $E_{\text{elst}}^{(10)} + E_{\text{exch}}^{(10)} + E_{\text{ind,resp}}^{(20)} + E_{\text{exch-ind,resp}}^{(20)}$. Adapted to the ISAPT context, $E_{\text{int}}^{\text{HF}}$ still signifies the difference in the HF energies of the **A–B** complex with and without interaction. In our ISAPT algorithms, on the noninteracting fragment side, the induction energy accounts for a polarization of orbitals from $\mathcal{B}_A$ in the electrostatic potential of $\mathcal{B}_B \oplus (1/2)\{\chi_y\}$ and the other way around. Accordingly, one should optimize the noninteracting occupied orbitals for **A** in the orthogonal complement of $\mathcal{B}_C$, with the embedding potential including interactions with

nuclei of C and the sum of electron densities coming from χ_x and the rest of $\mathcal{B}_C$ (recall that the latter has had two electrons removed and reassigned to A and B). Note that the embedding potential does not contain the interaction with the single-electron density from χ_y , the IHO at the more distant fragment boundary, which has been reassigned to fragment B; thus, the entire A–B interaction has been temporarily switched off.

On the interacting side, one optimizes the occupied orbitals of AB also in the orthogonal complement of $\mathcal{B}_C$, with the nuclei of C and the Coulomb and exchange operators corresponding to the density of $\mathcal{B}_C$ forming the HF embedding potential. Note that this potential is obtained from the same HF calculation for the entire molecule. Moreover, the reassignment of electrons on χ_x and χ_y does not matter this time because the HF system is embedded in the sum of densities coming from χ_x , χ_y , and the rest of $\mathcal{B}_C$. The orbitals χ_x and χ_y , while frozen in the HF optimization, do belong to the fragments A and B, so that their interaction with each other and with the nuclei of the other fragment needs to be included in the HF interaction energy (it is a constant term that does not depend on the orbitals being optimized). Unfortunately, in this algorithm, the treatment of the two electrons assigned to χ_x and χ_y is not consistent between the fragment calculations, where they are described by frozen orbitals, and the molecular one, where they belong to doubly occupied orbitals which are variationally optimized. As a result, we were not able to define a supermolecular δ_{HF} correction that is consistent with the proposed reassignment of link bond electrons and the resulting $E_{\text{elst}}^{(10)}$, $E_{\text{exch}}^{(10)}$, $E_{\text{ind,resp}}^{(20)}$, and $E_{\text{exch-ind,resp}}^{(20)}$ corrections. In the calculations below, the δ_{HF} term will be taken from ISAPT calculations with the default assignment of the entire link orbitals to C.

Embedding, Orthogonalization, and Self-Consistency. We note at this point that the embedding defined in the fragment HF calculations described above (omitting the IHO at the distant fragment boundary), while perfectly consistent with our electrostatic and induction energy expressions, is not the same as the embedding used to compute the initial HF orbitals spanning $\mathcal{B}_A$ and $\mathcal{B}_B$ (which, by default, always takes into account the entire linker including both χ_x and χ_y). If the linker is sufficiently large, the difference in the embedding density occurs far away from the fragment in question and its effect should be small; however, we have to consider whether an orbital reoptimization is in order before the computation of ISAPT corrections. On the other hand, χ_x and χ_y depend on the spaces $\mathcal{B}_A$ and $\mathcal{B}_B$ because they are orthogonalized to these spaces. This circular dependence suggests that iteration to self-consistency is the most rigorous way to implement the process: determine χ_x and χ_y which are orthogonal to the initial spaces $\mathcal{B}_A$ and $\mathcal{B}_B$ from the default link assignment algorithm, reoptimize $\mathcal{B}_A$ in an embedding excluding χ_y and $\mathcal{B}_B$ in an embedding excluding χ_x , update χ_x/χ_y by enforcing orthogonality to new $\mathcal{B}_A/\mathcal{B}_B$, and so on. However, a fully self-consistent algorithm is impractical—each iteration requires converged HF calculations for both fragments. On the other hand, as differences in embedding should be minor and occur relatively far away from a given fragment, this process should be very quickly convergent, and we can limit ourselves to zero, one, or two iterations. This leads to three variants of our approach:

Zero iterations: the occupied spaces $\mathcal{B}_A/\mathcal{B}_B$ are taken directly from the default ISAPT variant, that is, the HF orbitals

have been obtained using an embedding in the entire $\mathcal{B}_C$ space. These occupied spaces are used to construct χ_x and χ_y , and the ISAPT corrections are obtained from the resulting augmented density matrices and the polarization of the occupied spaces.

One iteration: after the IHOs χ_x and χ_y are constructed as above, the HF calculations for fragments are repeated, this time with the embedding potential for A excluding χ_y and the embedding potential for B excluding χ_x . This leads to new occupied spaces $\mathcal{B}'_A/\mathcal{B}'_B$, which in turn results in updated IHOs χ'_x and χ'_y orthogonal to the new occupied spaces. Now, all ISAPT corrections are obtained from the resulting augmented density matrices and the polarization of the occupied spaces $\mathcal{B}'_A/\mathcal{B}'_B$.

Two iterations: after the IHOs χ'_x and χ'_y are constructed as above, the HF calculations for fragments are repeated once again, this time with the embedding potential for A excluding χ'_y and the embedding potential for B excluding χ'_x . This leads to new occupied spaces $\mathcal{B}''_A/\mathcal{B}''_B$, which in turn results in updated IHOs χ''_x and χ''_y orthogonal to the new occupied spaces. Now, all ISAPT corrections are obtained as above, but using the occupied spaces $\mathcal{B}''_A/\mathcal{B}''_B$ and the IHOs χ''_x and χ''_y .

Determining Link Hybrid Orbitals: the SAO and SIAO Approaches. To complete the specification of our method, we have to define how the IHOs χ_x and χ_y are determined. We propose two algorithms for this, both of which proceed with a projection step (onto the fragment A or B) and an orthogonalization step (relative to all doubly occupied orbitals of this fragment). Once the specific IBO ϕ_{AC} responsible for the A–C link bond is identified, its projection onto A can be performed in one of two ways. In the first approach, inspired by the ALMO method,²⁷ ϕ_{AC} is represented in the AO basis and projected onto the space of basis functions centered on atoms of A only (all other coefficients of this orbital are zeroed). This approach will be termed Splitting of Atomic Orbitals (SAO). In the second approach, ϕ_{AC} is represented in the IAO basis,¹⁷ projected onto the space of IAOs centered on atoms of A only by zeroing all other coefficients, and transformed back to the AO basis. This approach will be termed Splitting of Intrinsic Atomic Orbitals (SIAO). In either method, the final IHO χ_x is obtained from the projected orbital by a Schmidt orthonormalization to all doubly occupied orbitals of A, that is, to the $\mathcal{B}_A$ space. Note that the SIAO projection is not strictly an ALMO (it has nonzero coefficients of basis functions on the other fragments resulting from the tails of the IAOs), and the SAO projection ceases to be an ALMO after orthogonalization. However, in either case, χ_x is still mostly localized on A. The same projection-then-orthogonalization procedure is performed at the other interfragment boundary to determine χ_y .

The final specification of the new ISAPT algorithm includes both the SAO/SIAO recipe to determine link IHOs and the level of self-consistency between those IHOs and the occupied orbital space expressed by the number of iterations described earlier in this section. For example, the SAO0 method involves the SAO projection algorithm resulting in χ_x/χ_y orthogonalized to the original occupied spaces $\mathcal{B}_A/\mathcal{B}_B$ (zero iterations), while the SIAO2 one uses the SIAO projection leading to χ''_x/χ''_y orthogonalized to the twice updated occupied spaces $\mathcal{B}''_A/\mathcal{B}''_B$. While the orthogonalization step is formally required to obtain a valid fragment density matrix when augmented by χ_x/χ_y , we will perform a limited set of numerical tests to investigate the practical consequences of not orthogonalizing (note that the

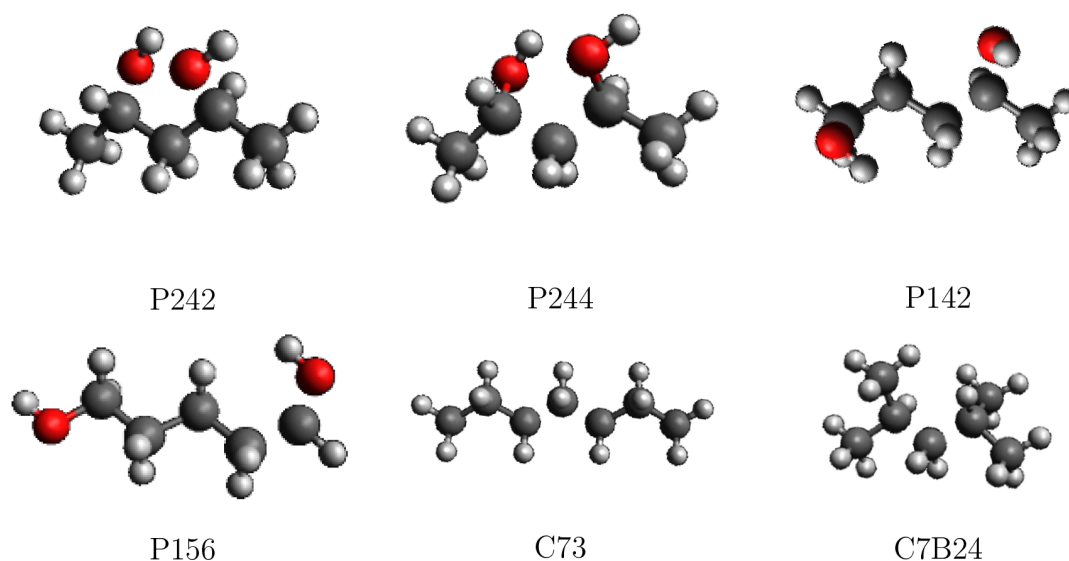


Figure 2. Some illustrative fragmentation patterns of 2,4-, 1,4-, and 1,5-pentandiol, *n*-heptane, and 2,4-dimethylpentane considered in this work. A missing bond signifies a fragment boundary.

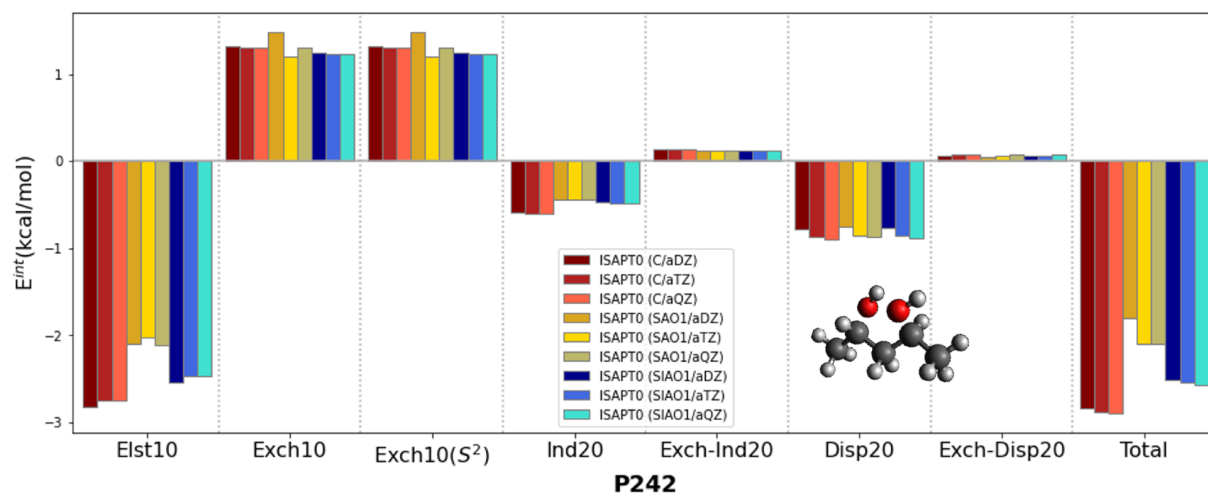


Figure 3. ISAPT energy components in 2,4-pentandiol (P242 fragmentation pattern) computed using the aXZ bases ($X = D, T, Q$) and three link assignments: C, SAO1, and SIAO1.

lack of orthogonalization makes the algorithms with 1 and 2 iterations identical as $\chi_X \equiv \chi'_X \equiv \chi''_X$ for $X = x, y$).

Finally, we note a common formal flaw of both selections of the IHOs χ_x and χ_y : they are not entirely contained in the occupied space $\mathcal{B}_C$. As a result, from the point of view of the linker C, the reassignment of one electron each to χ_x and χ_y involves subtracting some electron density that is not present in the system, leaving out a residual density matrix for C that is not positive definite (two negative eigenvalues). We do not see a simple way to avoid this flaw; however, it should be inconsequential in practice as only the density matrices of A and B are used to compute the ISAPT corrections and, for the purpose of recomputing the occupied orbital spaces $\mathcal{B}_A$ and $\mathcal{B}_B$, the repartitioning only changes the embedding potential at the more distant interfragment boundary.

RESULTS AND DISCUSSION

Our improved ISAPT algorithms have been implemented in a development version of the Psi4 software package.²⁸ The

modified Psi4 code is available on GitHub at <https://github.com/konpat/psi4/tree/isapt>. The interaction energy components within a number of representative molecules are analyzed and compared to the previously proposed ISAPT method by Parrish et al.¹⁴ In particular, pentandiol isomers, *n*-heptane, 2,4-dimethylpentane, and folded and unfolded bicyclic N-arylimide based molecular balances (halogens Cl, Br, and I interact noncovalently with aromatic fragments benzene, phenanthrene, pyrene, and ethylene)²⁹ are chosen for the study. All geometries are optimized at the MP2 level of theory employing the aug-cc-pVDZ basis set. To examine the accuracy of energy contributions, single-point ISAPT energy calculations are performed using three orbital bases, aug-cc-pVDZ, aug-cc-pVTZ, and aug-cc-pVQZ (aDZ, aTZ, aQZ for short). For bulky molecules, the N-arylimide balances, the systems are optimized using the B3LYP-D3(BJ) method³⁰ with the aDZ basis set (with cc-pVDZ-pp on the heavy iodine atom). The resulting ISAPT interaction energies are computed using the aDZ basis set, with cc-pVTZ centered on the iodine atom. Overall, each molecular system is tested with seven link

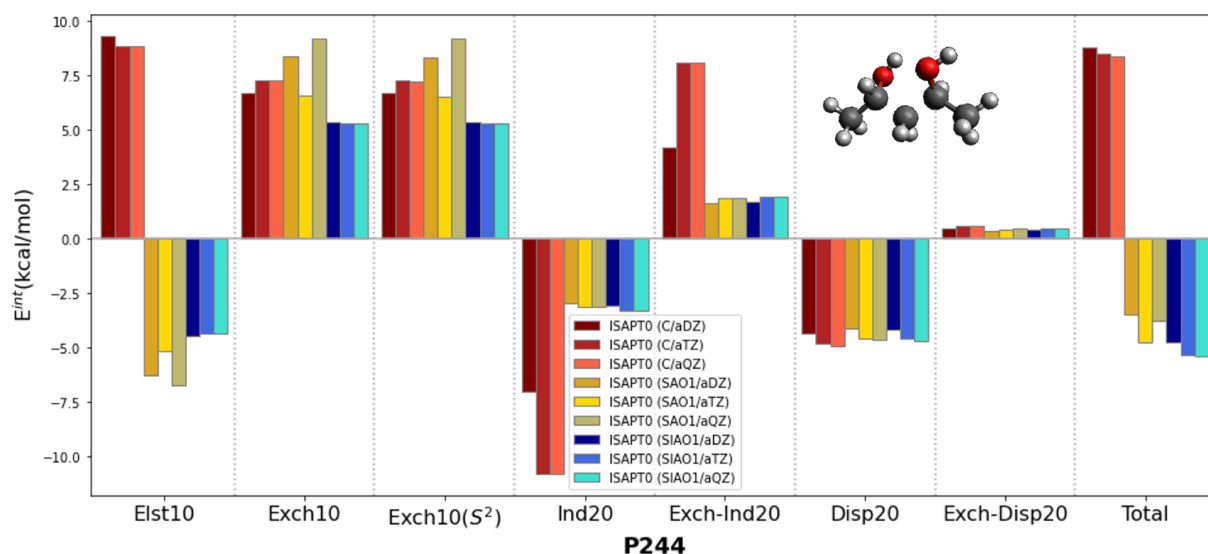


Figure 4. ISAPT energy components in 2,4-pentanediol (P244 fragmentation pattern) computed using the aXZ bases (X = D, T, Q) and three link assignments: C, SAO1, and SIAO1.

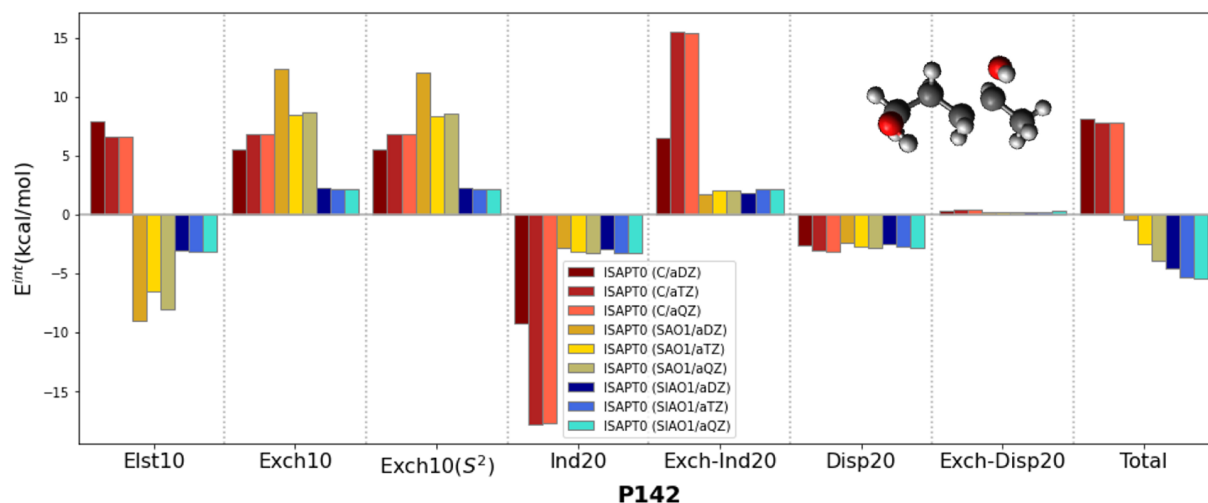


Figure 5. ISAPT energy components in 1,4-pentanediol (P142 fragmentation pattern) computed using the aXZ bases (X = D, T, Q) and three link assignments: C, SAO1, and SIAO1.

assignment options: C (original), SAO0, SAO1, SAO2, SIAO0, SIAO1, and SIAO2, with and without orthogonalizing the link orbitals to the fragment occupied space.

Pentanediol Isomers. The 2,4-, 1,4-, and 1,5-pentanediol molecules, featuring the OH...O distances ranging from 2.40 Å (an intramolecular hydrogen bond) to 6.30 Å, serve as convenient illustrative examples for a range of polar non-covalent intramolecular interactions. Four, six, and nine fragmentation patterns of the 2,4-, 1,4-, and 1,5-pentanediol systems, respectively, were studied. A few examples (P242, P244, P142, and P156) of these fragmentation patterns are presented in Figure 2; the remaining ones are shown in the Supporting Information.

The P242 model, well handled by the original ISAPT(C) approach, assists in the verification of the implemented methods by comparing to the data in ref 14. According to Figure 3, the ISAPT energy components computed by the SIAO1 method have the lowest absolute differences relative to the original C variant. The electrostatic term has a MAD, mean

absolute difference, of 0.28 kcal/mol with bases aDZ, aTZ, and aQZ. This MAD amounts to 0.70 kcal/mol for the SAO1 method. Other energy contributions are also in good agreement between different fragmentation patterns: the SAO1 and SIAO1 induction energies deviate from the C variant by 0.12–0.17 kcal/mol, the first-order exchange term varies by about 0.07 kcal/mol, and differences in the dispersion term do not exceed 0.02 kcal/mol, much smaller than the differences between basis sets.

On the other hand, for most of the other fragmentation patterns, including the minimal linker arrangement of 2,4-pentanediol (P244) as well as related short-linker configurations for other isomers (P142 and P156), the electrostatic energy in the original ISAPT(C) variant is strongly repulsive. This does not make physical sense, as one expects either strong electrostatic attraction due to favorable dipole–dipole interaction (P244) or weak attraction due to charge penetration (P142 and P156). As stated before, we hypothesize that the observed problematic behavior of

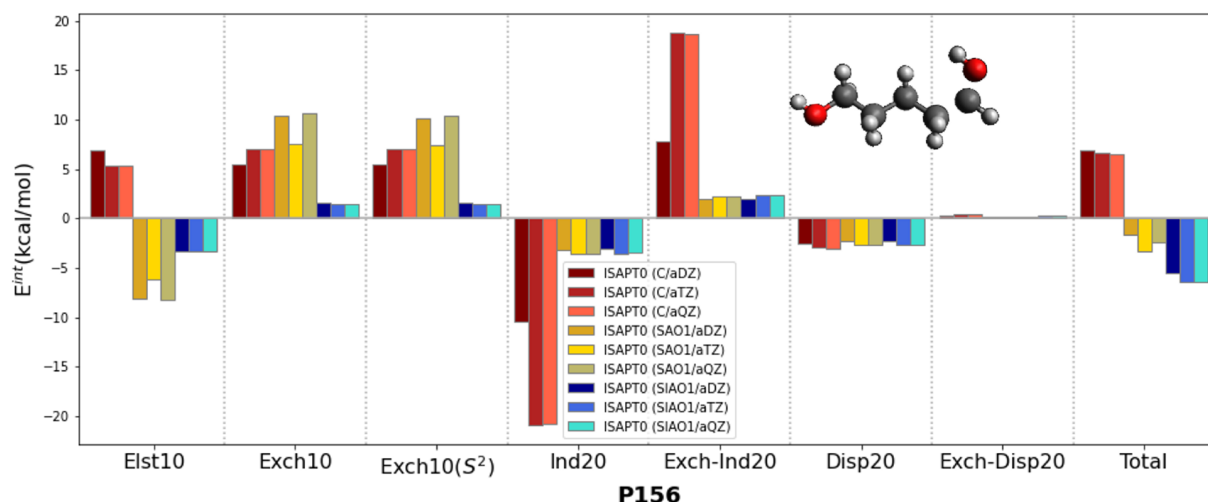


Figure 6. ISAPT energy components in 1,5-pentandiol (P156 fragmentation pattern) computed using the aXZ bases (X = D, T, Q) and three link assignments: C, SAO1, and SIAO1.

Table 1. Magnitude of the HF Dipole Moments (in a.u.) for Different Molecules and Their Fragments in the aDZ, aTZ, and aQZ Bases

System	Basis	ABC	Fragment A			Fragment B		
			C	SAO1	SIAO1	C	SAO1	SIAO1
P242	aDZ	1.057	0.951	0.670	0.716	0.936	0.666	0.691
	aTZ	1.056	0.950	0.644	0.714	0.935	0.633	0.689
	aQZ	1.056	0.950	0.657	0.713	0.935	0.630	0.689
P244	aDZ	1.057	0.833	0.699	0.688	1.312	0.720	0.922
	aTZ	1.056	0.833	0.704	0.687	1.313	0.765	0.921
	aQZ	1.056	0.833	0.744	0.687	1.313	0.740	0.921
P142	aDZ	1.040	1.319	0.676	0.915	0.893	0.625	0.697
	aTZ	1.040	1.319	0.680	0.913	0.892	0.611	0.694
	aQZ	1.039	1.319	0.638	0.913	0.892	0.628	0.694
P156	aDZ	0.959	1.186	0.630	0.893	0.893	0.625	0.705
	aTZ	0.958	1.185	0.667	0.889	0.892	0.617	0.702
	aQZ	0.957	1.184	0.604	0.889	0.892	0.612	0.701
C73	aDZ	0.047	0.808	0.095	0.389	0.813	0.097	0.392
	aTZ	0.046	0.810	0.048	0.388	0.815	0.056	0.391
	aQZ	0.046	0.810	0.110	0.388	0.815	0.126	0.391
C75	aDZ	0.047	0.788	0.115	0.339	0.813	0.080	0.352
	aTZ	0.046	0.790	0.047	0.339	0.815	0.037	0.351
	aQZ	0.046	0.790	0.041	0.339	0.815	0.117	0.351
C7B24	aDZ	0.045	0.799	0.091	0.385	0.799	0.091	0.385
	aTZ	0.044	0.800	0.178	0.384	0.800	0.177	0.384
	aQZ	0.044	0.800	0.100	0.384	0.800	0.100	0.384

ISAPT(C) electrostatics is connected to the unphysical dipole moments emerging at the A-C and B-C fragment boundaries. The results with the new ISAPT link assignments (Figures 4–6) indicate that the SAO1 and SIAO1 electrostatic energy has shifted to a physically justified negative value. Furthermore, while the second-order induction and exchange-induction components in original ISAPT(C) are both very large in magnitude and cancel each other to a large extent, the respective SAO1 and SIAO1 terms are much smaller. The large cancellation between induction and exchange-induction effects is well known from conventional intermolecular SAPT (and is related to the overall divergence of the perturbation series^{31,32}): the unsymmetrized (polarization) expansion allows the two molecules to overpolarize each other, an effect that has to be cancelled by the exchange terms brought about

by enforcing the full permutational symmetry of the wavefunction. The observation that, in ISAPT, such an overpolarization is made particularly egregious by the original C link assignment, once again suggests the unphysical dipole moments at interfragment boundaries, leading to an artificially enhanced electric field, as the culprit. As far as other ISAPT energy terms are concerned, the first-order exchange is relatively consistent between the C and SAO1 schemes but reduced in the SIAO1 one while the dispersion and exchange-dispersion energies are remarkably consistent across all three link assignments.

Figures 4–6 show that the link reassignment proposed in this work significantly improves the ISAPT energy components that are problematic in original ISAPT(C). We postulate that this improvement stems from a substantial reduction of the

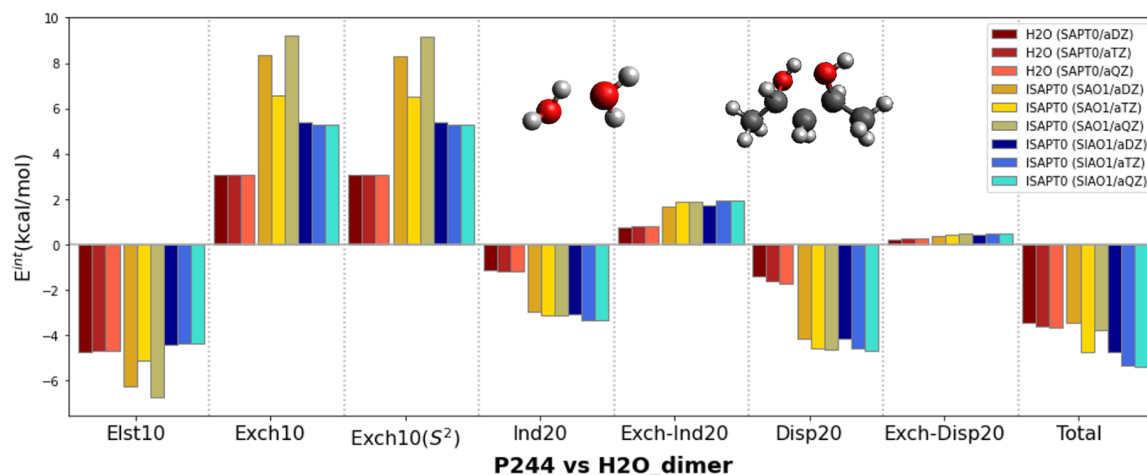


Figure 7. Comparison of energy components between the P244 fragmentation pattern of 2,4-pentandiol (computed with ISAPT/SAO1 and ISAPT/SIAO1) and the corresponding water dimer structure exhibiting the same hydrogen-bonded arrangement (computed with standard SAPT0).

artificial dipoles arising from the nonsymmetric carbon atoms next to the A-C and B-C linking bonds. Indeed, Table 1 shows that the SAO1 and SIAO1 dipole moments on fragments A and B, for different pentanediol isomers and fragmentation patterns, are about 15–20% smaller in magnitude than the fragment dipole moments from the original ISAPT0, and are consistent with respect to the basis set. As both fragments are obviously polar, we do not expect a reduction of dipole moments to near-zero in pentanediol molecules. However, it appears that the redistribution of link bonds to the corresponding IHOs on each fragment has crucially reduced the imbalance of the electron densities on the linking atoms, leading to more meaningful interaction energies.

Overall, with the new SAO1 and SIAO1 variants, the dominant attractive contributions to the intramolecular interaction energy in 2,4-pentandiol are electrostatics and dispersion as expected for a hydrogen-bonded system. The overall interaction energy for the P244 fragmentation pattern, Figure 4, is attractive as expected, amounting to -3.48 kcal/mol for SAO1/aDZ and -4.75 kcal/mol for SIAO1/aDZ, improving from the repulsive 8.80 kcal/mol value obtained with original ISAPT0(C)/aDZ. The SAO1 and SIAO1 variants predict that the fragments in P244 exhibit significantly stronger bonding than the fragments in P242 (Figure 3), which is consistent with the addition of aliphatic chains that amplify the attractive dispersion and charge penetration effects. Both models can also be compared with an analogous hydrogen-bonded configuration of the water dimer, involving the same OH...OH geometry with the remaining hydrogens added in the direction of the O-C bonds in 2,4-pentandiol at MP2/aDZ-optimized distances (Figure 7). One can see that the attractive electrostatic energy obtained from the intramolecular SIAO1 method is in good agreement with the water dimer value computed with standard intermolecular SAPT0. This agreement is accidentally too good: both electrostatic terms primarily stem from the dipole–dipole interaction of the very similar polar fragments, but 2,4-pentandiol includes more charge penetration due to the hydrocarbon chains. Indeed, the presence of the additional aliphatic chains in P244 increases both the interfragment density overlap (thus leading to larger first-order exchange) and the polarizability (enhancing the induction and dispersion terms).

When the basis set is enlarged from aDZ to aQZ, the ISAPT(SIAO1) energy components are highly consistent. As expected, the dispersion and exchange-dispersion corrections increase in magnitude as the basis set increases while other, uncorrelated corrections change very little. This observation is in perfect agreement with standard intermolecular SAPT, where the slow basis set convergence of dispersion and exchange-dispersion corrections is well documented and the possible remedies include midbond functions³³ and the explicitly correlated F12 approach.³⁴ Somewhat disappointingly, the same stable convergence pattern does not apply to the SAO1 link assignment. Conversely, Figures 3–6 indicate that the basis set convergence of the ISAPT(SAO1) electrostatic and first-order exchange terms is slow and erratic, and these convergence issues carry on to the total interaction energies. This ISAPT(SAO1) basis set instability is the main reason why we recommend the SIAO1 variant, which never suffers from such convergence issues, for all practical applications.

Some insights into the different basis set behavior of ISAPT(SAO1) and ISAPT(SIAO1) can be obtained by examining the corresponding link orbitals (IHOs). An example comparison of the link IBOs with the IHOs resulting from both schemes, for the *n*-heptane molecule and its fragmentation pattern later referred to as C74, is presented in Figure 8. For this example, we note that both schemes lead to IHOs that roughly resemble the carbon sp^3 hybrid orbitals pointing toward the other fragment; however, the SIAO1 orbitals appear to be more localized to the respective A/B fragments (note that all panels in Figure 8 use the same isosurface value). This would suggest that the two IHOs χ'_x and χ'_y should have smaller overlap for SIAO1 than for SAO1 (as before, the primes signify that one iteration towards self-consistency of orbital spaces and link IHOs has been performed within the SAO1/SIAO1 schemes). Indeed, we observe that the $\langle \chi'_x | \chi'_y \rangle$ overlap integral tends to be much smaller in magnitude for SIAO1 than for SAO1. For example, in the 2,4-pentandiol system, the P244 fragmentation scheme, and the aDZ, aTZ, and aQZ basis sets, $|\langle \chi'_x | \chi'_y \rangle|$ amounts to 0.133, 0.0632, and 0.0724, respectively, for SAO1 and 3.02×10^{-4} , 3.09×10^{-4} , and 3.11×10^{-4} , respectively, for SIAO1. Formally, neither

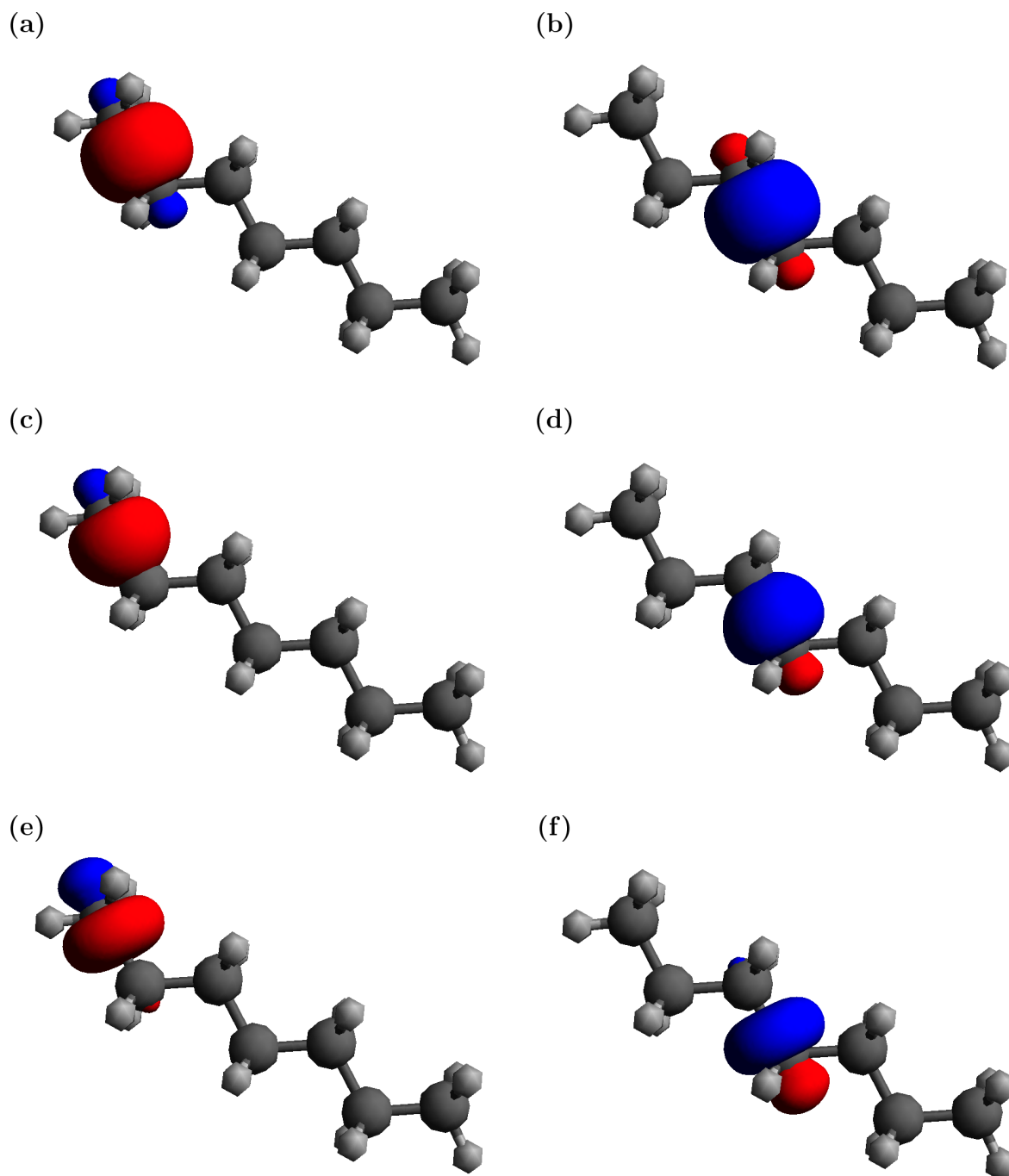


Figure 8. Orbital contours (at the same isosurface value of 0.05) of the complete link IBOs from ISAPT(C) (panels (a) and (b)) and the reassigned link IHOs from ISAPT(SAO1) (panels (c) and (d)) and from ISAPT(SIAO1) (panels (e) and (f)). The system is *n*-heptane in the C74 fragmentation pattern, and the basis set is aDZ.

SIAO1 nor SAO1 link hybrids are strictly localized on the A/B fragments: the IAOs have small tails on other atoms, and the orthogonalization does not strictly preserve the localization. The reason why the SIAO1 link hybrids are more strongly localized than the SAO1 ones is that the minimal basis used to construct IAOs¹⁷ does not give them the flexibility to improve the basis set description on atoms other than their center. On the contrary, in a sufficiently large AO basis, the functions centered on one atom actively improve the basis set description around other atoms, although this improvement varies erratically from basis to basis. As a result, the SAO

projection scheme leaves out larger tails than the SIAO one, but these tails are not stable with the basis set. In our opinion, this behavior leads to both a larger $|\langle \chi'_x | \chi'_y \rangle|$ overlap and a larger basis set instability for the SAO1 link assignment scheme relative to the SIAO1 one.

As the OH...O interacting distance increases when going from 2,4- to 1,4- and 1,5-pentanediols, the electrostatic contribution in a primarily dipole–dipole interaction decreases. Among the 1,4-pentanediol models, the P142 partitioning reports the most favorable electrostatic energy (Figure 5), which amounts to -3.04 kcal/mol for SIAO1 and

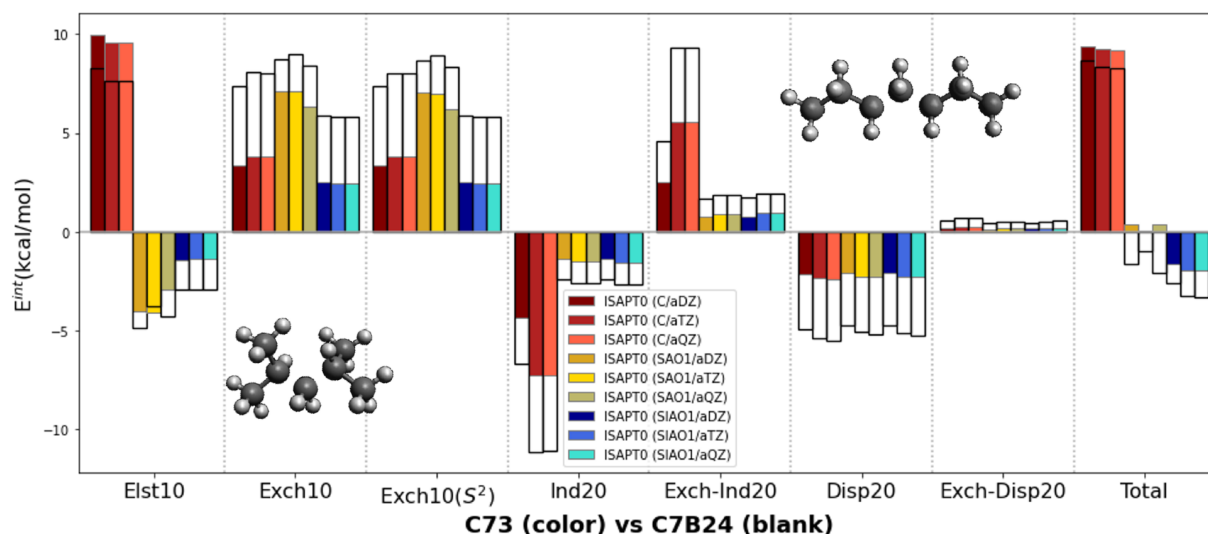


Figure 9. Comparison of the ISAPT0 energy components between *n*-heptane (C73, colored bars) and its branched isomer, 2,4-dimethylpentane (C7B24, empty bars) computed with the C, SAO1, and SIAO1 variants.

−9.01 kcal/mol for SAO1 using the aDZ basis. A notable reduction in the induction term compared to the original ISAPT(C) variant occurs, leading to a net interaction energy of −4.50 kcal/mol and −0.38 kcal/mol in SIAO1 and SAO1, respectively. Interestingly, a similar trend is observed in the P156 partitioning of the 1,5-pentanediol system (Figure 6), yet the electrostatic term can more likely be attributed to charge penetration in this case. The 1,5-pentanediol series shows that the fragments are mainly bound by the electrostatic force, along with a roughly equal contribution from the induction and dispersion terms that sum to the net attractive energy.

Alkanes. Due to the nonpolarity of the C–C bond and the negligible electronegativity difference involved in the C–H bonding, intramolecular interactions in hydrocarbons are dominated by London dispersion forces. Alkanes are excellent models for analyzing nonbonded intramolecular interactions, as it is those interactions that result in a higher thermodynamic stability of branched alkanes relative to their linear isomers.³⁵ Thus, linear and branched seven-carbon alkanes, *n*-heptane and 2,4-dimethylpentane (structures C73 and C7B24 in Figure 2, respectively) are chosen as models to study the performance of different ISAPT variants. It should be stressed that a single fragmentation of these alkanes is not enough to quantify the entire nonbonded intramolecular interaction, as some of the relevant interacting groups inevitably end up on the same fragment. However, the fragmentation patterns in C73 and C7B24 are designed to alleviate this issue as, in both cases, fragments A and B are propyl groups (1-propyl or 2-propyl). Thus, the A–B interaction misses the important 1,3-methyl-methyl stabilizing effects (protobranching³⁶) within each propyl group, but it misses the same number of such effects in both cases.

Both *n*-heptane and 2,4-dimethylpentane are nonpolar molecules, and their fragments obtained by cutting through C–C bonds should be nonpolar as well. However, as shown in Table 1, when the entire link IBO is assigned to fragment C, the noncovalently interacting fragments A and B acquire large dipole moments. As stated above, the origin of those unphysical dipole moments is the unbalanced charge distribution around the linking carbon atoms, where the electrons occupying only three out of the four sp^3 orbitals

belong to the fragment A(B). Fortunately, according to Table 1, the magnitude of the fragment dipole moments is strongly reduced in the new approaches, by 85–90% for SAO1 and 50–55% with SIAO1. In the SIAO1 case, this reduction is perfectly consistent in all tested basis sets, while the SAO1 dipole moments show some basis set fluctuations in line with the energy components. Thus, while the link bond reassignment proposed here does not completely eliminate the spurious multipoles at the interfragment boundary, it reduces their magnitude significantly. It is not clear why the SAO1 approach tends to provide a greater dipole moment reduction than the SIAO1 one.

We now turn to examining the ISAPT energy contributions for *n*-heptane and 2,4-dimethylpentane as a function of the link assignment, focusing on the difference between two systems which may shed light on the physical origins of the branched alkane stabilization. A glimpse at both sets of data (Figure 9) once again shows that the default ISAPT(C) variant does not provide physically meaningful results: the electrostatic energy is strongly repulsive while the induction and exchange-induction terms are very large in magnitude and strongly vary between aDZ and aTZ. The SAO1 link reassignment once again alleviates the issue of repulsive electrostatics. However, besides a fairly slow (but not terrible) basis set convergence, one sees another troubling property that puts the usefulness of ISAPT(SAO1) into doubt: the total C73 interaction energy is slightly repulsive in all tested basis sets. The SIAO1 link reassignment eliminates both of those issues, providing meaningful and quickly convergent ISAPT energy contributions and predicting a stronger interfragment stabilization for the branched isomer relative to the linear one. Thus, only the ISAPT(SIAO1) variant is suitable for providing meaningful insights into the nonbonded interactions enhancing the thermodynamic stability of 2,4-dimethylpentane relative to *n*-heptane.

The ISAPT(SIAO1) data in Figure 9 indicate that the overall attractive interaction energy is predominantly controlled by the dispersion contribution and secondarily by the electrostatic term, but the latter effect is more than counterbalanced by the repulsive first-order exchange. The 2,4-dimethylpentane system exhibits roughly twice the intra-

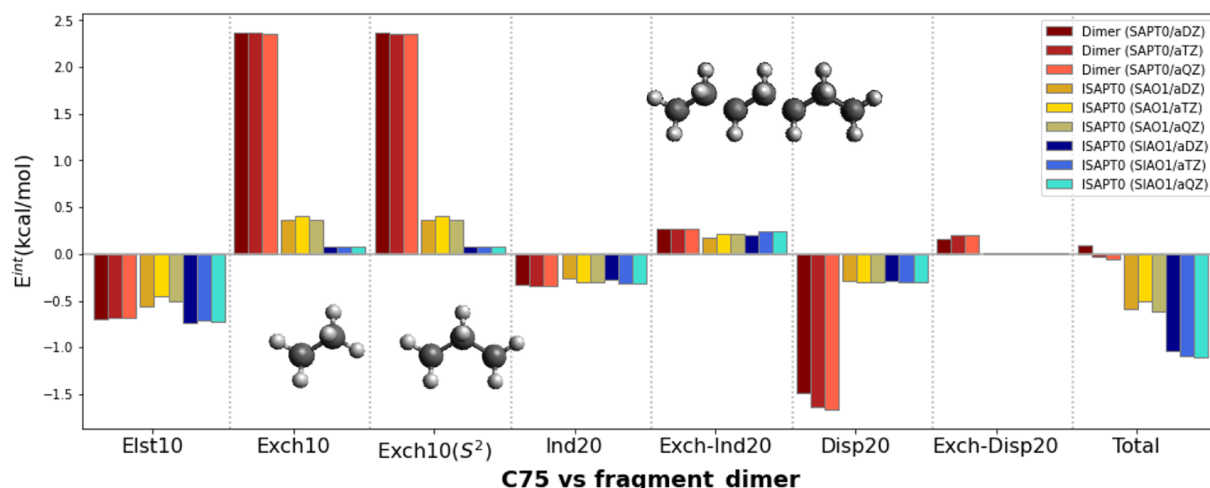


Figure 10. Comparison of energy components between the C75 model of *n*-heptane (computed with the ISAPT0/SAO1 and ISAPT0/SIAO1 variants) and the related intermolecular propane-ethane dimer (computed with standard SAPT0) in the aXZ bases, X = D, T, Q. The locations of two additional hydrogens in the dimer calculation are explicitly optimized at the MP2/aDZ level.

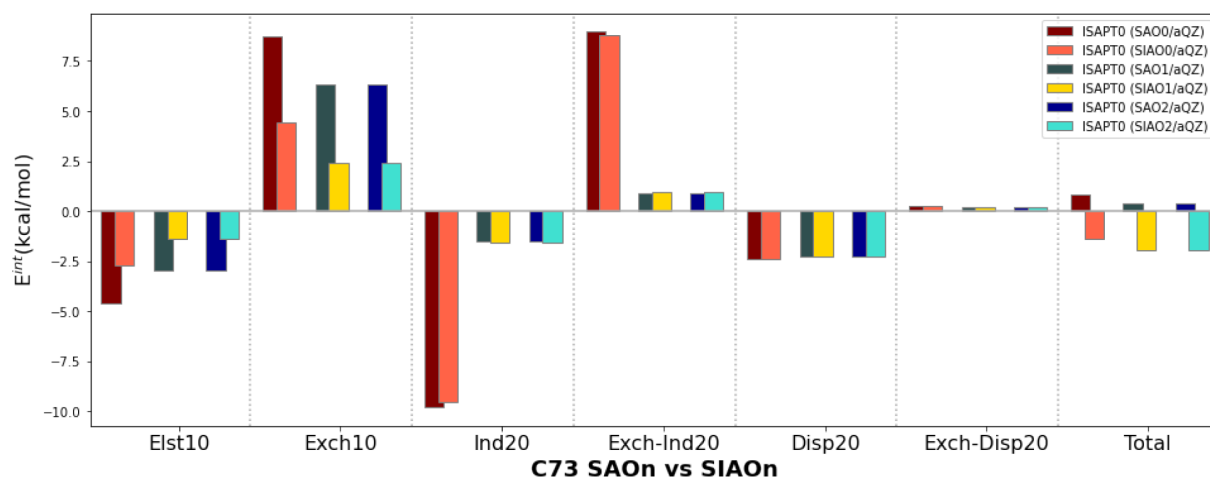


Figure 11. Comparison of interaction energies in the C73 model of *n*-heptane, computed using the aQZ basis set with six link assignment options SAOn and SIAOn, *n* = 0, 1, 2.

molecular electrostatic and dispersion energy of the *n*-heptane model (−2.95 and −4.76 kcal/mol versus −1.43 and −2.09 kcal/mol, respectively, in the aDZ basis set). This illustrates why branched alkanes are more thermodynamically stable than linear alkanes with the same carbon content³⁵—the contribution to the fragments' electron densities from multiple closely spaced methyl groups enhances the charge penetration term in the electrostatic energy as well as leads to a larger dispersion energy. Obviously, the larger interfragment overlap in the branched system increases the first-order exchange repulsion as well, and the overall change in interaction energy is the net result of the additional stabilizing terms in electrostatics and dispersion and destabilizing exchange contributions.

In addition, the intramolecular energy contributions in *n*-heptane can be compared to the standard SAPT terms for a related intermolecular interaction involving fragments A and B capped with hydrogen atoms. The C73 model with the small −CH₂− linking fragment is not suitable for this purpose as the capping hydrogens would end up too close to each other. Therefore, we switch to the C75 model with a slightly bigger −CH₂CH₂− linking fragment. For this model, the interacting

fragments A and B with an additional hydrogen on each are treated with the standard SAPT0 method in the same aDZ–aQZ bases. The resulting energy contributions are presented in Figure 10, showing that the intermolecular SAPT0 electrostatic energy well matches with the improved ISAPT0 variants, especially SIAO1. The close distance between the two capping hydrogens results in larger exchange and dispersion energies observed in the intermolecular system as expected. This attribution of increased first-order exchange and dispersion can be confirmed by an intermolecular F-SAPT calculation,³⁷ where each SAPT term is partitioned into contributions from a pair of fragments, the capping hydrogen and the rest of the molecule. In the F-SAPT calculation in the aTZ basis, the first-order exchange interaction energy of the added hydrogens and the partner monomer amounts to 1.71 kcal/mol, that is, 72% of the total dimer exchange energy of 2.36 kcal/mol. The uncapped fragments account for only 0.65 kcal/mol (28%) of exchange and −0.69 kcal/mol (48%) of dispersion energy, much closer to the intramolecular SAPT results. Overall, the improved ISAPT methods provide reasonable and reliable energy decomposition unlike the original ISAPT0(C) variant

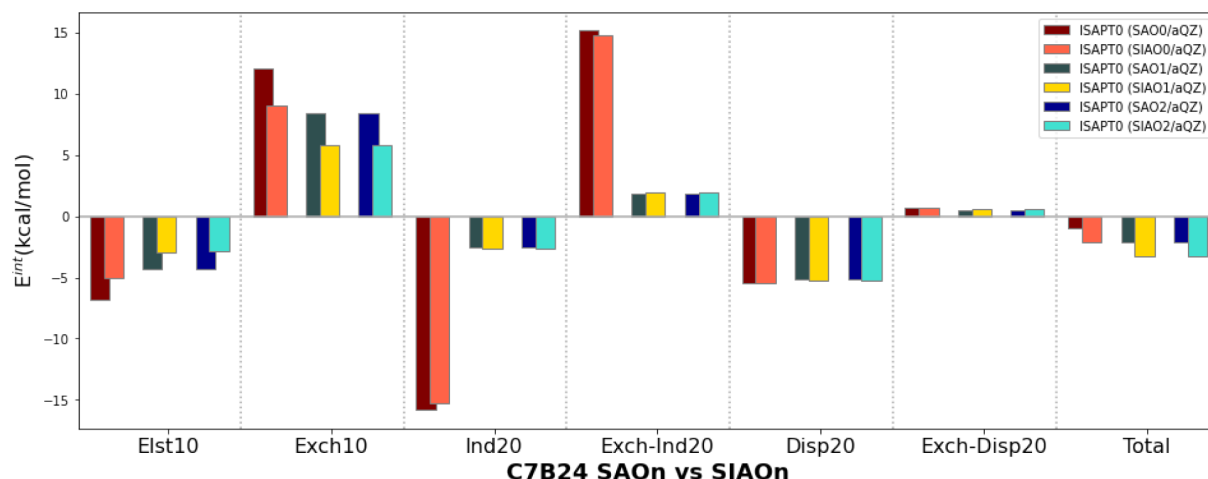


Figure 12. Comparison of interaction energies in the 2,4-dimethylpentane model C7B24, computed using the aQZ basis set with six link assignment options SAOn and SIAOn, $n = 0, 1, 2$.

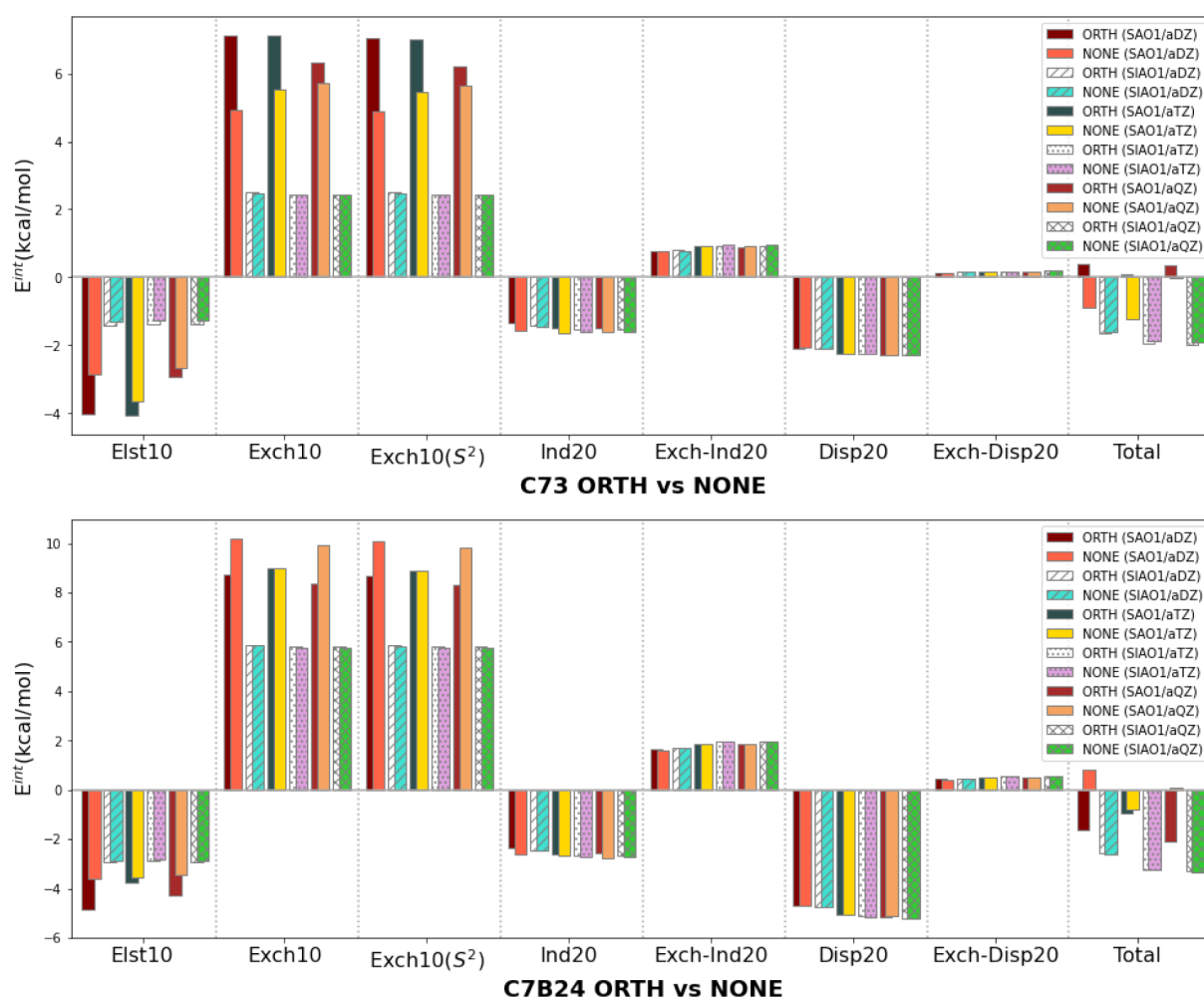


Figure 13. Comparison of ISAPT energy components within *n*-heptane (C73) and 2,4-dimethylpentane (C7B24) molecules with (data marked “ORTH”) and without (marked “NONE”) link orbital orthogonalization, computed using the SAO1 and SIAO1 variants with the aXZ bases, X = D, T, Q.

(which gives 1.42 kcal/mol for the electrostatic term in the C75 model).

Comparison of Different SAO and SIAO Variants, with and without Orthogonalization. The results

presented so far led us to designate the ISAPT(SIAO1) variant as the most meaningful one for practical calculations. We also made extensive comparisons to the ISAPT(SAO1) variant as well as the original ISAPT(C) one. We will now

illustrate how the performance of the method is influenced by the level of self-consistency in the determination of fragment and link occupied orbitals (that is, the choice between SAO0/SAO1/SAO2 or SIAO0/SIAO1/SIAO2), and by the orthogonalization of the link IHOs to the fragment occupied spaces or lack thereof. We will make explicit comparisons using the *n*-heptane and 2,4-dimethylpentane models (C73 and C7B24, respectively), but the conclusions are transferable to the other studied systems as well: analogous tables comparing different ISAPT “minor variants” for the pentanediol models are provided in the [Supporting Information](#).

A comparison of the new ISAPT energy contributions for the C73 and C7B24 models is presented in [Figures 11 and 12](#), respectively. In both SAO and SIAO formalisms, a dramatic change of ISAPT results, especially of induction and exchange-induction energies, is observed between the original HF orbitals (SAO0/SIAO0) and their one-iteration refinement (SAO1/SIAO1). However, the second iteration (SAO2/SIAO2) changes very little, confirming that the algorithm is essentially converged after one iteration. This is the reason why we have focused on the SAO1 and SIAO1 results so far, and why we recommend the SIAO1 variant for all practical calculations: some internal consistency between the fragment and link orbitals is clearly required (otherwise, as shown in [Figures 11 and 12](#), the induction effects blow up quite dramatically), but performing a single iteration of their mutual refinement is entirely sufficient. The SAO2 and SIAO2 methods show only up to ± 0.003 kcal/mol energy differences compared to the corresponding SAO1 and SIAO1 methods, regardless of the basis set size.

Another technical detail of the new ISAPT calculations that turns out to have very minor significance is the orthogonalization of the link IHOs to the occupied space for the fragment (that is, for the SAO1 and SIAO1 approaches, the orthogonalization of $\chi'_x(\chi'_y)$ to the space $\mathcal{B}'_A(\mathcal{B}'_B)$, respectively). All ISAPT(SAO) and ISAPT(SIAO) results presented so far have employed this orthogonalization, but we will now check what happens if the orthogonalization is skipped. A comparison of the ISAPT(SAO1) and ISAPT(SIAO1) energies for *n*-heptane and 2,4-dimethylpentane, with (denoted “ORTH”) and without (“NONE”) the orthogonalization of link IHOs to the fragment occupied space, is presented in [Figure 13](#). For the SAO1 variant, the electrostatic and first-order exchange components with and without orthogonalization follow the same trends, but the actual numerical values are quite different. On the contrary, the differences between the corresponding first-order ISAPT(SIAO1) contributions are nearly negligible, and so are the variations in second-order ISAPT energies for both approaches.

The last technical aspect of new ISAPT calculations that we need to investigate is the choice of spin coupling between the singly occupied link IHOs in the computation of exchange energies. At the first order, the parallel (PAR) and perpendicular (PERP) spin-coupled values of $E_{\text{exch}}^{(10)}$, as well as their averages (AVG), are presented in [Figure 14](#) for four systems: 2,4-pentanediol (P244), 1,4-pentanediol (P142), *n*-heptane (C73), and 2,4-dimethylpentane (C7B24). Due to the larger overlap of the link orbitals, the ISAPT(SAO1) first-order exchange energy is not just more erratic with respect to the basis set, but it is also much more sensitive to the spin coupling than ISAPT(SIAO1). As the overlap integral between the two IHOs is zero in the perpendicular mode, the parallel coupling

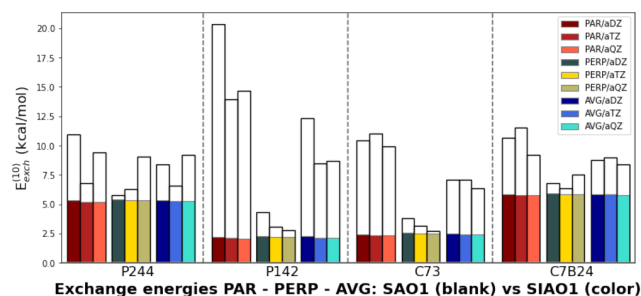


Figure 14. ISAPT exchange energy $E_{\text{exch}}^{(10)}$ from parallel (PAR), perpendicular (PERP), and average (AVG) spin couplings in models P244, P142, C73, and C7B24. The values are computed with SAO1 (empty bars) and SIAO1 (colored bars) in the aXZ bases, X = D,T,Q.

leads to a larger exchange energy in the SAO1 variant. In contrast, the PAR and PERP $E_{\text{exch}}^{(10)}$ values from ISAPT(SIAO1) are highly consistent for all systems studied. In second order (see the [Supporting Information](#)), both $E_{\text{exch-ind,resp}}^{(20)}$ and $E_{\text{exch-disp}}^{(20)}$ are practically independent on spin coupling: the absolute differences between the PAR and PERP values for the systems and basis sets included in [Figure 14](#) do not exceed 0.042 kcal/mol (SAO1) and 2.7×10^{-4} kcal/mol (SIAO1) for $E_{\text{exch-ind,resp}}^{(20)}$ and 0.0032 kcal/mol (SAO1) and 7.4×10^{-5} kcal/mol (SIAO1) for $E_{\text{exch-disp}}^{(20)}$. For all other ISAPT results in this work, the average of the parallel and perpendicular spin-coupling values is chosen as the total exchange contribution.

N-Arylimide Molecular Balances. Noncovalent intramolecular interactions of aromatic groups are crucial in chemical and biological processes, especially in rational drug design, new drug discovery, and the development of synthetic materials, sensors, and catalysts.^{29,38} A prime example of carefully designed intramolecular interactions are the highly versatile N-arylimide molecular torsion balances, an effective platform to quantify non-covalent halogen- π interactions in solution, via the folded (closed) and unfolded (open) conformational equilibrium. Recently, a number of bicyclic N-arylimide balances were synthesized and used to quantitatively probe halogen- π interactions by Sun and coworkers.²⁹ Understanding and tuning the delicate interplay of intramolecular nonbonded interactions between the open and closed conformations of molecular balances requires reliable insights from energy decomposition. Therefore, in this study, the intramolecular interaction energy components of closed and open states of several halogen-containing N-arylimide systems (see [Figure 15](#) for representative structures, and the [Supporting Information](#) for all geometries) are analyzed using the ISAPT(SIAO1) method and compared to the original ISAPT(C) approach. In the closed conformation, a halogen atom is positioned over an aromatic surface (benzene, phenanthrene, or pyrene). In the open conformation, the halogen-substituted benzene ring is rotated 180 degrees so that only a hydrogen atom points towards the aromatic surface. It should be noted that the open and closed structures are separately optimized; thus, their geometric difference involves some rearrangements in addition to the rotation of the substituted benzene ring. To separate the actual noncovalent intramolecular interactions from the geometry relaxation effects (the latter do not lend themselves to a SAPT-like decomposition), we also examine the “openR” structures, which are open-balance conformations obtained from the optimized closed structure directly by a 180-degree rotation of

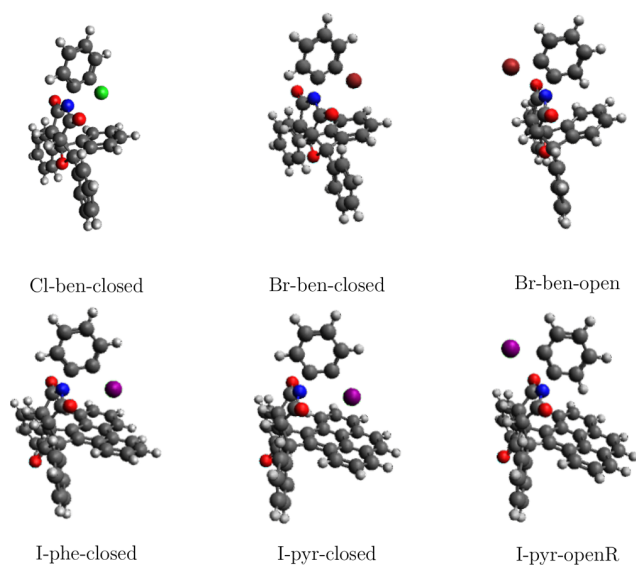


Figure 15. Representative open and closed structures of several molecular balances investigated in this work. The balances vary by the halogen atom (Cl, Br, or I) and its noncovalently interacting partner, the aromatic surface (ben = benzene, phe = phenanthrene, pyr = pyrene). A missing bond signifies a fragment boundary.

the halogen-substituted benzene ring (no geometry reoptimization is performed so the “openR” structure is not a local minimum). An analogous “closedR” structure, a closed conformation obtained by a 180-degree rotation of the open one, involves very short halogen- π distances and is strongly repulsive; therefore, we do not discuss it any further. In all structures, the noncovalent interaction energy of the halogen atom (fragment A) with the bulk of the molecule including the

aromatic surface (fragment B) is investigated, with a single phenylene ring constituting the linker fragment C, as depicted in Figure 15.

The ISAPT(C) and ISAPT(SIAO1) intramolecular interaction energy contributions for several N-arylimide molecular balances are presented in Table 2. The issues of the original C variant are quite evident. First, while the open and “openR” conformations display a grossly overestimated repulsive electrostatics, the ISAPT(C) $E_{\text{elst}}^{(10)}$ energy for closed conformations is usually overly attractive. While this has not happened before for the systems studied in this work, it is a manifestation of the very same issue of unphysical dipole moments at the interfragment boundary: after all, these dipole moments can be aligned unfavorably or favorably. Second, the ISAPT(C) induction and exchange-induction terms for closed conformations are dramatically large. When the individual contributions are so overestimated, it is hard to expect their sufficiently complete cancellation: indeed, the sum $E_{\text{ind,resp}}^{(20)} + E_{\text{exch-ind,resp}}^{(20)}$ is up to 12 times larger for ISAPT(C) than for ISAPT(SIAO1). The issues with the ISAPT(C) induction energy lead to sometimes unphysically large values of the δ_{HF} term—see the Supporting Information. Therefore, the total ISAPT molecular balance interaction energies presented in this section *do not* include δ_{HF} .

On the other hand, a switch to the SIAO1 link assignment solves all the issues of ISAPT(C). The induction and exchange-induction energies for closed conformations are corrected to physically meaningful values, the first-order exchange energy for the same structures is reduced about threefold and, consequently, the total intramolecular interaction becomes attractive in all cases. Just like for the other systems, the dispersion energies are consistent between ISAPT(C) and ISAPT(SIAO1) (<1 kcal/mol differences).

Table 2. Intramolecular Interaction Energy Components (in kcal/mol) of Representative Models of N-Arylimide Molecular Balances Computed at the ISAPT0/aDZ Level of Theory

System	Method	$E_{\text{elst}}^{(10)}$	$E_{\text{exch}}^{(10)}$	$E_{\text{ind,resp}}^{(20)}$	$E_{\text{exch-ind,resp}}^{(20)}$	$E_{\text{disp}}^{(20)}$	$E_{\text{exch-disp}}^{(20)}$	Total
Cl-ben-closed	C	−1.06	14.50	−72.82	70.97	−7.49	1.32	5.42
	SIAO1	−1.87	6.82	−5.90	5.32	−6.83	0.80	−1.66
Cl-ben-openR	C	5.43	2.64	−1.92	0.92	−3.09	0.27	4.27
	SIAO1	0.50	2.39	−1.08	0.56	−3.03	0.23	−0.43
Cl-ben-open	C	5.64	2.27	−1.69	0.77	−2.83	0.23	4.39
	SIAO1	0.63	2.03	−0.84	0.38	−2.78	0.19	−0.39
Br-ben-closed	C	−9.82	24.73	−351.01	344.47	−9.27	1.95	1.06
	SIAO1	−2.99	8.49	−11.10	10.35	−8.34	1.09	−2.50
Br-ben-openR	C	5.54	3.02	−4.36	3.31	−3.41	0.34	4.43
	SIAO1	0.56	2.55	−1.36	0.76	−3.32	0.27	−0.52
Br-ben-open	C	5.68	2.97	−3.88	2.86	−3.32	0.33	4.64
	SIAO1	0.59	2.52	−1.26	0.72	−3.24	0.26	−0.41
I-phe-closed	C	−7.54	22.93	−729.84	720.46	−12.42	2.27	−4.14
	SIAO1	−2.30	8.26	−9.16	8.23	−11.58	1.48	−5.07
I-phe-openR	C	4.87	3.96	−10.83	9.50	−4.43	0.52	3.58
	SIAO1	−0.16	3.17	−2.44	1.61	−4.29	0.42	−1.69
I-phe-open	C	5.50	3.50	−12.44	11.25	−3.91	0.43	4.32
	SIAO1	0.19	2.67	−1.82	1.14	−3.77	0.33	−1.25
I-pyr-closed	C	−5.91	20.95	−600.22	592.43	−12.46	2.15	−3.05
	SIAO1	−2.26	8.02	−8.74	7.82	−11.69	1.45	−5.40
I-pyr-openR	C	4.93	4.45	−10.55	9.15	−4.66	0.57	3.90
	SIAO1	−0.20	3.66	−2.77	1.89	−4.51	0.47	−1.46
I-pyr-open	C	5.50	3.70	−9.85	8.63	−4.04	0.46	4.40
	SIAO1	0.15	2.93	−1.97	1.28	−3.91	0.37	−1.16

The induction and exchange-induction corrections are no longer unreasonably large and their cancellation is more complete, leaving the dispersion energy as the primary binding effect for all systems and conformations. The ISAPT(SIAO1) electrostatic energy is weakly repulsive (<1 kcal/mol) for most open conformations and somewhat attractive (−1.9 to −3.0 kcal/mol) for the closed ones, but neither electrostatics nor induction match the binding strength of dispersion interactions (−2.9 to −3.9 kcal/mol for the open conformations and −6.8 to −11.7 kcal/mol for the closed ones, taking into account the small exchange-dispersion effects as well).

Interestingly, bromine-containing systems exhibit the strongest electrostatic attraction forces in the closed configuration, ranging from −0.71 to −2.99 kcal/mol compared with −1.14 to −2.31 kcal/mol in iodine systems. Furthermore, the benzene surface produces the strongest electrostatic energies for each halogen substituent, about 0.1–0.2 kcal/mol greater than in phenanthrene and pyrene systems. The total ISAPT(SIAO1) interaction energy in closed conformations shows that binding increases with both the size of the aromatic surface and the atomic number of the halogen (Figure 16).

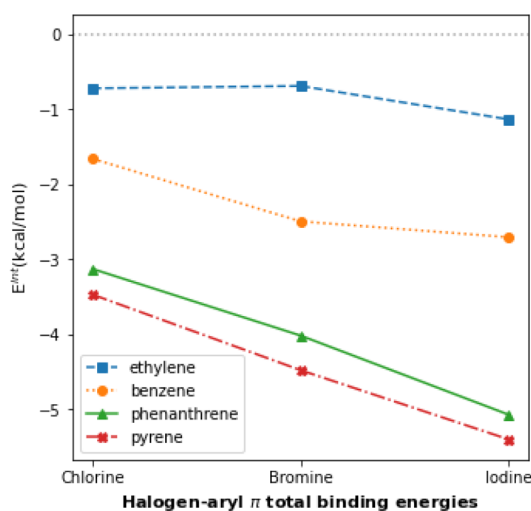


Figure 16. The ISAPT0(SIAO1)/aDZ total interaction energy of closed N-arylimide molecular balances with three halogen substituents, Cl, Br, and I, and four π surfaces, ethylene, benzene, phenanthrene, and pyrene.

For the open state of the molecular balance, the “openR” structure (where the -Ph-X fragment is rotated 180 degrees without reoptimization) has a slightly larger dihedral angle (by about 6 degrees) between the aromatic surface and the imide group connecting to the linker fragment than the “open” one. With the SIAO1 method, a slightly more repulsive first-order exchange and somewhat stronger induction and dispersion energies are observed in the “openR” systems. As a result, the total interaction energy of the “open” systems is up to 0.4 kcal/mol less attractive than of the “openR” systems, as shown in Table 2. However, the overall differences in ISAPT0 interaction energy components between the “open” and “openR” structures are quite minor, showing that geometric relaxation is not a crucial factor for the performance of these molecular balances; the difference in intramolecular non-bonded interactions, which can be readily studied with ISAPT(SIAO1), is a much more important quantity. The increase in the electron density and polarizability of the

halogen substituent from chlorine to iodine is correlated with their total binding energies. In the benzene based balances, the overall interaction energy increases from −0.43 kcal/mol for chlorine to −0.52 kcal/mol for bromine and to −0.98 kcal/mol for iodine in the “openR” configurations. The corresponding increase in the pyrene based balances is from −0.58 kcal/mol for Cl to −0.75 kcal/mol for Br and −1.46 kcal/mol for I.

As shown by the ISAPT0(SIAO1) decomposition, the closed conformations of N-arylimide molecular balances are mainly bound by the dispersion forces. The predicted relative folding energy can thus be computed as $\delta E_{\text{total}}^{\text{int}}(\text{closed} - \text{openR})$ as presented in Figure 17. At equilibrium, due to a

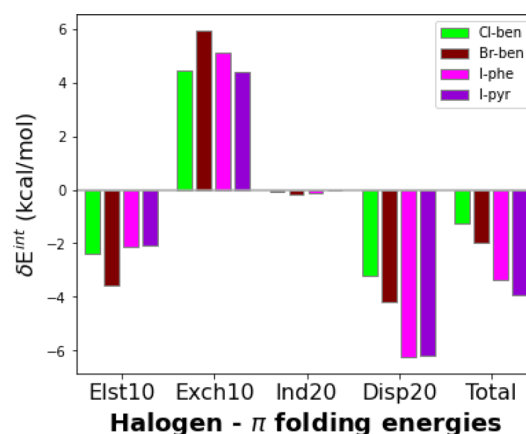


Figure 17. Decomposition of the folding energy (closed–openR) in representative molecular balance systems computed at the ISAPT0(SIAO1)/aDZ level of theory. The effects marked “Ind20” and “Disp20” represent the *complete* induction and dispersion components, including the respective exchange-induction and exchange-dispersion contributions.

close halogen- π contact, the closed structure is strongly stabilized by dispersion forces, which grow with both the halogen atomic number (and thus its polarizability) and the size of the π surface. The closed conformation is secondarily stabilized, relative to the open one, by the electrostatic contribution which is, however, more than counterbalanced by the exchange repulsion term. The effect of induction energy on the relative conformational stability is negligible.

SUMMARY

In this work, we solve the issue of unphysical interaction energy contributions in the original intramolecular SAPT (ISAPT) method of ref 14. This issue culminates in strongly repulsive electrostatic energies for systems where the electrostatics should clearly be favorable (such as those involving intramolecular hydrogen bonding) and while it does not show up in every possible fragmentation pattern, it plagues most of them (especially those with a small linker fragment C). We identified the cause of this issue to be the artificial dipole moments at the A–C and B–C boundaries, where the linking atom is missing electrons on one of its hybrid orbitals as the entire doubly occupied intrinsic bond orbital (IBO) is assigned to C. We propose to overcome this issue by partially undoing the coupling that leads from intrinsic hybrid orbitals (IHOs) to IBOs, and reassign one electron from C to each of A/B, placing it on a suitably constructed approximation to the missing hybrid orbital. This reassignment leads to new, updated density matrices for fragments A and B which can

be employed in standard SAPT0 formulas to compute updated electrostatic, induction, and dispersion energies (it is assumed that the link hybrid orbital does not participate in excitations). For the corresponding exchange corrections, a choice needs to be made whether the single electrons on the A–C and B–C link orbitals undergo parallel, perpendicular, or averaged spin coupling. This is a purely technical issue that has a minor effect on the ISAPT energy contributions.

The selection of the link IHOs involves a projection of the bond orbital onto a suitably designed fragment A/B space. We propose two algorithms for this projection: in the SAO variant, the target space is spanned by the AO basis functions centered on A/B, and in the SIAO one, the target space is spanned by the intrinsic atomic orbitals¹⁷ centered on A/B. Each of these two choices leads to several minor variants depending on the level of self-consistency between the noninteracting orbital space of, say, A (which should not include any interaction with the B link IHO) and the A link IHO (which should be orthogonal to the other occupied orbitals of A). We find out that a minimal level of this self-consistency is essential (the SAO0 and SIAO0 variants are not reliable), but one iteration toward self-consistency, as in the SAO1 and SIAO1 methods, is entirely sufficient.

We show that both ISAPT(SAO1) and ISAPT(SIAO1) resolve the issue with artificial dipole moments on the linking atoms. All ISAPT energy contributions now exhibit physically meaningful values for all fragmentation schemes. In particular, the electrostatic energy is now strongly attractive for intramolecular hydrogen-bonded systems and weakly attractive (due to charge penetration) for nonpolar fragments. The induction and exchange-induction energies are not nearly as large as in original ISAPT(C), which makes their sum more meaningful. The first-order exchange energy typically becomes less repulsive, and the dispersion and exchange-dispersion contributions are in good agreement with the original ISAPT(C) method. Finally, the δ_{HF} term, the only component that is not amenable to the proposed link reassignment, is taken from the original ISAPT(C) theory. In all tested molecular systems, ISAPT(SIAO1) is the best variant for calculating and interpreting intramolecular interactions. Unlike ISAPT(SAO1), the SIAO1 variant shows remarkable consistency between basis sets thanks to a much smaller overlap between the link IHOs when they are formed via a projection onto a fragment IAO space. While both ISAPT(SAO1) and ISAPT(SIAO1) significantly reduce the magnitude of the fragment dipole moments in systems with entirely nonpolar fragments, the former variant leads to a somewhat larger reduction.

The reliable intramolecular energy decomposition provided by the new ISAPT(SIAO1) approach enables one to shed light on the origins of various nonbonded interactions. In this work, we illustrate the new algorithms by comparing the energy components for three pentanediol isomers at a range of fragmentation patterns, comparing the ISAPT description of the intramolecular hydrogen bond in 2,4-pentanediol to the SAPT0 description of the intermolecular hydrogen bond in the water dimer. Next, we use ISAPT(SIAO1) to study alkanes, explaining why the branched isomer, 2,4-dimethylpentane, is more thermodynamically stable than the linear isomer, *n*-heptane. Finally, we investigate a family of N-arylimide molecular torsion balances, differing by the halogen atom and the aromatic surface, to examine the origins of the folding energy, that is, the nonbonded energy difference between the

closed and open conformer. The proposed ISAPT(SIAO1) approach is expected to provide a physically reasonable energy decomposition for any closed-shell molecule that can be separated into two noncovalently interacting fragments and a linker by cutting two single bonds. Therefore, we expect more valuable insights obtained from this method to emerge in the near future.

■ APPENDIX

Nonapproximated First-Order Exchange Energy in new ISAPT

In this Appendix, we will rederive the expression for the full, nonapproximated $E_{\text{exch}}^{(10)}$ ISAPT correction appropriate for the new SAO and SIAO variants where the singly occupied IHOs on fragments A and B do not have a definite spin. The original formula for this correction is given in terms of spinorbitals and still holds;²⁵ however, the despinning of this formula needs to proceed differently as the link spinorbitals involve both spin-up and spin-down contributions and thus couple the two spin blocks of the overlap matrix.

Following the formulation in ref 25, the complete nonexpanded first-order SAPT0 energy $E^{(10)} = E_{\text{elst}}^{(10)} + E_{\text{exch}}^{(10)}$ is expressed as

$$E^{(10)} = W_{AB} + \sum_{ir} B_{ir} D_{ri} + \sum_{jr} A_{jr} D_{rj} + \frac{1}{2} \sum_{ijrs} \langle ij || rs \rangle (D_{ri} D_{sj} - D_{si} D_{rj}) \quad (11)$$

In eq 11, *i* runs over the occupied spinorbitals of A, *j* runs over occupied spinorbitals of B, and *r*, *s* run over occupied spinorbitals of both fragments. Furthermore, W_{AB} is the constant intermolecular nuclear repulsion term, $B_{ir} = \langle \psi_i | V_B | \psi_r \rangle$ and $A_{jr} = \langle \psi_j | V_A | \psi_r \rangle$ are the matrix elements of the nuclear attraction potential of B/A, and $\langle ij || rs \rangle = \langle ij | rs \rangle - \langle ij | sr \rangle$ is the antisymmetrized two-electron integral in the physicists' notation. Finally, D_{rs} is the element of the inverse of the overlap matrix S_{rs} of all occupied spinorbitals, $\mathbf{D} = \mathbf{S}^{-1}$. It is the lack of the block diagonal character of **S**, and thus also **D**, due to spin-up and spin-down coupling that necessitates a somewhat different treatment of eq 11 in the ISAPT/SAO and ISAPT/SIAO cases.

Once again, the explicit form of **S** and **D** depends to some extent on the spin coupling between the electrons occupying the spinorbitals ψ_x and ψ_y , reassigned to fragments A and B in the SAO and SIAO approaches. We will consider both the parallel (eq 2) and perpendicular (eq 3) spin coupling of ψ_x and ψ_y . Let us assume that the occupied A and B spinorbitals ψ_r are ordered as $(\psi_i(l\uparrow), \psi_i(l\downarrow), \psi_x, \psi_j(l\uparrow), \psi_j(l\downarrow), \psi_y)$. In this basis, the matrix **S** has the following explicit block form for each spin coupling:

$$\mathbf{S}^{\parallel} = \begin{bmatrix} \delta_{ii'} & 0 & 0 & S_{ji} & 0 & S_{yi}/\sqrt{2} \\ 0 & \delta_{ii'} & 0 & 0 & S_{ji} & S_{yi}/\sqrt{2} \\ 0 & 0 & 1 & S_{jx}/\sqrt{2} & S_{jx}/\sqrt{2} & S_{yx} \\ S_{ij} & 0 & S_{xj}/\sqrt{2} & \delta_{jj'} & 0 & 0 \\ 0 & S_{ij} & S_{xj}/\sqrt{2} & 0 & \delta_{jj'} & 0 \\ S_{iy}/\sqrt{2} & S_{iy}/\sqrt{2} & S_{xy} & 0 & 0 & 1 \end{bmatrix} \quad (12)$$

$$\mathbf{S}^\perp = \begin{pmatrix} \delta_{ii'} & 0 & 0 & S_{ji} & 0 & S_{ji}/\sqrt{2} \\ 0 & \delta_{ii'} & 0 & 0 & S_{ji} & -S_{ji}/\sqrt{2} \\ 0 & 0 & 1 & S_{jx}/\sqrt{2} & S_{jx}/\sqrt{2} & 0 \\ S_{ij} & 0 & S_{xj}/\sqrt{2} & \delta_{jj'} & 0 & 0 \\ 0 & S_{ij} & S_{xj}/\sqrt{2} & 0 & \delta_{jj'} & 0 \\ S_{iy}/\sqrt{2} & -S_{iy}/\sqrt{2} & 0 & 0 & 0 & 1 \end{pmatrix} \quad (13)$$

One can see that, in either case, $\mathbf{S}$ does not have a block-diagonal character and thus needs to be inverted as a whole (therefore, a $(2N_{\text{occ},\text{A}} + 2N_{\text{occ},\text{B}} + 2) \times (2N_{\text{occ},\text{A}} + 2N_{\text{occ},\text{B}} + 2)$ matrix needs to be inverted instead of a $(N_{\text{occ},\text{A}} + N_{\text{occ},\text{B}}) \times (N_{\text{occ},\text{A}} + N_{\text{occ},\text{B}})$ one in original ISAPT). However, $\mathbf{S}$ and thus $\mathbf{D}$ remain symmetric (or, in the case of $\mathbf{S}^\perp$ and the ψ_y row/column, antisymmetric) with respect to a simultaneous flipping of all spins. As a result, the corresponding (up,up) and (down,down) elements of $\mathbf{D}^\parallel$ or $\mathbf{D}^\perp$ are identical, e.g., $D_{i\uparrow,j\uparrow}^\parallel = D_{i\downarrow,j\downarrow}^\parallel = D_{ij}^{\parallel,\text{ss}}$; similarly, $D_{i\uparrow,j\downarrow}^\parallel = D_{i\downarrow,j\uparrow}^\parallel = D_{ij}^{\parallel,\text{os}}$, with the superscripts 'ss' and 'os' indicating same-spin and opposite-spin terms, respectively. Now, going back to eq 11, we note that the B_{ir} and A_{ir} matrices are spin-diagonal so that only the same-spin block of $\mathbf{D}$ contributes to these terms. To despin the last term in eq 11, we have to break up the antisymmetrized two-electron integral and note that the $\langle ij|rs \rangle$ term requires the same spins within the (i, r) and (j, s) pairs while the $\langle ij|sr \rangle$ term imposes the same spins within the (i, s) and (j, r) pairs. As a result, the following non-zero terms contribute to the last sum in eq 11 (not including the 1/2 factor):

$$\begin{aligned} & \sum_{ijrs} \langle ij|rs \rangle (D_{ri}D_{sj} - D_{si}D_{rj}) \\ &= 4 \sum_{ijrs} \langle ij|rs \rangle D_{ri}^{\text{ss}} D_{sj}^{\text{ss}} - 2 \sum_{ijrs} \langle ij|sr \rangle D_{ri}^{\text{ss}} D_{sj}^{\text{ss}} \\ & \quad - 2 \sum_{ijrs} \langle ij|sr \rangle D_{ri}^{\text{os}} D_{sj}^{\text{os}} - 2 \sum_{ijrs} \langle ij|rs \rangle D_{si}^{\text{ss}} D_{rj}^{\text{ss}} \\ & \quad - 2 \sum_{ijrs} \langle ij|rs \rangle D_{si}^{\text{os}} D_{rj}^{\text{os}} + 4 \sum_{ijrs} \langle ij|sr \rangle D_{si}^{\text{ss}} D_{rj}^{\text{ss}} \end{aligned} \quad (14)$$

Special care needs to be taken when any of the indices i, j, r, s falls on the reassigned link spinorbital (x or y). Such an orbital

contributes (with a factor of $1/\sqrt{2}$) to both spin contributions of an integral, and the spin-flip symmetry of $\mathbf{D}$ is in general different for the parallel and perpendicular spin coupling. The $\mathbf{D}^\parallel$ matrix is symmetric in all cases, for example, $D_{xi\uparrow}^\parallel = D_{xi\downarrow}^\parallel$. On the other hand, $\mathbf{D}^\perp$ is symmetric when it comes to x ($D_{xi\uparrow}^\perp = D_{xi\downarrow}^\perp = D_{xi}^\perp$) but antisymmetric when y is involved ($D_{yi\uparrow}^\perp = D_{yi\downarrow}^\perp = -D_{yi}^\perp$); in particular, $D_{xy}^\perp = 0$.

The one- and two-electron integrals involving the link orbital ψ_y also differ between the parallel and perpendicular spin coupling. In the parallel case, the $\sum_{ir} B_{ir} D_{ri}$ term has the following contributions from the link orbitals, when the index i falls on x and/or the index r falls on x or y :

$$\begin{aligned} \sum_{ir} B_{ir} D_{ri}^\parallel &\rightsquigarrow 2 \sum_{ii'} B_{ii'} D_{ii'}^{\text{ss}} + \sqrt{2} \sum_i B_{ix} D_{xi} + 2 \sum_{ij} B_{ij} D_{ji}^{\text{ss}} \\ & \quad + \sqrt{2} \sum_i B_{iy} D_{yi} + \sqrt{2} \sum_i B_{xi} D_{ix} + B_{xx} D_{xx} \\ & \quad + \sqrt{2} \sum_j B_{xj} D_{jx} + B_{xy} D_{yx} \end{aligned} \quad (15)$$

In the perpendicular case, the same term becomes

$$\begin{aligned} \sum_{ir} B_{ir} D_{ri}^\perp &\rightsquigarrow 2 \sum_{ii'} B_{ii'} D_{ii'}^{\text{ss}} + \sqrt{2} \sum_i B_{ix} D_{xi} + 2 \sum_{ij} B_{ij} D_{ji}^{\text{ss}} \\ & \quad + \sqrt{2} \sum_i B_{iy} D_{yi} + \sqrt{2} \sum_i B_{xi} D_{ix} + B_{xx} D_{xx} \\ & \quad + \sqrt{2} \sum_j B_{xj} D_{jx} \end{aligned} \quad (16)$$

with the B_{xy} term canceling out in this case. The other one-electron term is completely analogous for the parallel spin coupling, and for the perpendicular one, once again, the term containing A_{yx} cancels (note that, numerically, the parallel and perpendicular formulas are different in every term because the matrices $\mathbf{S}$ and $\mathbf{D}$ are different). We see that every link index reduces the term prefactor by a factor of $\sqrt{2}$, in accordance with the weight of the spin-up and spin-down contributions to the overall link spinorbital. The same pattern is observed for the two-electron terms in eq 14, as will be illustrated by a complete decomposition of one of the terms involving both same-spin and opposite-spin contributions. Omitting the summation signs for brevity, we get

$$\begin{aligned}
\langle ij|rs\rangle D_{ij} D_{rs} &\rightsquigarrow 2\langle ij|i'i''\rangle D_{ii'}^{ss} D_{ij'}^{ss} + 2\langle ij|i'i''\rangle D_{ii'}^{os} D_{ij'}^{os} + 2\langle ij|j'i''\rangle D_{ji'}^{ss} D_{ij}^{ss} + 2\langle ij|j'i''\rangle D_{ji'}^{os} D_{ij}^{os} + \sqrt{2}\langle ij|xi'\rangle D_{ii'}^{ss} D_{xj} + \sqrt{2}\langle ij|xi'\rangle D_{ii'}^{os} D_{xj} + \sqrt{2}\langle ij|yi'\rangle D_{ii'}^{ss} D_{yj} \\
&\pm \sqrt{2}\langle ij|yi'\rangle D_{ii'}^{os} D_{yj} + 2\langle ij|i'i''\rangle D_{ii'}^{ss} D_{ij'}^{ss} + 2\langle ij|i'i''\rangle D_{ii'}^{os} D_{ij'}^{os} + 2\langle ij|j'i''\rangle D_{ji'}^{ss} D_{ij}^{ss} + 2\langle ij|j'i''\rangle D_{ji'}^{os} D_{ij}^{os} + \sqrt{2}\langle ij|xj'\rangle D_{ji'}^{ss} D_{xj} + \sqrt{2}\langle ij|xj'\rangle D_{ji'}^{os} D_{xj} + \sqrt{2}\langle ij|yj'\rangle D_{ji'}^{ss} D_{yj} \\
&\pm \sqrt{2}\langle ij|yj'\rangle D_{ji'}^{os} D_{yj} + \sqrt{2}\langle ij|i'i''\rangle D_{ii'}^{ss} D_{ij'}^{ss} + \sqrt{2}\langle ij|i'i''\rangle D_{ii'}^{os} D_{ij'}^{os} + \sqrt{2}\langle ij|j'i''\rangle D_{ji'}^{ss} D_{ij}^{ss} + \sqrt{2}\langle ij|j'i''\rangle D_{ji'}^{os} D_{ij}^{os} + 2\langle ij|xx\rangle D_{xx} D_{xj} + 2\langle ij|yx\rangle D_{xx} D_{yj} \\
&+ \sqrt{2}\langle ij|i'i''\rangle D_{ii'}^{ss} D_{ij'}^{ss} \pm \sqrt{2}\langle ij|i'i''\rangle D_{ii'}^{os} D_{ij'}^{os} + \sqrt{2}\langle ij|j'i''\rangle D_{ji'}^{ss} D_{ij}^{ss} \pm \sqrt{2}\langle ij|j'i''\rangle D_{ji'}^{os} D_{ij}^{os} + 2\langle ij|xy\rangle D_{xx} D_{xj} + 2\langle ij|yy\rangle D_{xx} D_{yj} + \sqrt{2}\langle xj|i'i''\rangle D_{ii'}^{ss} D_{ij} + \sqrt{2}\langle xj|i'i''\rangle D_{ii'}^{os} D_{ij} \\
&+ \sqrt{2}\langle xj|j'i''\rangle D_{ji'}^{ss} D_{ij} + \sqrt{2}\langle xj|j'i''\rangle D_{ji'}^{os} D_{ij} + 2\langle xj|xi\rangle D_{xx} D_{xj} + 2\langle xj|yi\rangle D_{xx} D_{yj} + \sqrt{2}\langle xj|ij'\rangle D_{ji'}^{ss} D_{ij} + \sqrt{2}\langle xj|ij'\rangle D_{ji'}^{os} D_{ij} + \sqrt{2}\langle xj|j'i''\rangle D_{ji'}^{ss} D_{ij} + \sqrt{2}\langle xj|j'i''\rangle D_{ji'}^{os} D_{ij} \\
&+ 2\langle xj|xj'\rangle D_{ji'}^{ss} D_{xj} + 2\langle xj|yj'\rangle D_{ji'}^{ss} D_{yj} + \langle xj|xi\rangle D_{xx} D_{xj} + \langle xj|xi\rangle D_{xx} D_{xj} + \langle xj|j'i''\rangle D_{ji'}^{ss} D_{ij} + \langle xj|j'i''\rangle D_{ji'}^{os} D_{ij} + \sqrt{2}\langle xj|xx\rangle D_{xx} D_{xj} + \sqrt{2}\langle xj|yx\rangle D_{xx} D_{yj} \\
&+ \langle xj|iy\rangle D_{xx} D_{ij} + \langle xj|iy\rangle D_{xx} D_{ij} + \langle xj|j'i''\rangle D_{ji'}^{ss} D_{ij} + \langle xj|j'i''\rangle D_{ji'}^{os} D_{ij} + \sqrt{2}\langle xj|xy\rangle D_{xx} D_{xj} + \sqrt{2}\langle xj|yy\rangle D_{xx} D_{yj} + \sqrt{2}\langle iy|i'i''\rangle D_{ii'}^{ss} D_{ij} \pm \sqrt{2}\langle iy|i'i''\rangle D_{ii'}^{os} D_{ij} \\
&+ \sqrt{2}\langle iy|j'i''\rangle D_{ji'}^{ss} D_{ij} \pm \sqrt{2}\langle iy|j'i''\rangle D_{ji'}^{os} D_{ij} + \langle iy|xi\rangle D_{xx} D_{xj} + \langle iy|xi\rangle D_{xx} D_{xj} + \langle iy|yi\rangle D_{xx} D_{yj} \pm \langle iy|yi\rangle D_{xx} D_{yj} + \sqrt{2}\langle iy|i'i''\rangle D_{ii'}^{ss} D_{ij} \pm \sqrt{2}\langle iy|i'i''\rangle D_{ii'}^{os} D_{ij} \\
&+ \sqrt{2}\langle iy|j'i''\rangle D_{ji'}^{ss} D_{ij} \pm \sqrt{2}\langle iy|j'i''\rangle D_{ji'}^{os} D_{ij} + \langle iy|xj\rangle D_{xx} D_{xj} + \langle iy|xj\rangle D_{xx} D_{xj} + \langle iy|yj\rangle D_{xx} D_{yj} \\
&\pm \langle iy|yj\rangle D_{xx} D_{yj} + 2\langle iy|i'i''\rangle D_{ii'}^{ss} D_{ij} + 2\langle iy|i'i''\rangle D_{ii'}^{os} D_{ij} + \sqrt{2}\langle iy|xx\rangle D_{xx} D_{xj} + \sqrt{2}\langle iy|yx\rangle D_{xx} D_{yj} + 2\langle iy|i'i''\rangle D_{ii'}^{ss} D_{ij} \\
&+ 2\langle iy|j'i''\rangle D_{ji'}^{ss} D_{ij} + \sqrt{2}\langle iy|xy\rangle D_{xx} D_{xj} + \sqrt{2}\langle iy|yy\rangle D_{xx} D_{yj} + 2\langle xj|i'i''\rangle D_{ii'}^{ss} D_{ij} + 2\langle xj|i'i''\rangle D_{ii'}^{os} D_{ij} + \sqrt{2}\langle xj|xi\rangle D_{xx} D_{xj} + \sqrt{2}\langle xj|yi\rangle D_{xx} D_{yj} \\
&+ 2\langle xj|ij'\rangle D_{ji'}^{ss} D_{xj} + 2\langle xj|ij'\rangle D_{ji'}^{os} D_{xj} + \sqrt{2}\langle xj|xx\rangle D_{xx} D_{xj} + \sqrt{2}\langle xj|yx\rangle D_{xx} D_{yj} + \sqrt{2}\langle xj|xy\rangle D_{xx} D_{xj} + \sqrt{2}\langle xj|yy\rangle D_{xx} D_{yj} + \langle xj|xx\rangle D_{xx} D_{xj} + \langle xj|yx\rangle D_{xx} D_{yj} \\
&+ \sqrt{2}\langle xj|iy\rangle D_{xx} D_{ij} + \sqrt{2}\langle xj|iy\rangle D_{xx} D_{ij} + \langle xj|xy\rangle D_{xx} D_{xj} + \langle xj|yy\rangle D_{xx} D_{yj}
\end{aligned} \quad (17)$$

We see that each occurrence of a link-orbital index x or y simply lowers the prefactor of a term by a factor of $\sqrt{2}$, from a total of 4 (twice the same-spin part plus twice the opposite-spin part) when no link indices are present to 1 when all four indices pertain to link orbitals. The same behavior occurs for all other two-electron terms. The variant of eq 17 for the perpendicular spin coupling omits all terms containing $D_{xy}^{\pm} = D_{yx}^{\pm} = 0$, and also the two-electron integral is zero if x and y are the pair of indices belonging to the same electron, e.g., $\langle xj|yi\rangle^{\pm}$

= 0. As a result, all terms underlined in eq 17 are present only in the parallel spin coupling and they are omitted in the perpendicular spin coupling variant. Moreover, the terms involving both y and the opposite-spin D matrix elements change sign; thus the $\pm$ signs in eq 17 which signify a plus sign for the parallel spin coupling and a minus sign for the perpendicular one. An analogous formula for a Coulomb-type integral term, once again in the parallel and perpendicular version, reads

$$\begin{aligned}
\langle ij|rs\rangle D_{ij} D_{rs} &\rightsquigarrow 4\langle ij|i'i''\rangle D_{ii'}^{ss} D_{ij'}^{ss} + 4\langle ij|i'i''\rangle D_{ii'}^{os} D_{ij'}^{os} + 2\sqrt{2}\langle ij|i'i''\rangle D_{ii'}^{ss} D_{xj} + 2\sqrt{2}\langle ij|i'i''\rangle D_{ii'}^{os} D_{xj} + 4\langle ij|j'i''\rangle D_{ji'}^{ss} D_{ij}^{ss} + 4\langle ij|j'i''\rangle D_{ji'}^{os} D_{ij}^{os} + 2\sqrt{2}\langle ij|j'i''\rangle D_{ji'}^{ss} D_{xj} \\
&+ 2\sqrt{2}\langle ij|j'i''\rangle D_{ji'}^{os} D_{xj} + 2\sqrt{2}\langle ij|xi'\rangle D_{ii'}^{ss} D_{xj} + 2\sqrt{2}\langle ij|xi'\rangle D_{ii'}^{os} D_{xj} + 2\langle ij|xy\rangle D_{xx} D_{xj} + 2\langle ij|yx\rangle D_{xx} D_{yj} + 2\sqrt{2}\langle ij|yi'\rangle D_{ii'}^{ss} D_{ij} + 2\sqrt{2}\langle ij|yi'\rangle D_{ii'}^{os} D_{ij} \\
&+ 2\langle ij|xy\rangle D_{xx} D_{xj} + 2\sqrt{2}\langle xj|i'i''\rangle D_{ii'}^{ss} D_{ij} + 2\sqrt{2}\langle xj|i'i''\rangle D_{ii'}^{os} D_{ij} + 2\langle xj|xy\rangle D_{xx} D_{xj} + 2\sqrt{2}\langle xj|j'i''\rangle D_{ji'}^{ss} D_{ij} + 2\sqrt{2}\langle xj|j'i''\rangle D_{ji'}^{os} D_{ij} + 2\langle xj|xj'\rangle D_{ji'}^{ss} D_{xj} \\
&+ 2\langle xj|yj'\rangle D_{ji'}^{ss} D_{yj} + 2\langle xj|xi\rangle D_{xx} D_{xj} + 2\langle xj|xj'\rangle D_{ji'}^{ss} D_{xj} + \sqrt{2}\langle xj|xx\rangle D_{xx} D_{xj} + \sqrt{2}\langle xj|xy\rangle D_{xx} D_{xj} \\
&+ 2\langle xj|yi\rangle D_{xx} D_{yj} + 2\langle xj|yj'\rangle D_{ji'}^{ss} D_{yj} + \sqrt{2}\langle xj|xy\rangle D_{xx} D_{xj} + \sqrt{2}\langle xj|yy\rangle D_{xx} D_{yj} + 2\sqrt{2}\langle iy|i'i''\rangle D_{ii'}^{ss} D_{ij} + 2\sqrt{2}\langle iy|i'i''\rangle D_{ii'}^{os} D_{ij} + 2\sqrt{2}\langle iy|j'i''\rangle D_{ji'}^{ss} D_{ij} \\
&+ 2\sqrt{2}\langle iy|j'i''\rangle D_{ji'}^{os} D_{ij} + 2\sqrt{2}\langle iy|xi\rangle D_{xx} D_{xj} + 2\langle iy|xy\rangle D_{xx} D_{xj} + 2\langle iy|yi\rangle D_{xx} D_{yj} + 2\sqrt{2}\langle iy|xi'\rangle D_{ii'}^{ss} D_{xj} + 2\sqrt{2}\langle iy|xi'\rangle D_{ii'}^{os} D_{xj} + \sqrt{2}\langle iy|xx\rangle D_{xx} D_{xj} + \sqrt{2}\langle iy|yx\rangle D_{xx} D_{yj} \\
&+ 2\langle iy|xy\rangle D_{xx} D_{xj} + \sqrt{2}\langle iy|xy\rangle D_{xx} D_{xj} + \sqrt{2}\langle iy|yy\rangle D_{xx} D_{yj} + 2\langle xj|i'i''\rangle D_{ii'}^{ss} D_{ij} + 2\langle xj|i'i''\rangle D_{ii'}^{os} D_{ij} + \sqrt{2}\langle xj|xi\rangle D_{xx} D_{xj} + \sqrt{2}\langle xj|yi\rangle D_{xx} D_{yj} + 2\langle xj|ij'\rangle D_{ji'}^{ss} D_{xj} \\
&+ 2\langle xj|ij'\rangle D_{ji'}^{os} D_{xj} + \sqrt{2}\langle xj|xx\rangle D_{xx} D_{xj} + \sqrt{2}\langle xj|yx\rangle D_{xx} D_{yj} + \sqrt{2}\langle xj|xy\rangle D_{xx} D_{xj} + \sqrt{2}\langle xj|yy\rangle D_{xx} D_{yj} + \langle xj|xx\rangle D_{xx} D_{xj} + \langle xj|yx\rangle D_{xx} D_{yj} \\
&+ \sqrt{2}\langle xj|iy\rangle D_{xx} D_{ij} + \sqrt{2}\langle xj|iy\rangle D_{xx} D_{ij} + \langle xj|xy\rangle D_{xx} D_{xj} + \langle xj|yy\rangle D_{xx} D_{yj}
\end{aligned} \quad (18)$$

with the underlined terms omitted for the perpendicular spin coupling.

The molecular-orbital formulas for the complete, non-approximated $E^{(10)}$ correction still need to be recast into an atomic-orbital (AO) form so that the operations involving two-electron integrals can be efficiently implemented using generalized Coulomb and exchange matrices.⁴⁰ To this end, we will use capital letters K, L, M, N to denote the AO basis functions; note that the set of AOs spans the entire molecule and is used to expand occupied orbitals on all fragments. The coefficient C_{rK} represents the weight of basis function K in the molecular orbital ψ_r that is, $\psi_r = \sum_K C_{rK} \phi_K^{AO}$. Further, we define the back-transformed inverse-overlap matrix blocks $\mathbf{D}^{vw,spin}$, where v, w are orbital type indices (i, j, r , or s) and the spin case $spin$ is either 'ss' or 'os', as in the following example:

$$\begin{aligned}
(\mathbf{D}^{ri,ss})_{KL} &= \sum_r C_{rK} D_{ri}^{ss} C_{rL} = \sum_{i'} C_{i'K} D_{i'i}^{ss} C_{i'L} \\
&+ \frac{1}{\sqrt{2}} \sum_i C_{xK} D_{xi} C_{iL} + \sum_{ji} C_{jK} D_{ji}^{ss} C_{iL} + \frac{1}{\sqrt{2}} \sum_i C_{yK} D_{yi} C_{iL} \\
&+ \frac{1}{\sqrt{2}} \sum_i C_{iK} D_{ix} C_{xL} + \frac{1}{2} C_{xK} D_{xx} C_{xL} + \frac{1}{\sqrt{2}} \sum_j C_{jK} D_{jx} C_{xL} \\
&+ \frac{1}{2} C_{yK} D_{yx} C_{xL}
\end{aligned} \quad (19)$$

A second example shows the differences between the parallel spin coupling (upper signs) and perpendicular spin coupling (lower signs) in the opposite-spin case (the same-spin formulas are identical for both couplings):

$$\begin{aligned}
 (\mathbf{D}^{sj,os})_{KL} &= \sum_{sj} C_{sK} D_{sj}^{os} C_{jL} \\
 &= \sum_{ij} C_{iK} D_{ij}^{os} C_{jL} + \frac{1}{\sqrt{2}} \sum_i C_{xK} D_{xj} C_{jL} \\
 &\quad + \sum_{jj'} C_{jK} D_{jj'}^{os} C_{j'L} \pm \frac{1}{\sqrt{2}} \sum_i C_{yK} D_{yj} C_{jL} \\
 &\quad \pm \frac{1}{\sqrt{2}} \sum_i C_{iK} D_{iy} C_{yL} + \frac{1}{2} C_{xK} D_{xy} C_{yL} \\
 &\quad \pm \frac{1}{\sqrt{2}} \sum_j C_{jK} D_{jy} C_{yL} \pm \frac{1}{2} C_{yK} D_{yy} C_{yL}
 \end{aligned} \quad (20)$$

(additionally, as stated before, for the perpendicular spin coupling $D_{xy} = 0$). The intermediate expression in quotes would be appropriate in a closed-shell case where each occupied orbital of **A** and **B** holds two electrons, and it has been used before.^{28,41} However, in our case we need to break down summations over individual-fragment indices into doubly occupied orbitals and the link orbital ($i \rightsquigarrow (i, x)$, $j \rightsquigarrow (j, y)$), and the summations over both-fragment indices into all four above parts ($r, s \rightsquigarrow (i, x, j, y)$). To accommodate the correct numerical prefactors for terms involving link orbitals in eqs 15–17, each term in eqs 19 and 20 is reduced by a factor of $\sqrt{2}$ for every link orbital x, y present. One should note that v, w in $\mathbf{D}^{vw,spin}$ are not matrix indices but only denote the specific molecular orbital subblock over which the summation in eq 19 extends. Regardless of the v, w type, all these matrices are of the same size $N_{AO} \times N_{AO}$. Finally, as implied in eqs 19 and 20, the ss/os designation only applies if neither of the indices represents a link orbital.

The definition of $\mathbf{D}^{vw,spin}$ in eqs 19 and 20 allows folding all contributions to any single term in eq 14, for example, the 64 terms in eq 17, into one tensor contraction. The same is achieved for the 8 contributions in eq 15 and the analogous 8 contributions involving the A_{jr} matrix. This leads to the following AO formula for the complete first-order ISAPT interaction energy with the reassignment of link orbitals:

$$\begin{aligned}
 E^{(10)} &= W_{AB} + 2\mathbf{B} \cdot \mathbf{D}^{ri,ss} + 2\mathbf{A} \cdot \mathbf{D}^{rj,ss} + 2\mathbf{D}^{ri,ss} \cdot \mathbf{J}[\mathbf{D}^{sj,ss}] \\
 &\quad - \mathbf{D}^{ri,ss} \cdot \mathbf{K}[\mathbf{D}^{sj,ss}]^T - \mathbf{D}^{ri,os} \cdot \mathbf{K}[\mathbf{D}^{sj,os}]^T \\
 &\quad - \mathbf{D}^{si,ss} \cdot \mathbf{K}[\mathbf{D}^{rj,ss}]^T - \mathbf{D}^{si,os} \cdot \mathbf{K}[\mathbf{D}^{rj,os}]^T \\
 &\quad + 2\mathbf{D}^{si,ss} \cdot \mathbf{J}[\mathbf{D}^{rj,ss}] \\
 &= W_{AB} + 2\mathbf{B} \cdot \mathbf{D}^{ri,ss} + 2\mathbf{A} \cdot \mathbf{D}^{rj,ss} + 4\mathbf{D}^{ri,ss} \cdot \mathbf{J}[\mathbf{D}^{sj,ss}] \\
 &\quad - 2\mathbf{D}^{ri,ss} \cdot \mathbf{K}[\mathbf{D}^{sj,ss}]^T - 2\mathbf{D}^{ri,os} \cdot \mathbf{K}[\mathbf{D}^{sj,os}]^T
 \end{aligned} \quad (21)$$

where the back-transformed inverse-overlap matrix blocks $\mathbf{D}^{vw,spin}$, augmented by link orbital contributions, are defined in eqs 19 and 20, and the generalized Coulomb and exchange matrices $\mathbf{J}[\mathbf{X}]$ and $\mathbf{K}[\mathbf{X}]$ have been defined in eq 7. The final transformation in eq 21 makes use of the fact that the indices r and s are equivalent (have the same summation range), so after summing over the MO indices, for example, $\mathbf{D}^{ri,ss} \equiv \mathbf{D}^{si,ss}$. Equation 21 has the same form for the parallel and perpendicular spin coupling, but the $\mathbf{D}^{vw,spin}$ matrix blocks, as explained above, are defined and computed differently.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jpca.2c06465>.

- Derivation of the modified $E_{\text{exch-disp}}^{(20)}$ expressions involving parallel and perpendicular spin couplings of link orbitals, and additional tables and figures (PDF)
- Cartesian coordinates of all systems studied in this work (TXT)

■ AUTHOR INFORMATION

Corresponding Author

Konrad Patkowski – Department of Chemistry and Biochemistry, Auburn University, Auburn, Alabama 36849, United States; orcid.org/0000-0002-4468-207X; Email: patkowski@auburn.edu

Author

Du Luu – Department of Chemistry and Biochemistry, Auburn University, Auburn, Alabama 36849, United States

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acs.jpca.2c06465>

Notes

The authors declare no competing financial interest. The modified Psi4 code including new ISAPT variants is available at <https://github.com/konpat/psi4/tree/isapt>.

■ ACKNOWLEDGMENTS

This work was supported by the U.S. National Science Foundation award CHE-1955328.

■ REFERENCES

- (1) Jeziorski, B.; Moszyński, R.; Szalewicz, K. Perturbation Theory Approach to Intermolecular Potential Energy Surfaces of van der Waals Complexes. *Chem. Rev.* **1994**, *94*, 1887–1930.
- (2) Szalewicz, K.; Patkowski, K.; Jeziorski, B. Intermolecular Interactions via Perturbation Theory: From Diatoms to Biomolecules. *Struct. Bonding (Berlin)* **2005**, *116*, 43–117.
- (3) Szalewicz, K. Symmetry-adapted perturbation theory of intermolecular forces. *WIREs Comput. Mol. Sci.* **2012**, *2*, 254–272.
- (4) Hohenstein, E. G.; Sherrill, C. D. Wavefunction Methods for Noncovalent Interactions. *WIREs Comput. Mol. Sci.* **2012**, *2*, 304–326.
- (5) Jansen, G. Symmetry-adapted perturbation theory based on density functional theory for noncovalent interactions. *WIREs Comput. Mol. Sci.* **2014**, *4*, 127–144.
- (6) Patkowski, K. Recent developments in symmetry-adapted perturbation theory. *WIREs Comput. Mol. Sci.* **2020**, *10*, No. e1452.
- (7) Garcia, J.; Podeszwa, R.; Szalewicz, K. SAPT codes for calculations of intermolecular interaction energies. *J. Chem. Phys.* **2020**, *152*, 184109.
- (8) Mitoraj, M. P.; Michalak, A.; Ziegler, T. A Combined Charge and Energy Decomposition Scheme for Bond Analysis. *J. Chem. Theory Comput.* **2009**, *9*, 962–975.
- (9) Levine, D. S.; Horn, P. R.; Mao, Y.; Head-Gordon, M. Variational Energy Decomposition Analysis of Chemical Bonding. 1. Spin-Pure Analysis of Single Bonds. *J. Chem. Theory Comput.* **2016**, *12*, 4812–4820.
- (10) Levine, D. S.; Head-Gordon, M. Energy decomposition analysis of single bonds within Kohn-Sham density functional theory. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 12649–12656.
- (11) Meitei, O. R.; Hesselmann, A. On the Stability of Cyclophane Derivates Using a Molecular Fragmentation Method. *ChemPhysChem* **2016**, *17*, 3863–3874.

- (12) Meitei, O. R.; Hesselmann, A. Intramolecular Interactions in Sterically Crowded Hydrocarbon Molecules. *J. Comput. Chem.* **2017**, *38*, 2500–2508.
- (13) Gonthier, J. F.; Corminboeuf, C. Exploration of zeroth-order wavefunctions and energies as a first step toward intramolecular symmetry-adapted perturbation theory. *J. Chem. Phys.* **2014**, *140*, 154107.
- (14) Parrish, R. M.; Gonthier, J. F.; Corminboeuf, C.; Sherrill, C. D. Communication: Practical intramolecular symmetry adapted perturbation theory via Hartree-Fock embedding. *J. Chem. Phys.* **2015**, *143*, 051103.
- (15) Pastorczak, E.; Prlj, A.; Gonthier, J. F.; Corminboeuf, C. Intramolecular symmetry-adapted perturbation theory with a single-determinant wavefunction. *J. Chem. Phys.* **2015**, *143*, 224107.
- (16) Mayer, I. Towards a “Chemical” Hamiltonian. *Int. J. Quantum Chem.* **1983**, *23*, 341–363.
- (17) Knizia, G. Intrinsic Atomic Orbitals: An Unbiased Bridge between Quantum Theory and Chemical Concepts. *J. Chem. Theory Comput.* **2013**, *9*, 4834–4843.
- (18) Schreiner, P. R.; Chernish, L. V.; Gunchenko, P. A.; Tikhonchuk, E. Y.; Hausmann, H.; Serafin, M.; Schlecht, S.; Dahl, J. E. P.; Carlson, R. M. K.; Fokin, A. A. Overcoming lability of extremely long alkane carbon-carbon bonds through dispersion forces. *Nature* **2011**, *477*, 308–311.
- (19) Hohenstein, E. G.; Duan, J.; Sherrill, C. D. Origin of the Surprising Enhancement of Electrostatic Energies by Electron-Donating Substituents in Substituted Sandwich Benzene Dimers. *J. Am. Chem. Soc.* **2011**, *133*, 13244–13247.
- (20) Gryn’ova, G.; Corminboeuf, C. Implications of Charge Penetration for Heteroatom-Containing Organic Semiconductors. *J. Phys. Chem. Lett.* **2016**, *7*, 5198–5204.
- (21) Moszyński, R.; Jeziorski, B.; Rybak, S.; Szalewicz, K.; Williams, H. L. Many-body theory of exchange effects in intermolecular interactions. Density matrix approach and applications to He-F⁺, He-HF, H₂-HF, and Ar-H₂ dimers. *J. Chem. Phys.* **1994**, *100*, 5080–5093.
- (22) Hesselmann, A.; Jansen, G.; Schütz, M. Density-functional theory-symmetry-adapted intermolecular perturbation theory with density fitting: A new efficient method to study intermolecular interaction energies. *J. Chem. Phys.* **2005**, *122*, 014103.
- (23) Patkowski, K.; Żuchowski, P. S.; Smith, D. G. A. First-order symmetry-adapted perturbation theory for multiplet splittings. *J. Chem. Phys.* **2018**, *148*, 164110.
- (24) Lao, K. U.; Herbert, J. M. Atomic Orbital Implementation of Extended Symmetry-Adapted Perturbation Theory (XSAPT) and Benchmark Calculations for Large Supramolecular Complexes. *J. Chem. Theory Comput.* **2018**, *14*, 2955–2978.
- (25) Schäffer, R.; Jansen, G. Intermolecular exchange-induction energies without overlap expansion. *Theor. Chem. Acc.* **2012**, *131*, 1235.
- (26) Smith, D. G. A.; Burns, L. A.; Sirianni, D. A.; Nascimento, D. R.; Kumar, A.; James, A. M.; Schriber, J. B.; Zhang, T.; Zhang, B.; Abbott, A. S.; et al. Psi4NumPy: An Interactive Quantum Chemistry Programming Environment for Reference Implementations and Rapid Development. *J. Chem. Theory Comput.* **2018**, *14*, 3504–3511.
- (27) Khaliullin, R. Z.; Cobar, E. A.; Lochan, R. C.; Bell, A. T.; Head-Gordon, M. Unravelling the Origin of Intermolecular Interactions Using Absolutely Localized Molecular Orbitals. *J. Phys. Chem. A* **2007**, *111*, 8753–8765.
- (28) Smith, D. G. A.; Burns, L. A.; Simmonett, A. C.; Parrish, R. M.; Schieber, M. C.; Galvelis, R.; Kraus, P.; Kruse, H.; Di Remigio, R.; Alenaizan, A.; et al. Psi4 1.4: Open-source software for high-throughput quantum chemistry. *J. Chem. Phys.* **2020**, *152*, 184108.
- (29) Sun, H.; Horatscheck, A.; Martos, V.; Bartetzko, M.; Uhrig, U.; Lentz, D.; Schmieder, P.; Nazaré, M. Direct Experimental Evidence for Halogen-Aryl π Interactions in Solution from Molecular Torsion Balances. *Angew. Chem. Int. Ed.* **2017**, *129*, 6554–6558.
- (30) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465.
- (31) Patkowski, K.; Jeziorski, B.; Szalewicz, K. Unified treatment of chemical and van der Waals forces via symmetry-adapted perturbation expansion. *J. Chem. Phys.* **2004**, *120*, 6849–6862.
- (32) Misquitta, A. J. Charge Transfer from Regularized Symmetry-Adapted Perturbation Theory. *J. Chem. Theory Comput.* **2013**, *9*, 5313–5326.
- (33) Williams, H. L.; Mas, E. M.; Szalewicz, K.; Jeziorski, B. On the effectiveness of monomer-, dimer-, and bond-centered basis functions in calculations of intermolecular interaction energies. *J. Chem. Phys.* **1995**, *103*, 7374–7391.
- (34) Kodrycka, M.; Patkowski, K. Efficient Density-Fitted Explicitly Correlated Dispersion and Exchange Dispersion Energies. *J. Chem. Theory Comput.* **2021**, *17*, 1435–1456.
- (35) Steinmann, S. N.; Wodrich, M. D.; Corminboeuf, C. Overcoming Systematic DFT Errors for Hydrocarbon Reaction Energies. *Theor. Chem. Acc.* **2010**, *127*, 429–442.
- (36) Wodrich, M. D.; Wannere, C. S.; Mo, Y.; Jarowski, P. D.; Houk, K. N.; v. R. Schleyer, P. The Concept of Protobranching and Its Many Paradigm Shifting Implications for Energy Evaluations. *Chem. Eur. J.* **2007**, *13*, 7731–7744.
- (37) Parrish, R. M.; Parker, T. M.; Sherrill, C. D. Chemical Assignment of Symmetry-Adapted Perturbation Theory Interaction Energy Components: The Functional-Group SAPT Partition. *J. Chem. Theory Comput.* **2014**, *10*, 4417–4431.
- (38) Li, P.; Vik, E. C.; Shimizu, K. D. N-Arylimide Molecular Balances: A Comprehensive Platform for Studying Aromatic Interactions in Solution. *Acc. Chem. Res.* **2020**, *53*, 2705–2714.
- (39) Jeziorski, B.; Bulski, M.; Piela, L. First-Order Perturbation Treatment of the Short-Range Repulsion in a System of Many Closed-Shell Atoms or Molecules. *Int. J. Quantum Chem.* **1976**, *10*, 281–297.
- (40) Parrish, R. M.; Burns, L. A.; Smith, D. G. A.; Simmonett, A. C.; DePrince, A. E., III; Hohenstein, E. G.; Bozkaya, U.; Sokolov, A. Y.; Di Remigio, R.; Richard, R. M.; et al. Psi4 1.1: An Open-Source Electronic Structure Program Emphasizing Automation, Advanced Libraries, and Interoperability. *J. Chem. Theory Comput.* **2017**, *13*, 3185–3197.
- (41) Waldrop, J. M.; Patkowski, K. Nonapproximated third-order exchange induction energy in symmetry-adapted perturbation theory. *J. Chem. Phys.* **2021**, *154*, 024103.

Chapter 4

Range Separation of the Interaction Potential in Intermolecular and Intramolecular Symmetry-Adapted Perturbation Theory

This chapter was published as Luu, D.; Corminboeuf, C.; Patkowski, K. *JCTC* (2024).

The ISAPT method [49] relies on a fragmentation of the molecule into two noncovalently interacting fragments A and B embedded into a covalently bonded linker C . A recent refinement of the fragmentation algorithm has led to consistently meaningful ISAPT interaction energy contributions. [39] However, the total ISAPT interaction energy lacks the contributions arising from the linker ($A-C$, $B-C$, and $C-C$) as well as the ones internal to any single fragment ($A-A$ and $B-B$). Thus, the currently available ISAPT approach provides highly useful insights into the interaction between specific nonbonded fragments, but it does not rise to the challenge of completely describing the noncovalent interaction within a molecule. This study suggests range separation as a promising first step toward developing a fragmentation-free ISAPT algorithm. The purely long-range SAPT and ISAPT contributions provide reliable approximations to the complete non-range separated quantities, indicating that the nonbonded intramolecular interaction can be associated with only the long-range interaction potential without any specific fragmentation pattern.

In any SAPT variant, the perturbation operator V is the intermonomer interaction operator, containing Coulombic attraction/repulsion terms between all particles (nuclei and electrons) of A with those of B . The range separation of the Coulomb potential $1/r$, including the Gaussian and error function splittings, is used to check whether the long-range $v_l(r)$ and short-range $v_s(r)$ of V can be identified with noncovalent and covalent interactions, respectively, and the neglect of the latter does not appreciably influence the SAPT0/ISAPT corrections. [40] All range separation-based SAPT and ISAPT methods are implemented into the Psi4 software.

[63] The range separation-based methods involve one-electron and both one- and two-electron long-ranged integrals replacing V throughout the entire SAPT/ISAPT calculations, namely the RSEP-eN and RSEP methods respectively.

The relationship between the separation parameters η (for the Gaussian function splitting) and ω (for the error function splitting) needs to be carefully described. Since the unit of η is r^{-2} and of ω is r^{-1} , a specific value of η is expected to correspond to $k\omega^2$ for some constant k . Therefore, to compare the behavior of the two range separations, the values of ω are taken as equal to the square roots of the corresponding values of η .

The observation is that the long-range part of the interaction potential $v_l(r)$ is sufficient for a faithful description of noncovalent interactions. Overall, the full RSEP scheme using the error function splitting provides the best agreement with conventional SAPT0, e.g. Figs. 4.1-4.2 for the ammonia-chlorine monofluoride complex. The first-order range-separated electrostatic and exchange energies for model intermolecular complexes, see the Supporting Information of Ref. [40], exhibit smooth convergence to the full non-range separated energies when the separation parameter $\eta(\omega)$ goes to infinity. The first-order terms with range separation are close to the nonapproximated values at near-minimum and larger intermolecular separations as long as $\eta \geq 3.0 \text{ bohr}^{-2}$ or $\omega \geq \sqrt{3.0} \text{ bohr}^{-1}$. The full range separation RSEP overperforms its one-electron-only variant RSEP-eN for most systems. In addition, the RSEP second-order induction and exchange-induction terms are highly consistent between different values of $\eta(\omega)$, much more so than the RSEP-eN ones, translating to a more accurate description of the sum $E_{ind,resp}^{(20)} + E_{exch-ind,resp}^{(20)}$. The dispersion and exchange dispersion energies exhibit virtually no dependence on $\eta(\omega)$ in the RSEP-eN variant (as the two-electron integrals are unchanged) and moderate $\eta(\omega)$ dependence for RSEP.

The application of range separation to the ISAPT/SIAO1 linker assignment scheme [39, 40] shows similar smooth and mostly monotonic convergence to the non-range separated values. Due to a stronger interaction, the values of $\eta \leq 5.0 \text{ bohr}^{-2}$ often overestimate or flip the sign of the electrostatic energy. The total interaction energy is affected similarly, as electrostatic energy is typically the leading contributor to the ISAPT energy changes with η and all other

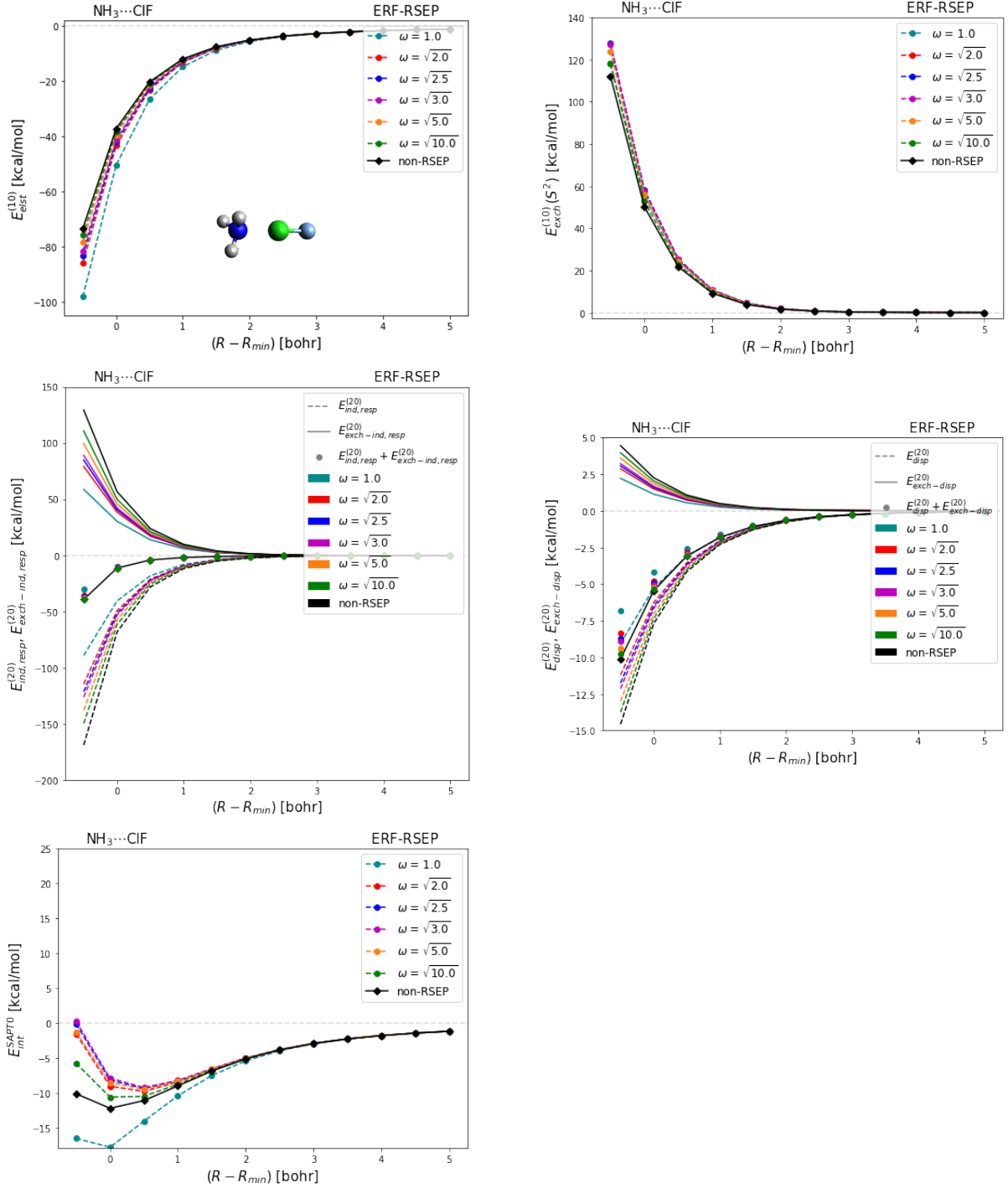


Figure 4.1: The SAPT0(RSEP/erf) interaction energy components in the ammonia-chlorine monofluoride complex computed with various values of ω (bohr^{-1}) in the jun-cc-pVDZ basis set. The intermolecular center-of-mass distance R is given relative to the global minimum value $R_{\min} = 5.569$ bohr.

first- and second-order terms converge faster. Overall, the $\eta \geq 8.0$ bohr^{-2} values afford reasonably accurate range-separated ISAPT contributions, although small residual errors remain even for $\eta = 20.0$ bohr^{-2} . The error-function range separation provides superior consistency

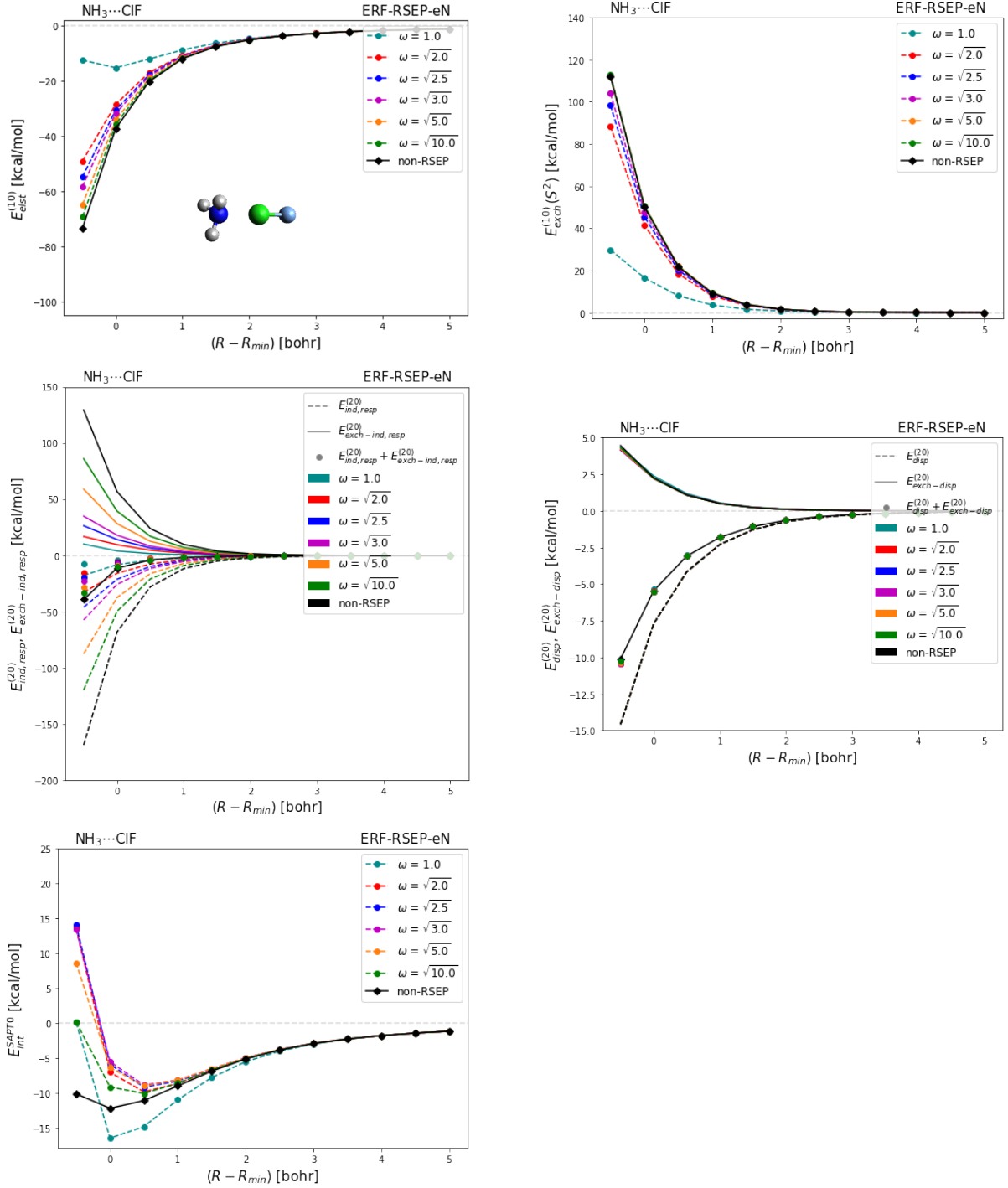


Figure 4.2: The SAPT0(RSEP-eN/erf) interaction energy components in the ammonia-chlorine monofluoride complex computed with various values of ω (bohr^{-1}) in the jun-cc-pVDZ basis set. The intermolecular center-of-mass distance R is given relative to the global minimum value $R_{\min} = 5.569$ bohr.

and fairly rapid convergence in all interaction energy components. The relatively small energy variance between ω values illustrates that the long-ranged error-function potential is the optimal selection for the range separation-based ISAPT method.

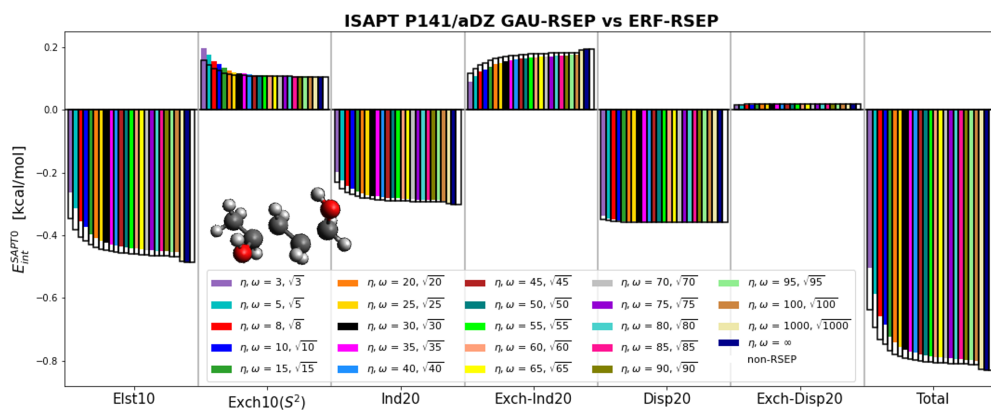


Figure 4.3: The components of the range-separated ISAPT/SIAO1 interaction energy of the 1,4-pentandiol molecule (fragmented as pictured) computed with different values of η (bohr⁻², colored bars) and ω (bohr⁻¹, empty bars) in the aDZ basis set. The energy results are compared to those of the traditional ISAPT/SIAO1 method.

Range Separation of the Interaction Potential in Intermolecular and Intramolecular Symmetry-Adapted Perturbation Theory

Du Luu, Clemence Corminboeuf, and Konrad Patkowski*

Cite This: *J. Chem. Theory Comput.* 2024, 20, 7884–7903

Read Online

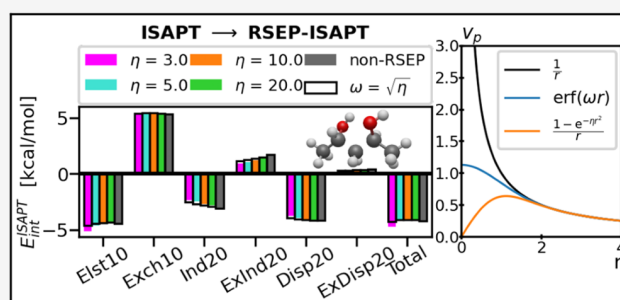
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: Symmetry-adapted perturbation theory (SAPT) is a popular and versatile tool to compute and decompose noncovalent interaction energies between molecules. The intramolecular SAPT (ISAPT) variant provides a similar energy decomposition between two nonbonded fragments of the same molecule, covalently connected by a third fragment. In this work, we explore an alternative approach where the noncovalent interaction is singled out by a range separation of the Coulomb potential. We investigate two common splittings of the $1/r$ potential into long-range and short-range parts based on the Gaussian and error functions, and approximate either the entire intermolecular/interfragment interaction or only its attractive terms by the long-range contribution. These range separation schemes are tested for a number of intermolecular and intramolecular complexes. We find that the energy corrections from range-separated SAPT or ISAPT are in reasonable agreement with complete SAPT/ISAPT data. This result should be contrasted with the inability of the long-range multipole expansion to describe crucial short-range charge penetration and exchange effects; it shows that the long-range interaction potential does not just recover the asymptotic interaction energy but also provides a useful account of short-range terms. The best consistency is attained for the error-function separation applied to all interaction terms, both attractive and repulsive. This study is the first step toward a fragmentation-free decomposition of intramolecular nonbonded energy.



1. INTRODUCTION

Symmetry-adapted perturbation theory (SAPT)^{1–5} is a popular method for studying intermolecular interactions. The key strength of SAPT is its ability to decompose the interaction energy into physically meaningful components, including electrostatics, induction, dispersion, and their exchange counterparts. The SAPT approaches based on a Møller–Plesset^{6,7} or density functional theory (DFT)^{8,9} description of intramolecular electron correlation are well developed, efficiently implemented workhorses for the ab initio treatment of intermolecular forces. Many specialized SAPT variants have been proposed and are the subject of ongoing research, targeting, for example, the interactions between multireference molecules¹⁰ and spin splittings due to exchange interaction¹¹ — see ref 12 for a recent review of other developments.

Notably, some of the new SAPT variants allow the decomposition of nonbonded interactions between fragments of the same molecule.^{13–17} Understanding noncovalent intramolecular interactions is crucial for conformational analysis, chemical reactivity, and molecular design. The intramolecular SAPT (ISAPT) method¹⁶ relies on a fragmentation of the molecule into two noncovalently interacting fragments A and B embedded into a covalently bonded linker C. A recent refinement of the fragmentation algorithm has led to consistently meaningful ISAPT interaction energy contributions.¹⁷ However, the total ISAPT interaction energy lacks the

contributions arising from the linker (A–C, B–C, and C–C) as well as the ones internal to any single fragment (A–A and B–B). Thus, the currently available ISAPT approach provides highly useful insights into the interaction between specific nonbonded fragments, but it does not rise to the challenge of completely describing the noncovalent interaction within a molecule. This limits the utility of ISAPT for systems such as hexaphenylethane and its derivatives¹⁸ which exhibit significant stabilization from intramolecular dispersion involving all six bulky groups rather than any individual pair of single-bonded fragments.

In the development of effective computational strategies, the range separation technique is one of versatile tools to delineate the interaction range and to improve convergence of the interaction energy. The key idea is the splitting of a Coulomb potential $1/r$ into its long-range and short-range parts that are treated differently. Density functionals incorporating range separation and long-range Hartree–Fock (HF) exchange avoid some common issues such as excessive electron delocalization

Received: May 8, 2024

Revised: July 25, 2024

Accepted: August 21, 2024

Published: September 10, 2024



and the resulting overstabilization of charge-transfer states,^{19–21} and they are also useful in the context of SAPT(DFT) as one may avoid an explicit asymptotic correction.^{22–24} On the other hand, density functionals with short-range HF exchange only improve the computational efficiency for periodic systems.²⁵ Another application of the range separation technique occurs in the regularized SAPT method, addressing the challenges associated with the divergence of the perturbation series^{26–28} as well as allowing a decomposition of induction energy into polarization and charge-transfer contributions.²⁹ Several specific forms of partitioning the interaction potential into short-range and long-range components are available, with the Gaussian and error-function based splittings being the most popular (one should also mention the hybrid “erfgau” partitioning).³⁰ The strong point of both splittings is the relatively easy evaluation of the resulting one- and two-electron integrals.

One can expect the long-ranged Coulomb potential to account for nearly all aspects of the noncovalent interaction energy including its complete long-range expansion and some (hopefully good) approximation to the short-range overlap effects. In this work, inspired by a similar long-range vs short-range splitting in the context of supermolecular interaction energies,³¹ we systematically study the convergence behavior of intermolecular and intramolecular SAPT energy contributions using only the long-ranged potential in the interaction operator. We investigate a diverse collection of range-separated SAPT methods, including one-electron or both one- and two-electron long-ranged integrals evaluated with Gaussian and error-function splittings. We observe that, for a broad range of the separation parameter values, the purely long-range SAPT and ISAPT contributions provide reliable approximations to the complete non-range separated quantities, indicating that the energy contribution resulting from noncovalent interactions can be identified solely via range separation of the interaction potential. This is not useful in the intermolecular context, where the noncovalently interacting units are clearly defined. However, it is a potential game changer for ISAPT, where the results so far have been tied to a specific fragmentation of the molecule into the A–B–C subsystems, with the ISAPT decomposition only addressing the A–B noncovalent interaction. If one wants to quantify and decompose the *entire* nonbonded energy in a molecule, the noncovalent and covalent contributions have to be separated in a way that does not assume any specific fragmentation pattern. This work provides such an alternative separation, indicating that the nonbonded intramolecular interaction can be, to a good accuracy, associated with the long-range interaction potential only. In this way, our proof of concept study suggests range separation as a promising first step toward developing a fragmentation-free ISAPT algorithm.

All range separation-based SAPT and ISAPT methods proposed here are implemented into the PSI4 software³² and available on GitHub: <https://github.com/dluu12/psi4/tree/rsep>.

2. THEORY

The inner workings of the intermolecular symmetry-adapted perturbation theory (SAPT) are described in a number of reviews; see, for example, refs 1 and 5. In this work, we will examine the simplest level of SAPT that entirely neglects intramonomer electron correlation and is usually termed SAPT0. The SAPT0 interaction energy is decomposed as

$$E_{\text{int}}^{\text{SAPT0}} = (E_{\text{elst}}^{(10)}) + (E_{\text{exch}}^{(10)}) + (E_{\text{ind,resp}}^{(20)} + E_{\text{exch-ind,resp}}^{(20)} + \delta E_{\text{int}}^{\text{HF}}) + (E_{\text{disp}}^{(20)} + E_{\text{exch-disp}}^{(20)}) \quad (1)$$

The SAPT0 corrections $E^{(kl)}$, grouped by parentheses into the total electrostatic, exchange, induction, and dispersion terms, are labeled by their order k in the intermonomer interaction operator and their order l in the intramonomer correlation operator. The additional subscript “resp” denotes the coupled perturbed Hartree–Fock response (relaxation) of monomer orbitals in the electric field of the interacting partner, and $\delta E_{\text{int}}^{\text{HF}}$ is an estimate of higher-order induction and exchange-induction effects from the supermolecular HF approach. While many higher-level intermolecular SAPT variants are available, employing different strategies to capture intramonomer correlation,^{6–9,33} intramolecular SAPT (ISAPT) is only available at the SAPT0 level of perturbation theory.^{16,17}

In any SAPT variant, the perturbation operator V is the intermonomer interaction operator, containing Coulombic attraction/repulsion terms between all particles (nuclei and electrons) of A with those of B. In this work, we will examine the effects of separating V into its long-range and short-range part by a partitioning of the Coulomb potential,

$$\frac{1}{r} = v_l(r) + v_s(r) \quad (2)$$

using either the Gaussian function splitting

$$v_l(r) = \frac{1 - e^{-\eta^2}}{r} \quad v_s(r) = \frac{e^{-\eta^2}}{r} \quad (3)$$

or the error-function (erf) splitting

$$v_l(r) = \frac{\text{erf}(\omega r)}{r} \quad v_s(r) = \frac{1 - \text{erf}(\omega r)}{r} \quad (4)$$

The parameters ω (in units of (distance)^{−1}) and η (in units of (distance)^{−2}) define the width of the range separation. The erf splitting of eq 4 is ubiquitous in range-separated density functional theory (DFT).^{20,21} The Gaussian splitting of eq 3 is less prevalent but has been used before in the context of SAPT to regularize the attractive singularities in the Coulomb potential^{28,29} (more on that process below). Overall, we expect that $v_l(r)$ will play a much bigger role in any SAPT corrections than $v_s(r)$: the long-range potential should be able to recover the entire asymptotic expansion of interaction energy in inverse powers of the intermolecular distance R as well as at least some of the overlap-dependent, exponentially vanishing terms. Therefore, we will examine the role of the approximation $1/r \approx v_l(r)$, that is, the neglect of the short-range part of the interaction operator V , on various SAPT corrections. It is possible to do the opposite and approximate $1/r$ by its short-range part $v_s(r)$ (in fact, both of these approximations have been tested before for the supermolecular interaction energy curves of rare gas dimers);³¹ however, we will not pursue this direction as the long-range interaction is obviously crucial to obtain a reasonable, asymptotically meaningful interaction energy.

The short-range Coulomb interaction can be neglected in different parts of the interaction operator. For example, only the attractive, one-electron part of V can be replaced by its long-range component:

$$V_{1e} = - \sum_{A \in A} \sum_{j \in B} Z_A v_1(r_{Aj}) - \sum_{B \in B} \sum_{i \in A} Z_B v_1(r_{Bi}) + \sum_{i \in A} \sum_{j \in B} \frac{1}{r_{ij}} + \sum_{A \in A} \sum_{B \in B} \frac{Z_A Z_B}{r_{AB}} \quad (5)$$

The rationale behind eq 5 is that the attractive singularities in V are more consequential than the repulsive ones, as the electrons from one monomer are pulled into the Coulomb wells around the other monomer's nuclei (note that the perturbation wave functions exist in a space spanned by products of monomer wave functions, and the permutation symmetry for electron interchanges between monomers is not enforced). This overpolarization of monomers resulting from the attractive singularities is the reason for the overall divergence of the SAPT series,^{28,34} which manifests itself in practice by the induction and exchange-induction energies increasing in magnitude with increasing order of SAPT (note that the dispersion energy does not depend on one-electron integrals so it is not directly affected). Accordingly, the removal (*regularization*) of attractive singularities in eq 5 was proposed to make the SAPT expansion convergent.^{26–28} While an outright neglect of $v_s(r)$ in eq 5 obviously alters the limit of the perturbation series, this operator can be brought back by the so-called *strong symmetry forcing*.¹ The resulting perturbation expansion was shown for a model Li–H system to converge to the full configuration interaction (FCI) interaction energy.²⁸ A different application of regularized SAPT based on eq 5 was proposed by Misquitta²⁹ and involves the (approximate) partitioning of induction energy into polarization and charge transfer (charge delocalization) contributions. The charge-transfer term involves the tunneling of electrons into the other monomer's attractive Coulomb wells, so it is suppressed by regularization of eq 5 while the polarization term remains.

Alternatively, the short-range part of the Coulomb potential can be neglected in all terms in V , attractive and repulsive:

$$V_{1e2e} = - \sum_{A \in A} \sum_{j \in B} Z_A v_1(r_{Aj}) - \sum_{B \in B} \sum_{i \in A} Z_B v_1(r_{Bi}) + \sum_{i \in A} \sum_{j \in B} v_1(r_{ij}) + \sum_{A \in A} \sum_{B \in B} Z_A Z_B v_1(r_{AB}) \quad (6)$$

This form of perturbation might potentially be preferable over V_{1e} of eq 5 as the attractive and repulsive terms are treated on an equal footing. It is well-known that the SAPT terms, most notably the electrostatic energy, are a result of a delicate balance between the attractive and repulsive Coulomb forces, and this balance could be disturbed when only the attractive forces are attenuated at short range. The potential V_{1e2e} was briefly tested in the context of regularized SAPT for the H–H complex,²⁶ leading to similar results as V_{1e} , however, more testing for larger complexes is warranted.

Let us point out some important differences in the goals and implementation of the range separated SAPT in this work and the regularized SAPT approach of refs 26 and 28, specifically, the regularized symmetrized Rayleigh–Schrödinger (R-SRS) theory adapted in ref 29. On the technical side, we will be neglecting the attractive (eq 5) or all (eq 6) short-range terms completely, replacing V by either V_{1e} or V_{1e2e} throughout the entire SAPT calculation. On the other hand, the R-SRS perturbation theory employs V_{1e} (V_{1e2e}) in the calculation of wave function corrections (where convergence needs to be enforced) but switches to the entire V in the subsequent

computation of the SRS energy terms from the wave function corrections. The R-SRS approach is specifically geared at computing the induction and exchange-induction terms, where it can be used to achieve a convergent series (ref 28) or to separate the polarization and charge-transfer contributions (ref 29). However, the range separation studied here applies to all SAPT0/ISAPT terms except for $\delta E_{\text{int}}^{\text{HF}}$. The main goal of this range separation is to check whether the long-range and short-range part of V can be identified with noncovalent and covalent interactions, respectively, so that the neglect of the latter does not appreciably influence the SAPT0/ISAPT corrections.

In the ISAPT method,¹⁶ the molecule is partitioned into noncovalently interacting fragments A and B and a linker C that connects to A and B by single covalent bonds. Then, the occupied orbitals and density matrices of noninteracting A and B are obtained by solving the Hartree–Fock equations for each fragment embedded in the Fock matrix of linker C. Those orbitals and density matrices are then used in standard intermolecular SAPT0 expressions to compute different noncovalent intramolecular energy contributions. It should be noted that the ISAPT accuracy is sensitive to the specific algorithm used to partition the electron density into fragments, and the original partitioning of ref 16 often leads to meaningless results due to unphysical dipole moments emerging at the interfragment boundaries. To remedy this problem, we have recently proposed¹⁷ a number of alternative partitionings aimed at reducing those spurious dipoles. Numerical tests indicated that one of the new algorithms, which assigns one electron from the interfragment bonding orbital to A (B) and places it onto an orbital obtained by a projection onto the space of intrinsic atomic orbitals (IAOs)³⁵ of a fragment, exhibits both meaningful ISAPT partitioning for all fragmentation patterns and smooth basis set convergence. This algorithm, termed SIAO1 as it Splits link orbitals in the IAO space by 1 iteration of a self-consistent algorithm, was recommended in ref 17 and will be employed in this work. Here, we examine how the replacement of the full interfragment interaction operator V by either V_{1e} or V_{1e2e} affects the ISAPT energy contributions. As noncovalently interacting fragments are typically closer together than noncovalently interacting separate molecules, we expect ISAPT to be a harder test case than SAPT0, placing more severe restrictions on the range separation parameter η (ω).

3. RESULTS AND DISCUSSION

The range-separated SAPT approach requires new one- and two-electron integrals involving either $v_1(r)$ or $v_s(r)$ (integrals of the other type are then obtained by subtraction from the standard integrals involving $1/r$). Specifically, in the Gaussian splitting approach of eq 3 we implemented the $v_s(r)$ integrals, and for the error-function splitting of eq 4 we programmed the $v_1(r)$ integrals. Both two-electron integral types were already available in the `PSI4` code used for our development:³² the error-function one as a key building block of range-separated density functional integrals^{20,21} and the Gaussian one as a standard integral in the explicitly correlated F12 theory.^{36,37} On the other hand, the corresponding one-electron integrals were derived and programmed by us utilizing the McMurchie–Davidson angular momentum recursion³⁸ – see Appendix for formulas. The resulting integrals were compared to our older unpublished code utilizing the Obara–Saika recursion³⁹ to verify the correctness of our implementation. The integrals were implemented into a

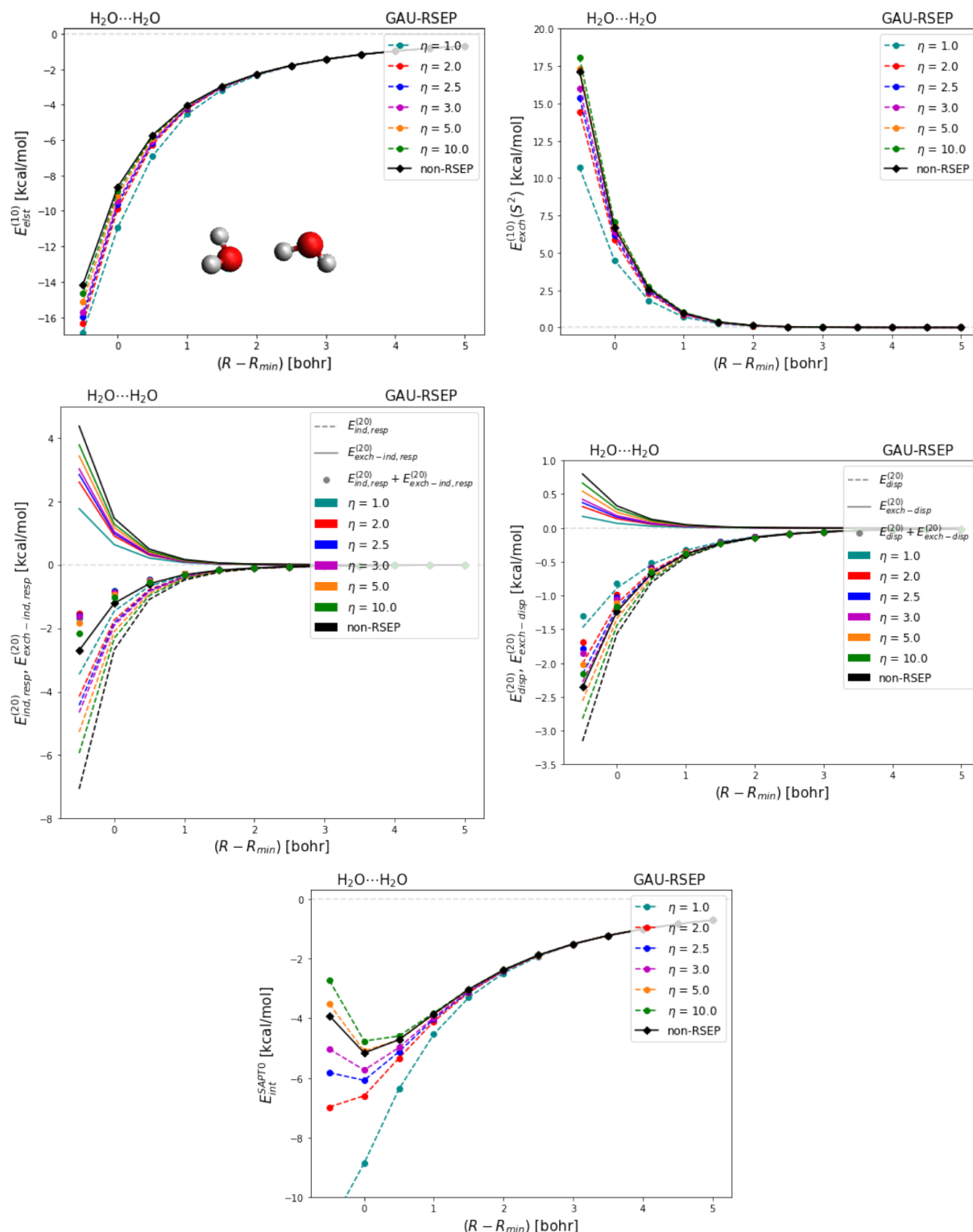


Figure 1. SAPT0(RSEP/gau) interaction energy components in the water dimer computed with various values of η (bohr^{-2}) in the jun-cc-pVDZ basis set. The intermolecular center-of-mass distance R is given relative to the global minimum value $R_{\min} = 5.525$ bohr.

development version of Psi4 and are accessible from both the C++ and Python layers.

All the range-separated intermolecular SAPT approaches utilizing the $v_1(r)$ Coulomb potential, partitioned with either the Gaussian function (eq 3) or the error function (eq 4), were examined and analyzed using the Psi4NumPy framework⁴⁰ and

subsequently programmed into the density-fitted ISAPT/SAPT0 code in Psi4. Four noncovalent systems were chosen to study the behavior of the interaction energy as a function of the intermolecular distance R : the methane dimer, water dimer, the ammonia-chlorine monofluoride complex, and the water-fluoride ion one. The geometry optimizations for the

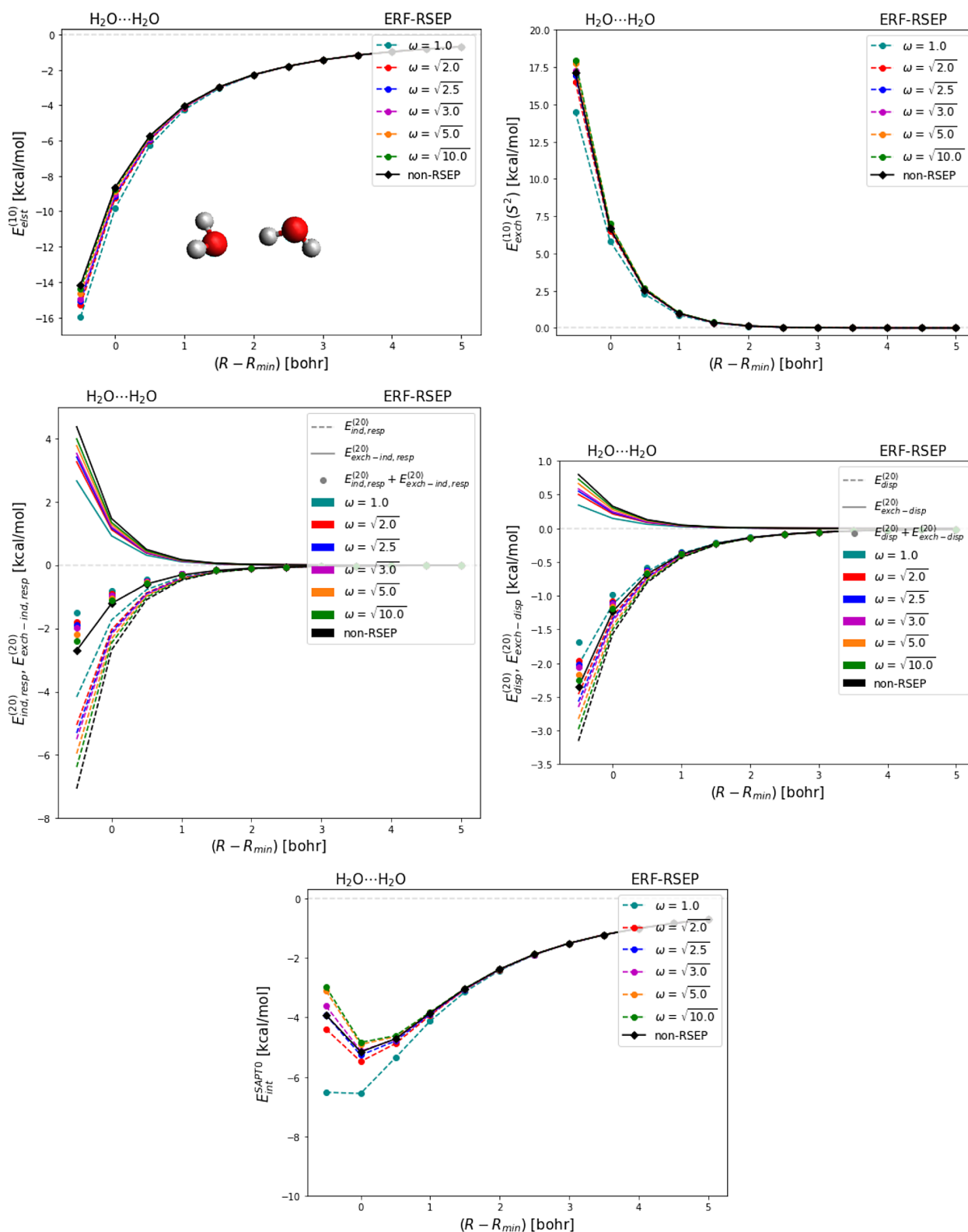


Figure 2. SAPT0(RSEP/erf) interaction energy components in the water dimer computed with various values of ω (bohr^{-1}) in the jun-cc-pVDZ basis set. The intermolecular center-of-mass distance R is given relative to the global minimum value $R_{\min} = 5.525$ bohr.

$\text{H}_2\text{O} \cdots \text{H}_2\text{O}$, $\text{H}_2\text{O} \cdots \text{F}^-$, and $\text{NH}_3 \cdots \text{ClF}$ systems are performed at the MP2 level of theory using the aug-cc-pVDZ basis set. The structure of the $\text{CH}_4 \cdots \text{CH}_4$ complex is taken from the S22⁴¹ database. As the distance R between the centers of masses of the

two monomers is changed to produce potential energy curves, the geometry of each monomer remains the same as in the global-minimum structure of the dimer. The single-point interaction energies as functions of R with multiple range

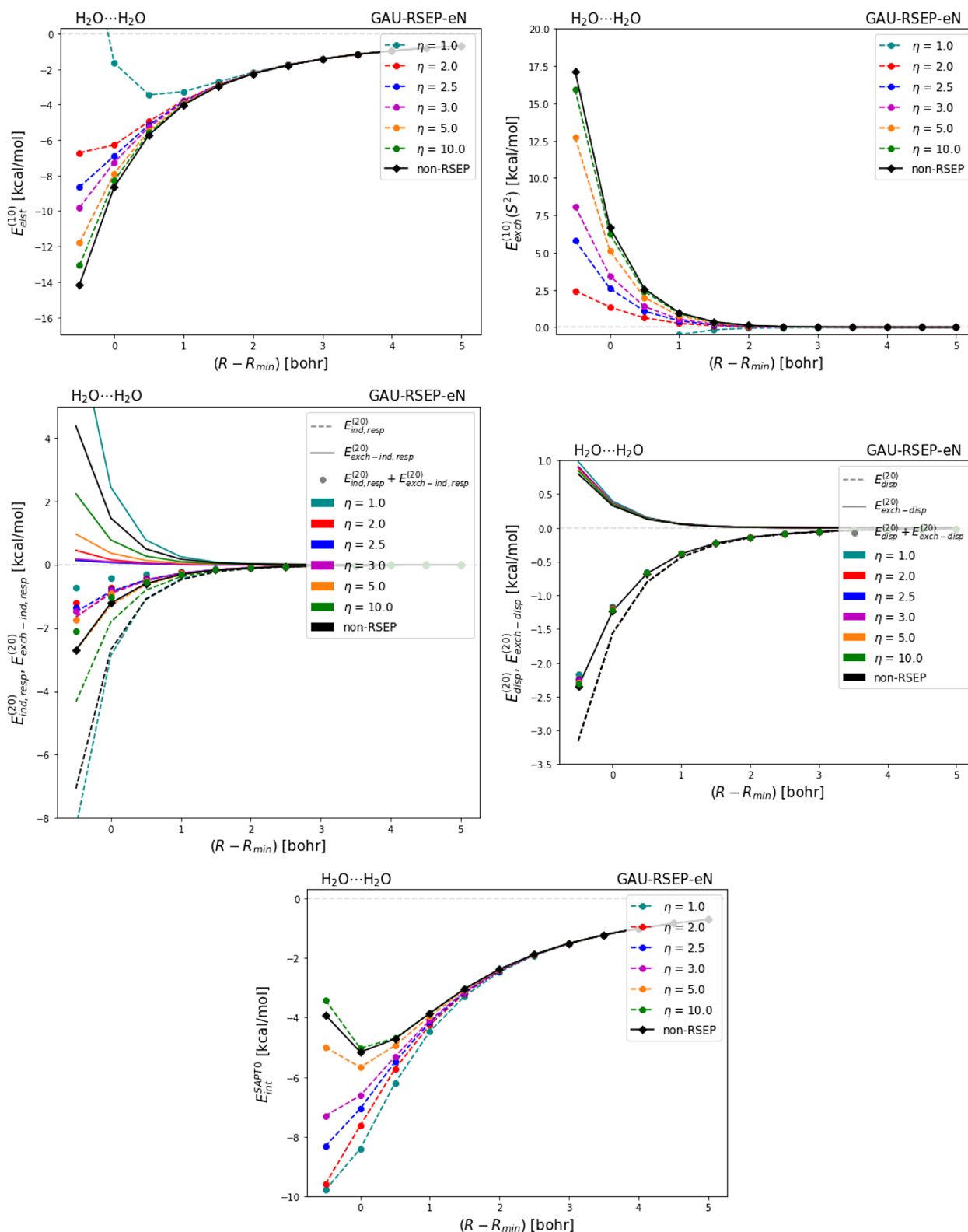


Figure 3. SAPT0(RSEP-eN/gau) interaction energy components in the water dimer computed with various values of η (bohr^{-2}) in the jun-cc-pVDZ basis set. The intermolecular center-of-mass distance R is given relative to the global minimum value $R_{\min} = 5.525$ bohr.

separation parameters are computed using the SAPT0/jun-cc-pVDZ method. For the $\text{H}_2\text{O} \cdots \text{H}_2\text{O}$ and $\text{NH}_3 \cdots \text{ClF}$ systems, we also performed full range-separated (eq 6) SAPT0 calculations in the jun-cc-pVTZ and jun-cc-pVQZ bases,

obtaining very similar convergence with the range separation parameters $\eta(\omega)$; see the [Supporting Information](#) for the pertinent figures. The “delta HF” contribution⁵ in SAPT0 is obtained without range separation, as it would not be possible to

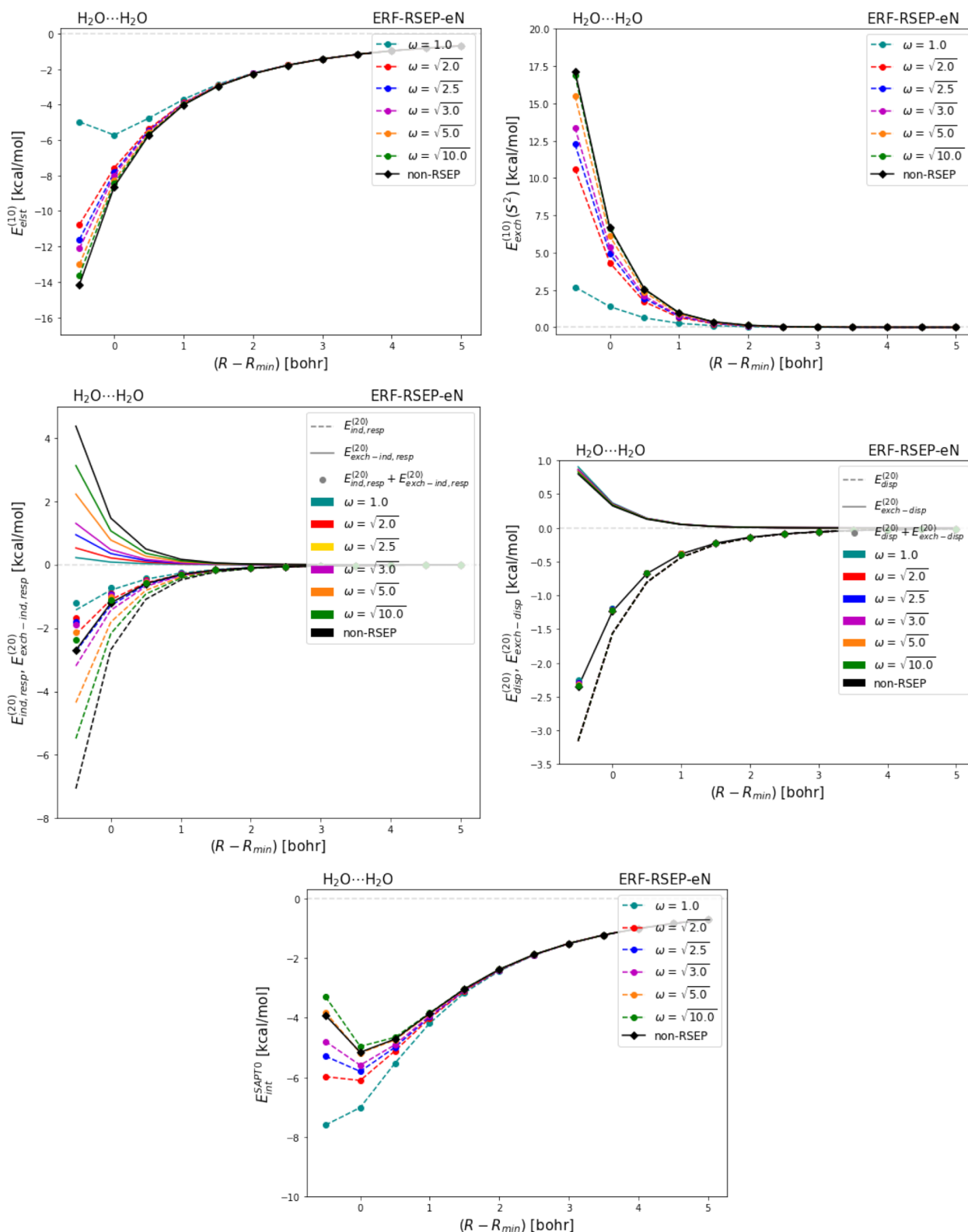


Figure 4. SAPT0(RSEP-eN/erf) interaction energy components in the water dimer computed with various values of ω (bohr $^{-1}$) in the jun-cc-pVDZ basis set. The intermolecular center-of-mass distance R is given relative to the global minimum value $R_{min} = 5.525$ bohr.

compute the HF energy of the entire complex in a manner consistent with range-separated SAPT corrections (using standard integrals for intramonomer terms and long-range ones for intermonomer terms). For consistency with the

standard practice, this contribution is included in all total SAPT0 interaction energies presented below, but it is the same for all values of η (ω) so it does not influence the convergence of

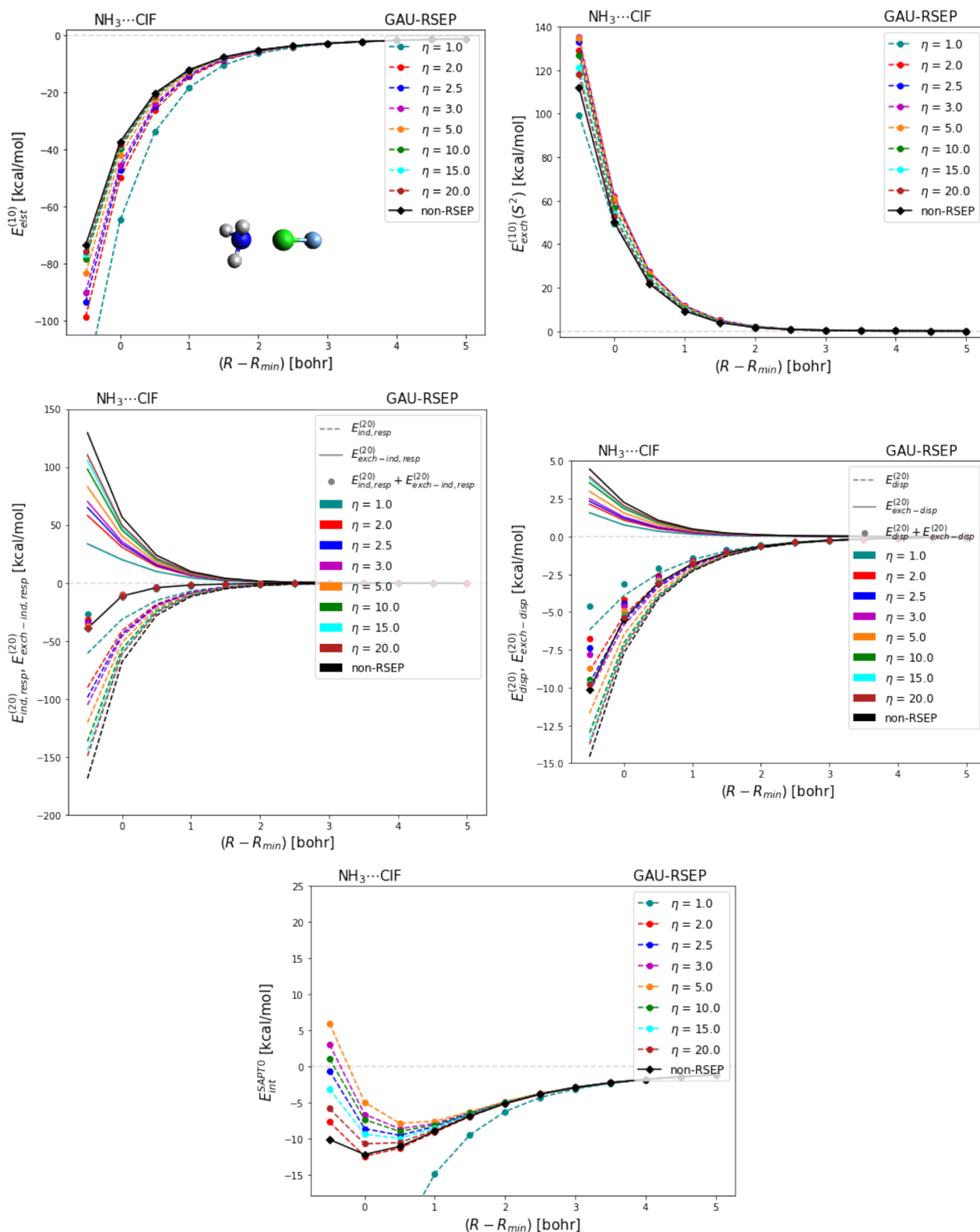


Figure 5. SAPT0(RSEP/gau) interaction energy components in the ammonia-chlorine monofluoride complex computed with various values of η (bohr^{-2}) in the jun-cc-pVDZ basis set. The intermolecular center-of-mass distance R is given relative to the global minimum value $R_{\min} = 5.569$ bohr.

the range-separated quantities. The ISAPT data presented in this work do not contain the “delta HF” term.

The relationship between the separation parameters η (for the Gaussian function splitting) and ω (for the error function

splitting) needs to be carefully described. Since the unit of η is r^{-2} and of ω is r^{-1} , a specific value of η is expected to correspond to $k\omega^2$ for some constant k . Therefore, to compare the behavior

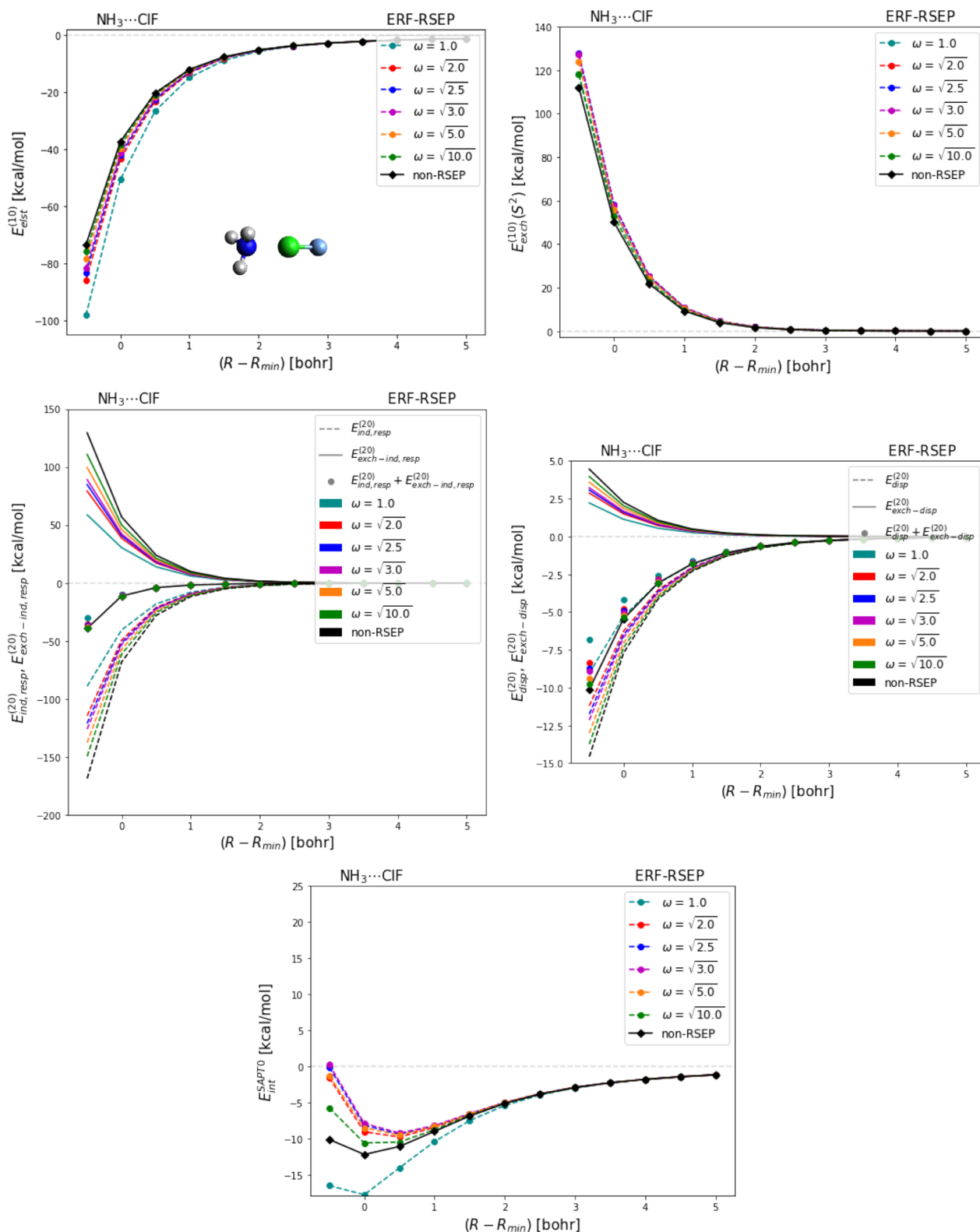


Figure 6. SAPT0(RSEP/erf) interaction energy components in the ammonia-chlorine monofluoride complex computed with various values of ω (bohr^{-1}) in the jun-cc-pVDZ basis set. The intermolecular center-of-mass distance R is given relative to the global minimum value $R_{\min} = 5.569$ bohr.

of the two range separations, we will investigate the values of ω equal to the square roots of the corresponding values of η .

Figures 1–4 present the interaction energies in the water dimer computed with the full range-separated SAPT (eq 6,

denoted RSEP for range separated electrostatic potential) and the one-electron range-separated SAPT (eq 5, denoted RSEP-eN). The corresponding figures for the $\text{NH}_3 \cdots \text{CIF}$ complex are shown in Figures 5–8, and those for the two remaining systems

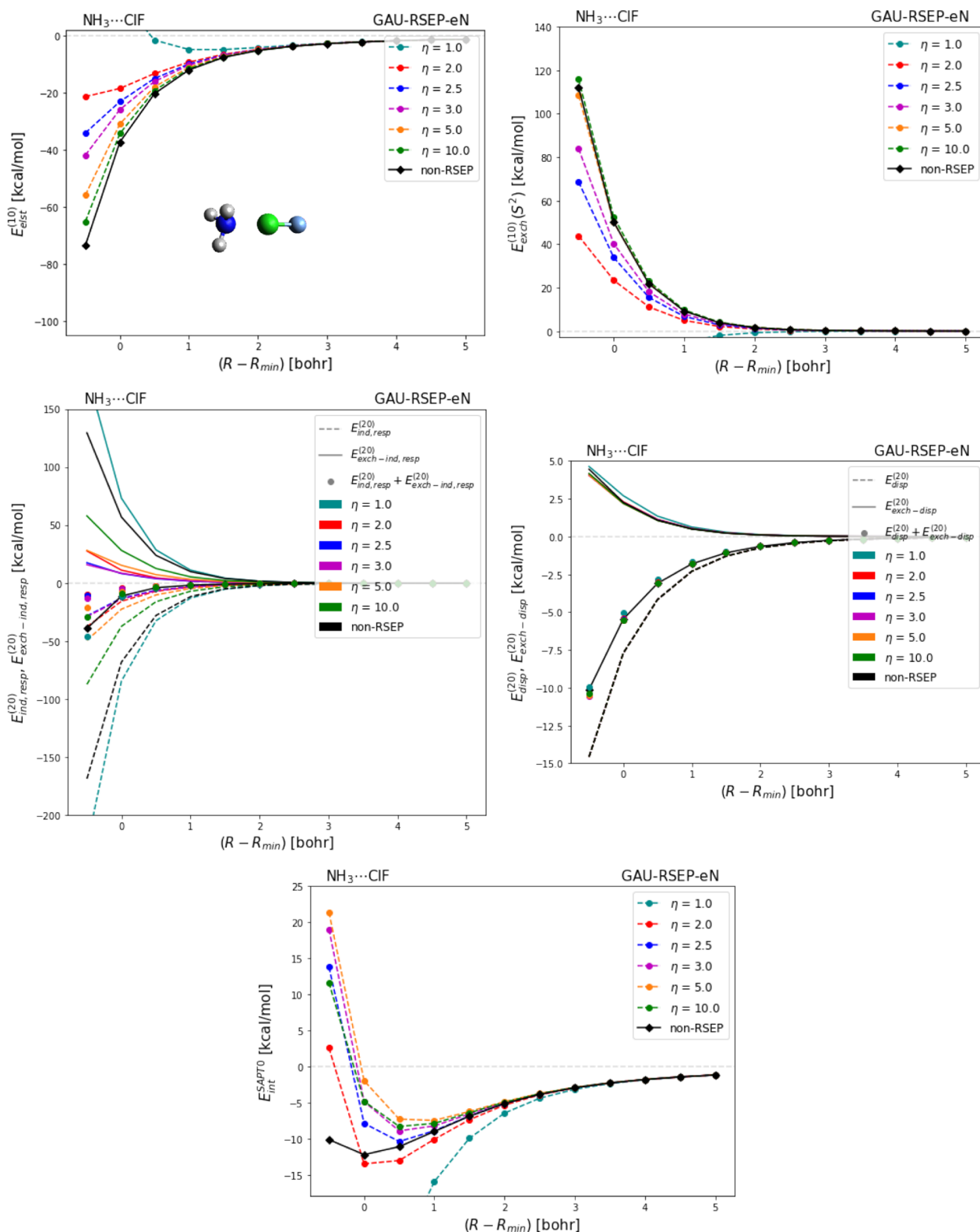


Figure 7. SAPT0(RSEP-eN/gau) interaction energy components in the ammonia-chlorine monofluoride complex computed with various values of η (bohr^{-2}) in the jun-cc-pVDZ basis set. The intermolecular center-of-mass distance R is given relative to the global minimum value $R_{\min} = 5.569$ bohr.

are provided in the [Supporting Information](#). As expected, all range-separated variants converge to the full SAPT0 values at large R . More importantly, however, the long-ranged SAPT terms maintain close agreement with the full SAPT numbers at

intermediate R , down to the van der Waals minimum separation or even slightly closer, except for $\eta = 1.0$ bohr^{-2} and $\omega = 1.0$ bohr^{-1} where the range separation is clearly too strong. This is exactly the physically reasonable range of separation param-

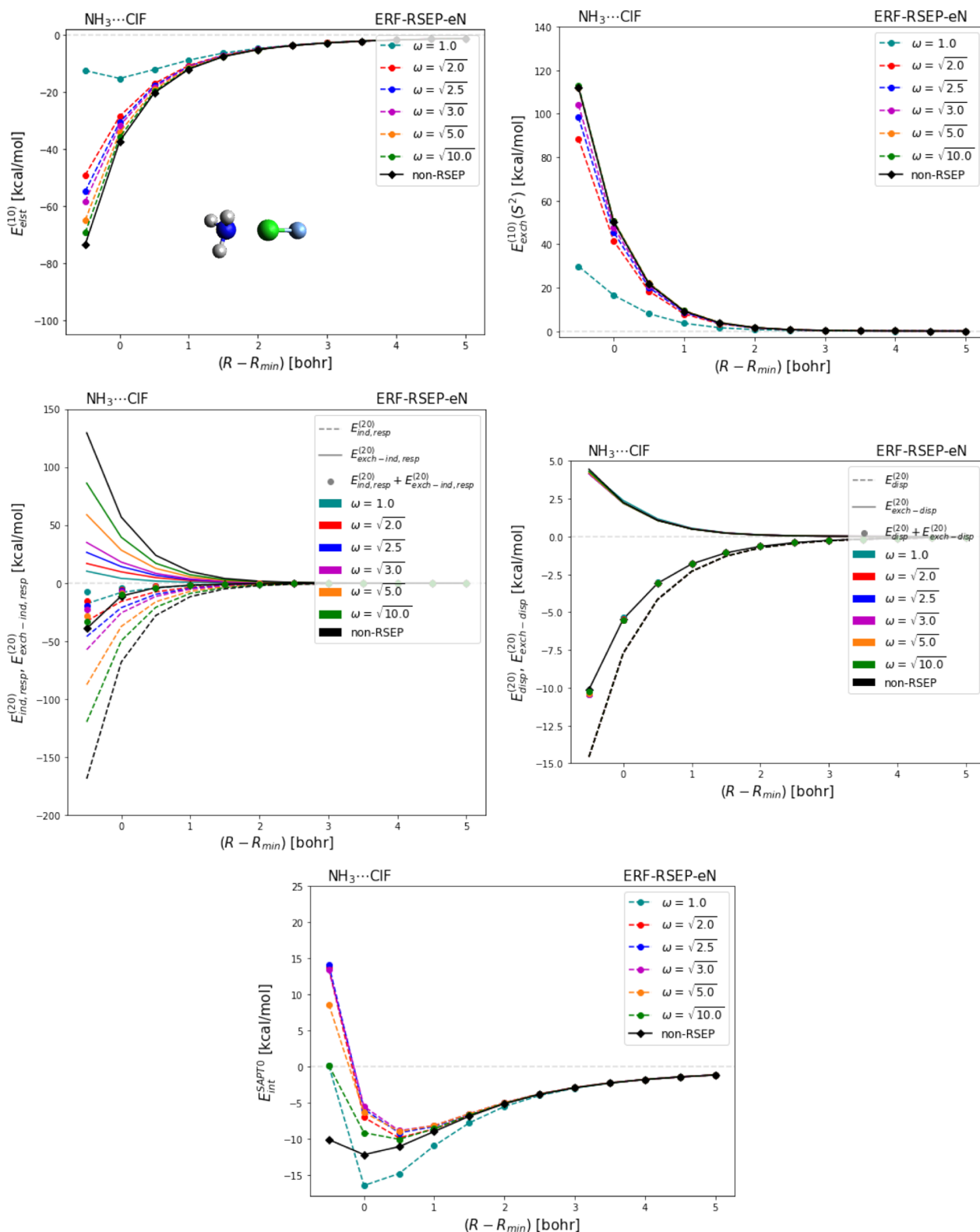


Figure 8. SAPT0(RSEP-eN/erf) interaction energy components in the ammonia-chlorine monofluoride complex computed with various values of ω (bohr^{-1}) in the jun-cc-pVDZ basis set. The intermolecular center-of-mass distance R is given relative to the global minimum value $R_{\min} = 5.569$ bohr.

ters—note that Misquitta suggested $\eta = 3.0 \text{ bohr}^{-2}$ to separate polarization and charge transfer within the R-SRS method.²⁹

The regularized first-order energy corrections show a systematic behavior in all tested systems as the separation

parameter η (ω) is varied. When this parameter approaches infinity, both RSEP and RSEP-eN recover the conventional SAPT0 values as expected. Interestingly, the electrostatic energies converge to the SAPT0 values from below for RSEP

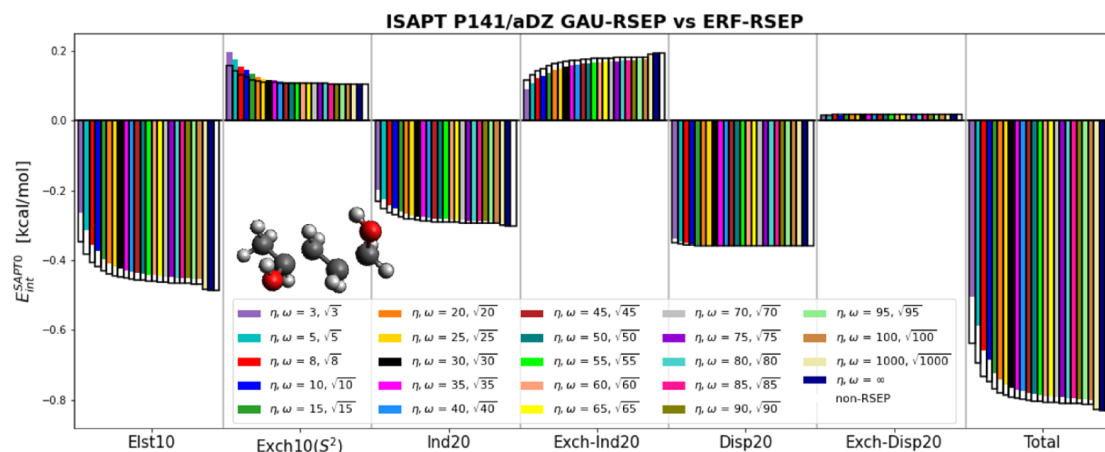


Figure 9. Components of the range-separated ISAPT/SIAO1 interaction energy of the 1,4-pentanediol molecule (fragmented as pictured) computed with different values of η (bohr⁻², colored bars) and ω (bohr⁻¹, empty bars) in the aDZ basis set. The energy results are compared to those of the traditional ISAPT/SIAO1 method.

and from above for RSEP-eN for both range separations, that is, the RSEP method gives more attractive electrostatic energy compared to the conventional value. In all tested models, the long-ranged error function exhibits a faster energy convergence than the Gaussian function. A desired feature of range separation is a relatively minor variance with the separation length parameter, indicating that the long-range Coulomb potential is responsible for nearly the entire noncovalent interaction. From this perspective, considering all $\eta \in [2, \infty)$ bohr⁻² and $\omega \in [\sqrt{2}, \infty)$ bohr⁻¹, the variance of both RSEP and RSEP-eN electrostatic energies is quite minor, with full RSEP performing better for all systems except CH₄...CH₄; see the [Supporting Information](#) for more detail.

The first-order exchange energy for the RSEP-eN scheme converges from below with increasing η (ω), in the opposite direction to electrostatic energy (thus, the range-separated sum $E_{\text{elst}}^{(10)} + E_{\text{exch}}^{(10)}$ exhibits some degree of error cancellation). For the full range separation, the exchange energies for different η (ω) are much closer to each other and to the full SAPT0 value. Nevertheless, the convergence is sometimes nonmonotonic, like for the NH₃...ClF system (the upper right panels in [Figures 5](#) and [6](#)).

The second-order induction and exchange-induction energies, as well as their sum $E_{\text{ind,resp}}^{(20)} + E_{\text{exch-ind,resp}}^{(20)}$, as functions of η (ω) for all three systems are illustrated in the second row of [Figures 1–8](#). The $\eta \rightarrow \infty$ ($\omega \rightarrow \infty$) convergence is mostly monotonic—the induction energy converges from above and the exchange-induction one converges from below, with a possible exception of the $\eta = 1.0$ bohr⁻² and $\omega = 1.0$ bohr⁻¹ data (the latter are highly inaccurate and might even exhibit a wrong sign). These directions of convergence imply that the sum $E_{\text{ind,resp}}^{(20)} + E_{\text{exch-ind,resp}}^{(20)}$ exhibits some error cancellation; indeed, the range-separated sums are not only much smaller in magnitude than the separate $E_{\text{ind,resp}}^{(20)}$ and $E_{\text{exch-ind,resp}}^{(20)}$ terms (as commonly observed in standard SAPT), but their range of variability is usually also much smaller.

For all systems and both range separation schemes, the RSEP induction and exchange-induction terms are highly consistent between different values of η (ω), exhibiting much less variance than RSEP-eN. In contrast, RSEP-eN leads to a smaller variance

for the sum of these two terms for the methane dimer and the ionic bonding system H₂O...F⁻ in both range separation functions. For the other systems, the sum exhibits less variance with the full RSEP range separation: by a minimal amount for H₂O...H₂O and by a significant factor for NH₃...ClF. Overall, the RSEP-eN variant could likely perform better for estimating the charge-transfer energy²⁹ as a larger reduction in magnitude of the mutually canceling induction and exchange-induction terms is afforded. For the purpose of approximating the interaction potential by its long-range term, RSEP appears to offer superior accuracy, in particular for the potentially most difficult NH₃...ClF complex.

We now move on to the range-separated second-order dispersion and exchange-dispersion energies displayed in the middle right panel of [Figures 1–8](#). As expected, all RSEP-eN dispersion terms are virtually identical to the conventional SAPT0 dispersion, and the corresponding RSEP-eN exchange-dispersion data exhibit very little variance with respect to η (ω). This behavior is due to the matrix elements involved in the computation of dispersion corrections not depending on the one-electron integrals. In contrast, the RSEP corrections do exhibit some variance with η (ω), monotonically converging to the conventional SAPT0 data from above for $E_{\text{disp}}^{(20)}$ and from below for $E_{\text{exch-disp}}^{(20)}$. The cancellation of range-separation effects in the sum $E_{\text{disp}}^{(20)} + E_{\text{exch-disp}}^{(20)}$ is reasonable but not quite perfect, and the RSEP sum exhibits some variation with the separation parameter, especially for the Gaussian regularization. In all models, the error function provides a better consistency on the total dispersion energy. Nevertheless, in the practically important range $\eta = 3.0$ bohr⁻² and above, RSEP can produce reasonably accurate (dispersion plus exchange-dispersion) energies at the van der Waals minima and beyond.

The total range-separated SAPT0-level interaction energies for our four example systems are presented at the bottom center of [Figures 1–8](#) and in the [Supporting Information](#). As in [eq 1](#), the total interaction energy includes the delta Hartree–Fock term that is taken here from the calculation without range separation. For all systems and range separation schemes, the differences in electrostatic energy and in first-order exchange energy, partially balancing each other, have the biggest contribution to the differences between range-separated and

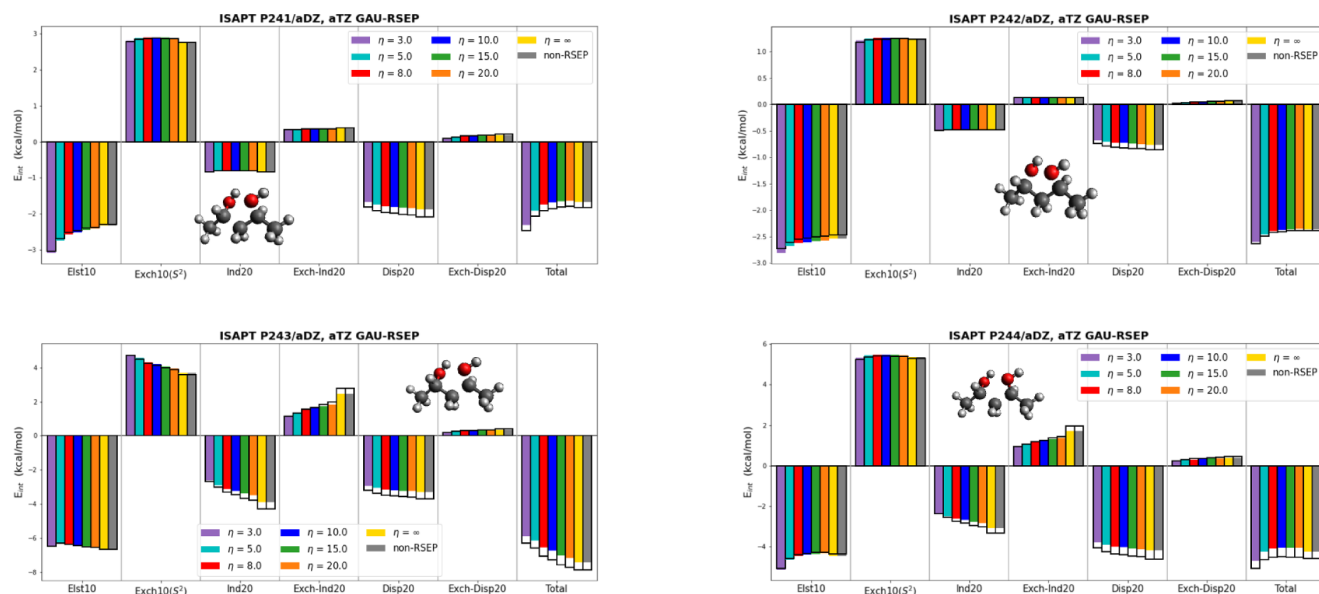


Figure 10. Components of the ISAPT/SIAO1 (RSEP/gau) interaction energy for 4 fragmentation models (pictured) of the 2,4-pentandiol molecule computed with different values of η (bohr⁻²) in the aDZ (colored bars) and aTZ (empty bars) basis sets. The energy results are compared to those of the traditional ISAPT/SIAO1 method.

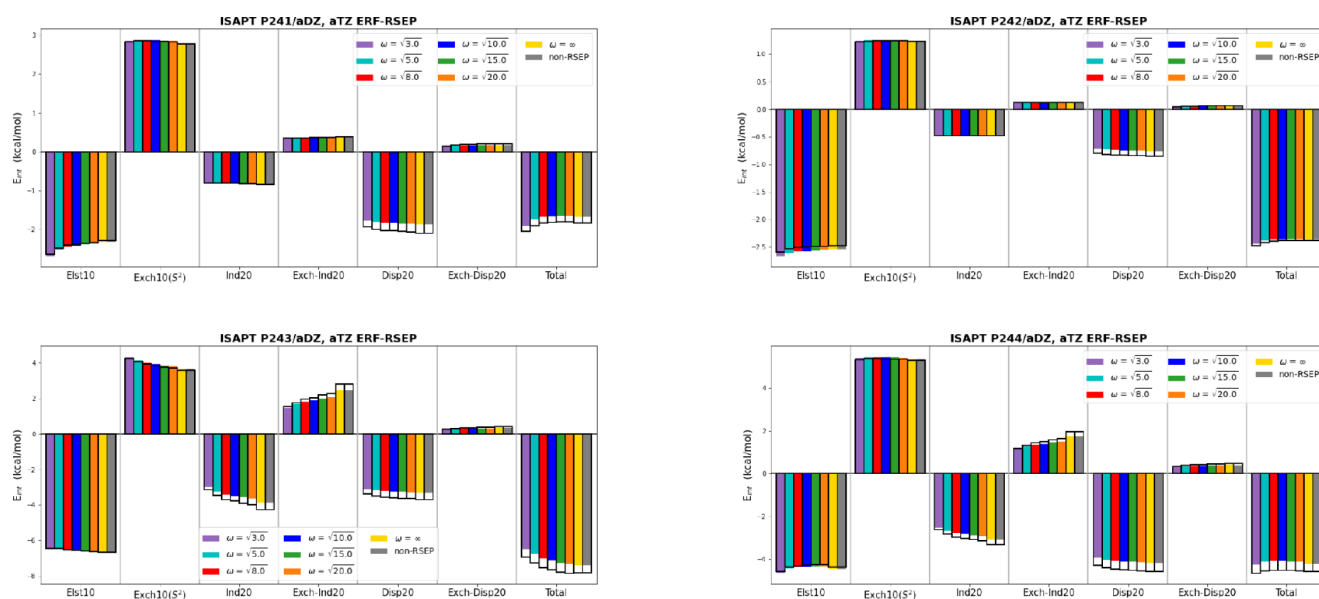


Figure 11. Components of the ISAPT/SIAO1 (RSEP/erf) interaction energy for 4 fragmentation models (pictured) of the 2,4-pentandiol molecule computed with different values of ω (bohr⁻¹) in the aDZ (colored bars) and aTZ (empty bars) basis sets. The energy results are compared to those of the traditional ISAPT/SIAO1 method.

conventional SAPT0. For the $\text{NH}_3 \cdots \text{ClF}$ system at the shortest R , as the monomers get a bit closer to each other than the equilibrium distance, most range separated SAPT variants erroneously give a repulsive total interaction energy. In this case, only the full error function separation gives an attractive or near-zero total interaction energy for all ω . In comparison, in the full Gaussian range separation as well as in the RSEP-eN method and either scheme, the attractive total energy is not observed even when $\eta = 10.0$ bohr⁻² (see Figures 5 and 7). Overall, the full RSEP range separation using the error function provides the best agreement with conventional SAPT0 for a wide range of ω values. The Gaussian range separation with parameter η typically shows somewhat larger total energy discrepancies than the error

function one with $\omega = \sqrt{\eta}$, but still exhibits acceptable accuracy in a wide range of η . The RSEP-eN scheme in general gives larger total energy errors than the RSEP one, indicating that a more balanced description of the attractive and repulsive Coulomb interactions is more accurate than a milder but unbalanced approximation.

The consistency of the range-separated SAPT0 interaction energy components with varied η (ω) is quite promising. As the next step and an even more stringent test, we decided to examine the influence of the purely long-ranged interaction potential V_{1e2e} in the RSEP/gau and RSEP/erf versions, on the accuracy of the ISAPT/SIAO1 intramolecular scheme. The long-ranged one-electron and two-electron potential integrals are applied

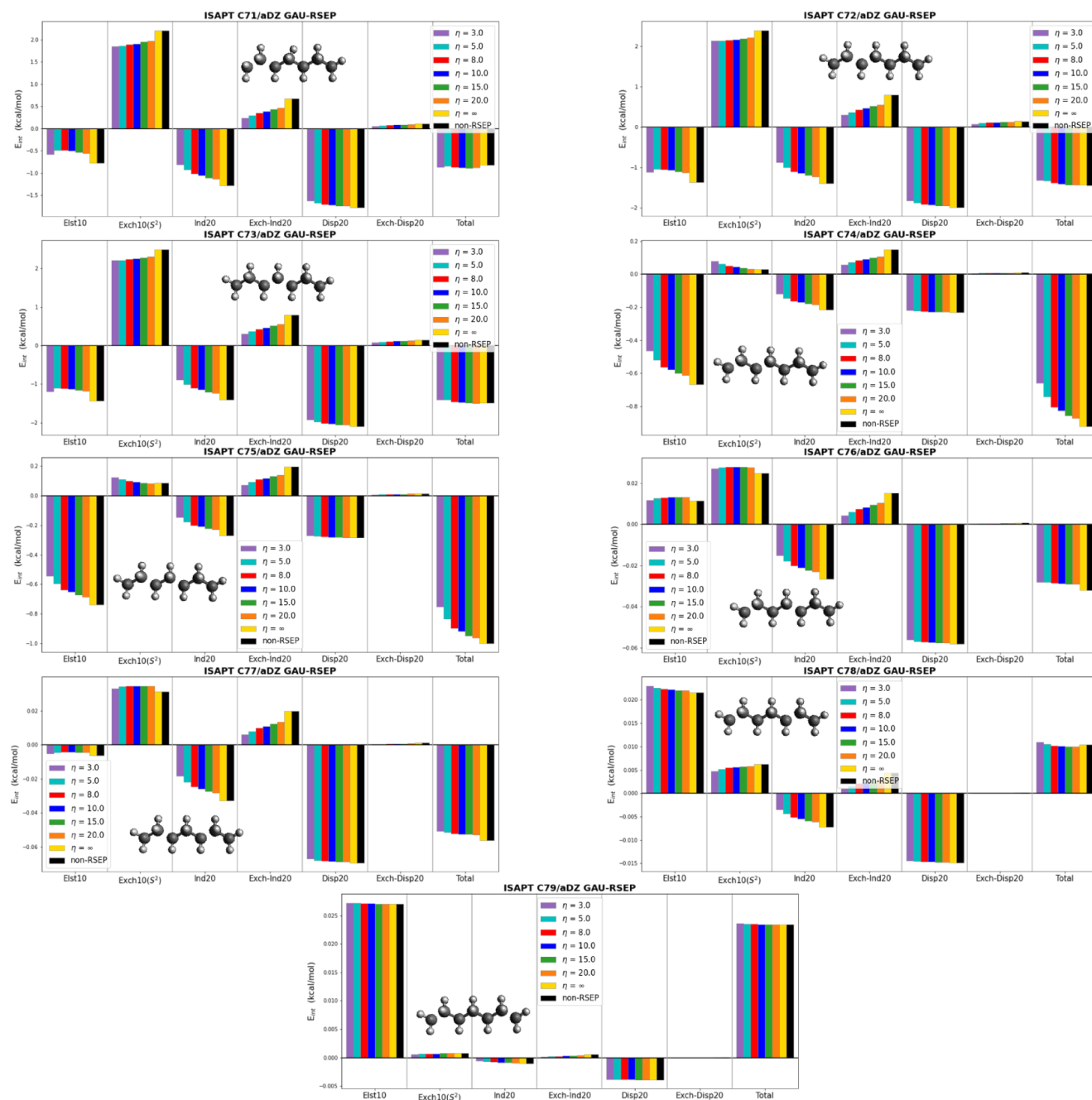


Figure 12. Components of the ISAPT/SIAO1 (RSEP/gau) interaction energy for 9 fragmentation models (pictured) of the *n*-heptane molecule computed with different values of η (bohr⁻²) in the aDZ basis set. The energy results are compared to those of the traditional ISAPT/SIAO1 method.

throughout the calculation of the energy corrections. We investigate a large subset of intramolecular interactions studied in ref 17 including multiple fragmentations of the family of pentanediol (1,4-; 1,5-; 2,4-) molecules as well as *n*-heptane. The resulting convergence of ISAPT/SIAO1 energy contributions as functions of η (ω) is presented in Figures 9–13. Note that the interaction energies for different fragmentation patterns are very different from each other: this is expected as some fragments are closer and others are relatively far away, separated by a large linker.

To perform range-separated intramolecular calculations, we implemented the RSEP/gau and RSEP/erf algorithms into the ISAPT/SIAO1 method in the PSI4 code³² with density fitting.

The most complex part of the implementation was the calculation of long-range generalized Coulomb and exchange (JK) matrices which are pervasive in atomic orbital-basis SAPT expressions. To verify the correctness of our code, we implemented the RSEP algorithm into two density-fitted JK algorithms present in PSI4, the memory-based one (memDF) and the disk-based one (diskDF). The data in Figures 9–13, and for other fragmentation models presented in the Supporting Information, were obtained with the memDF algorithm. We repeated all of the same calculations using an alternative diskDF algorithm and confirmed that all ISAPT energies are the same (the absolute difference in total ISAPT values does not exceed 2.09×10^{-6} kcal/mol). It should be stressed that, while we used

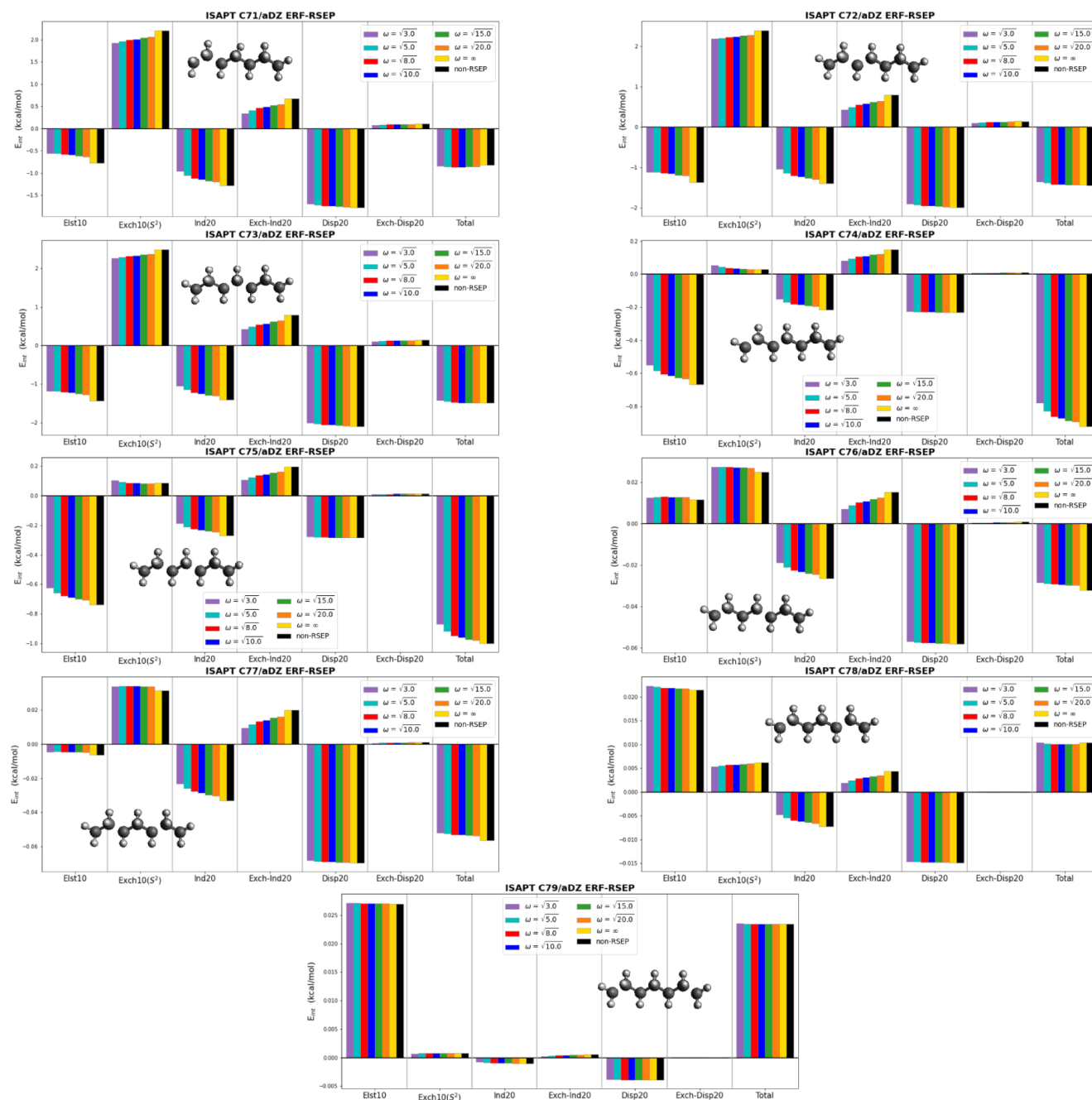


Figure 13. Components of the ISAPT/SIAO1 (RSEP/erf) interaction energy for 9 fragmentation models (pictured) of the *n*-heptane molecule computed with different values of ω (bohr⁻¹) in the aDZ basis set. The energy results are compared to those of the traditional ISAPT/SIAO1 method.

the aDZ orbital basis in all ISAPT calculations, the default aDZ-RI choice for the DF basis used to expand dispersion amplitudes was not always adequate and sometimes led to erratic dispersion energies as a function of η . Therefore, all range-separated ISAPT/SIAO1/aDZ results presented here have been obtained with a larger aTZ-RI dispersion auxiliary basis. The fragmentation pattern names for all systems, given in the header of each figure, follow the definitions of ref 17; in addition, we display the fragmentation in each figure as a ball-and-stick model with the two interfragment bonds removed.

First, we verified that the ISAPT/SIAO1 (RSEP/gau) and ISAPT/SIAO1 (RSEP/erf) interaction energies exhibit the correct limiting behavior at zero and infinity. The energy

corrections collapse to zero with zeroed long-ranged interaction potentials when $\eta = 0.0$ ($\omega = 0.0$), and accurately recover the non-range separated ISAPT/SIAO1 data when $\eta(\omega)$ is on the order of 10 000 000 bohr⁻² (10000000 bohr⁻¹) (shown as $\eta = \infty$, $\omega = \infty$ on energy graphs, Figures 9–13). The interaction energy components exhibit reasonable convergence toward the conventional ISAPT/SIAO1 results, with a somewhat smoother and quicker convergence observed in the RSEP/erf variant. In the ISAPT/SIAO1 (RSEP/gau) method, the calculations using $\eta \leq 5$ bohr⁻² may over- or underestimate the electrostatic energy due to the change of the Coulomb potential being too drastic. The total interaction energy is affected similarly, and its changes with η typically follow the changes in

the electrostatic energy. On the other hand, the ISAPT/SIAO1 (RSEP/erf) electrostatic energy exhibits less variance in ω for all tested models, but still exhibits some small discrepancies when $\omega \leq \sqrt{5}$ bohr⁻¹. Notably, the variance in the electrostatic energy resulting from the small values of $\eta(\omega)$ is unrelated to the size of the linker fragment in all tested models. For example, the linker fragment of the P244 model is $-\text{CH}_2-$; consequently, the occupied orbitals of two fragments $\text{C}_2\text{H}_4\text{OH}-$, which are embedded in the Coulomb and exchange operators of the fragment linker, participate in the calculation of all energy corrections at the SAPT0 level. In contrast, the original ISAPT/C linker partitioning algorithm¹⁶ works mostly well for large linker fragments but fails for small linkers.¹⁷

For the Gaussian splitting with $\eta \geq 8$ bohr⁻² and the error-function splitting with $\omega \geq \sqrt{8}$ bohr⁻¹, a convergent pattern emerges and the electrostatic energies are in qualitative agreement with the conventional values, even though the convergence of the small residual short-range effects can be quite slow. At large η , the difference between $1/r$ and $v_1(r)$ is nonnegligible only at very small r , so the residual effects missing in the long-range electrostatic energy are the consequence of (i) nonzero electron density of one fragment at the location of the other fragment's nuclei, and (ii) nonzero probability of finding a pair of electrons from different fragments at the same location in space. We observed that the effects (i) and (ii) are of the same order of magnitude. The large- η and large- ω convergence in the 1,4-pentanediol case and the P141 fragmentation is illustrated in Figure 9 for the aDZ basis, and in the Supporting Information for the aTZ and aQZ ones. In this case, the effect (i) dominates, resulting in a missing attractive nucleus-electron contribution at finite η and the electrostatic energy converging to its $\eta \rightarrow \infty$ limit from above. For this system, we further broke down the electrostatic interaction energy into its nuclear-electron and electron–electron contributions, illustrating the $\eta(\omega)$ dependence of each term—see the Supporting Information for the relevant figures (the long-range nuclear repulsion term is already converged to the non-range separated value by $\eta = 10$ bohr⁻²). Thus, the slow convergence of the small residual terms in the electrostatic energy is a consequence of the incomplete cancellation between the nuclear-electron and electron–electron residual electrostatic energies, where either effect may prevail. In the comparison to the Gaussian splitting (colored bars in Figure 9), the error-function one (empty bars) produces faster convergence of the electrostatic energy at large ω , with some cancellation between the residual one- and two-electron effects taking place as well (see the Supporting Information for the relevant figures).

Going beyond the electrostatic energy, the convergence of other ISAPT corrections to their non-range separated values is mostly quick, monotonic, and systematic. The induction and exchange-induction terms may slightly underestimate their conventional values even at $\eta = 20$ bohr⁻², although the missing residual effects cancel to a large extent between $E_{\text{ind,resp}}^{(20)}$ and $E_{\text{exch-ind,resp}}^{(20)}$. The convergence of dispersion energy with $\eta(\omega)$ is notably quick and systematic. The first-order exchange term also converges monotonously to its limit, however, a small residual contribution is occasionally still missing at $\eta = 20$ bohr⁻². However, all nonelectrostatic terms vary with the range separation parameter less than the electrostatic energy, so the convergence of the total ISAPT interaction energy to its non-range separated value closely follows the convergence pattern of

$E_{\text{elst}}^{(10)}$. While most of the range-separated ISAPT results have been obtained in the aDZ basis set, for the 2,4-pentanediol system we additionally present the aTZ results (displayed as empty bars in Figures 10 and 11), and the aQZ data in the Supporting Information. As expected, the aTZ data display a somewhat larger magnitude of the induction and dispersion energies as well as their corresponding exchange counterparts, but overall the influence of range separation on the basis set convergence is only minor. In all tested systems, when ω is equivalent to the η value, e.g., $\eta = 20.0$ bohr⁻² versus $\omega = \sqrt{20.0}$ bohr⁻¹, the induction and exchange induction energies have smaller errors when using the error function. Overall, we observe that the neglect of the short-range Coulomb potential still leads to reasonable intramolecular SAPT energy contributions, especially within the error-function range separation. On the other hand, the Gaussian range separation used in the ISAPT/SIAO1 method may prefer larger values of the parameter, $\eta \geq 8$ bohr⁻², than for the analogous intermolecular range-separated SAPT approach.

4. CONCLUSIONS

Symmetry-adapted perturbation theory is commonly used to compute and decompose noncovalent interaction energies between two molecules or molecular fragments. SAPT does provide both a long-range multipole expansion of interaction energy and short-range, overlap-dependent effects. An alternative separation between long-range and short-range interaction, common in the DFT context, is via an explicit splitting of the Coulomb operator $1/r$ into its long-range ($v_l(r)$) and short-range ($v_s(r)$) parts. Such a separation has been used in SAPT as well, but only in the specific contexts of preventing divergence of the perturbation series for a model system²⁸ and separating induction effects into polarization and charge transfer.²⁹ In this work, we follow a different direction and investigate the behavior of intermolecular (SAPT0) and intramolecular (ISAPT) energy terms when the full Coulomb interaction operator is approximated by its long-range part. In a sense, we build on the study of Chen et al.³¹ who performed a similar approximation in various supermolecular interaction energies for rare gas dimers, but we are able to examine each SAPT and ISAPT contribution separately. Two previously proposed forms of splitting the Coulomb operator are employed, a Gaussian range separation and an error-function one.

Our main observation is that the long-range part of the interaction potential $v_l(r)$ is sufficient for a faithful description of noncovalent interactions. This result could not have been taken for granted and should be contrasted with the insufficiency of the long-range multipole expansion, which completely misses both charge penetration contributions and all exchange terms. In our case, the first-order range-separated electrostatic and exchange energies for model intermolecular complexes exhibit smooth convergence to the full non-range separated energies when the separation parameter $\eta(\omega)$ goes to infinity. Moreover, the first-order terms with range separation are close to the nonapproximated values at near-minimum and larger intermolecular separations as long as $\eta \geq 3.0$ bohr⁻² or $\omega \geq \sqrt{3.0}$ bohr⁻¹. It is also notable that the full range separation RSEP overperforms its one-electron-only variant RSEP-eN for most systems, indicating that the approximation can be made more severe (neglecting more terms) as long as the attractive and repulsive contributions are treated in a balanced way. Similarly, the RSEP second-order induction and exchange-induction

terms are highly consistent between different values of η (ω), much more so than the RSEP-eN ones. This often, but not always, translates to a more accurate description of the sum $E_{\text{ind,resp}}^{(20)} + E_{\text{exch-ind,resp}}^{(20)}$ as well. The dispersion and exchange dispersion energies exhibit virtually no dependence on η (ω) in the RSEP-eN variant (as the two-electron integrals are unchanged) and moderate η (ω) dependence for RSEP. As far as total SAPT0(RSEP) energies are concerned, the differences in first-order electrostatic and exchange terms, while partially canceling each other, dominate the residual error of the range-separated approximation; this error is small at the van der Waals minimum and larger separations. Overall, the full RSEP scheme using the error function splitting provides the best agreement with conventional SAPT0.

The application of range separation to intramolecular SAPT (in the SIAO1 linker assignment scheme of ref 17) shows similar smooth and mostly monotonic convergence to the non-range separated values, but due to a stronger interaction, the requirements for the range separation parameter are more strict. The $\eta \leq 5 \text{ bohr}^{-2}$ choice, representing a very crude approximation in the ISAPT case, often overestimates or flips the sign of the electrostatic energy. The total interaction energy is affected similarly, as electrostatic energy is typically the leading contributor to the ISAPT energy changes with η ; all other first- and second-order terms converge faster. Overall, the $\eta \geq 8 \text{ bohr}^{-2}$ values afford reasonably accurate range-separated ISAPT contributions, although small residual errors remain even for $\eta = 20 \text{ bohr}^{-2}$. The error-function range separation provides superior consistency and fairly rapid convergence in all interaction energy components. The relatively small energy variance between ω values illustrates that the long-ranged error-function potential is the optimal selection for the range separation-based ISAPT method.

The results of this work demonstrate that all noncovalent interaction energy contributions can be well described by only the long-range part of the Coulomb potential. When the interaction gets particularly strong, such as for intermolecular complexes at repulsive short-range separations or for some intramolecular complexes, the requirements on the range separation parameter become stricter, but the long-range treatment remains accurate at least qualitatively, and often quantitatively. The use of a purely long-range potential is not fundamentally useful for intermolecular SAPT (in fact, it somewhat slows down the algorithm as two sets of integrals are needed) and the SAPT0 calculations in this work are meant as a proof of principle. However, range separation in intramolecular SAPT is expected to provide an alternative way of separating covalent and noncovalent interactions that is not reliant on a specific molecular fragmentation. This suggests that one may construct a range-separated ISAPT algorithm that yields the entire nonbonded energy within a molecule rather than an interaction energy between two specific fragments. On a separate note, the long-range potentials V_{1e} and V_{1e2e} can also be explored as alternatives to the R-SRS scheme used to partition induction energy into polarization and charge transfer.²⁹ Work in both of these directions is in progress in our group.

APPENDIX

A. Appendix: Range-Separated One-Electron Integrals

The McMurchie-Davidson recursion scheme³⁸ employed in the present work to compute one-electron integrals involving range-separated potentials relies on expanding the products of two

Cartesian basis functions into one-center Hermite Gaussians. Such a Hermite Gaussian with center $\mathbf{P}=(P_x, P_y, P_z)$ and exponent p is defined as⁴²

$$\Lambda_{tuv}(\mathbf{r}, p, \mathbf{P}) = \left(\frac{\partial}{\partial P_x}\right)^t \left(\frac{\partial}{\partial P_y}\right)^u \left(\frac{\partial}{\partial P_z}\right)^v \exp(-p(\mathbf{r} - \mathbf{P})^2) \quad (7)$$

Now, in case of the standard Coulomb potential proportional to $1/|\mathbf{r} - \mathbf{C}|$ ($\mathbf{C}=(C_x, C_y, C_z)$ is the location of the nucleus), the nuclear-attraction integral for a Hermite Gaussian $\Lambda_{tuv}(\mathbf{r}, p, \mathbf{P})$ is expressed as

$$V_{tuv}^{\text{Coul}} = \int \Lambda_{tuv}(\mathbf{r}, p, \mathbf{P}) \frac{1}{|\mathbf{r} - \mathbf{C}|} d\mathbf{r} = \frac{2\pi}{p} R_{tuv}^{0,\text{Coul}}(p, \mathbf{R}_{PC}) \quad (8)$$

where the integral $R_{tuv}^{0,\text{Coul}}(p, \mathbf{R}_{PC})$ is a special case of an auxiliary nuclear attraction integral

$$R_{tuv}^{n,\text{Coul}}(p, \mathbf{R}_{PC}) = \left(\frac{\partial}{\partial P_x}\right)^t \left(\frac{\partial}{\partial P_y}\right)^u \left(\frac{\partial}{\partial P_z}\right)^v R_{000}^{n,\text{Coul}}(p, \mathbf{R}_{PC}) \quad (9)$$

$$R_{000}^{n,\text{Coul}}(p, \mathbf{R}_{PC}) = (-2p)^n F_n(p R_{PC}^2) \quad (10)$$

$$-\mathbf{R}_{PC} = (P_x - C_x, P_y - C_y, P_z - C_z) \equiv (X_{PC}, Y_{PC}, Z_{PC}) \quad (11)$$

(note the convention used above for the sign of the coordinates of $\mathbf{R}_{PC}$). In the above equations, $F_n(x)$ is the Boys function of order n , defined by

$$F_n(x) = \int_0^1 \exp(-xt^2) t^{2n} dt \quad (12)$$

Note that⁴²

$$\frac{dF_n(x)}{dx} = -F_{n+1}(x) \quad (13)$$

The (t, u, v) -order Hermite Gaussian derivatives can be related to the (x, y, z) angular momenta of the product of Cartesian Gaussian basis functions, and are computed from $R_{000}^{n,\text{Coul}}$ using the recursive relations^{38,42}

$$R_{t+1,u,v}^{n,\text{Coul}}(p, \mathbf{R}_{PC}) = t R_{t-1,u,v}^{n+1,\text{Coul}}(p, \mathbf{R}_{PC}) + X_{PC} R_{tuv}^{n+1,\text{Coul}}(p, \mathbf{R}_{PC}) \quad (14)$$

$$R_{t,u+1,v}^{n,\text{Coul}}(p, \mathbf{R}_{PC}) = u R_{t,u-1,v}^{n+1,\text{Coul}}(p, \mathbf{R}_{PC}) + Y_{PC} R_{tuv}^{n+1,\text{Coul}}(p, \mathbf{R}_{PC}) \quad (15)$$

$$R_{t,u,v+1}^{n,\text{Coul}}(p, \mathbf{R}_{PC}) = v R_{t,u,v-1}^{n+1,\text{Coul}}(p, \mathbf{R}_{PC}) + Z_{PC} R_{tuv}^{n+1,\text{Coul}}(p, \mathbf{R}_{PC}) \quad (16)$$

We will now present the analogous recursive formulas for the new one-electron integrals appearing in the two range separation schemes.

B. The Gaussian range separation

The nuclear attraction integral using the $v_s(r) = \frac{e^{-\eta r^2}}{r}$ short-range Gaussian potential is

$$V_{tuv}^{\text{Gau}} = \int \Lambda_{tuv}(\mathbf{r}, p, \mathbf{P}) \frac{e^{-\eta(\mathbf{r}-\mathbf{C})^2}}{|\mathbf{r}-\mathbf{C}|} d\mathbf{r} \\ = \frac{2\pi}{p} R_{tuv}^{0,\text{Gau}} \left(\frac{p^2}{p+\eta}, \mathbf{R}_{PC} \right) \quad (17)$$

where we start the derivation of $R_{tuv}^{0,\text{Gau}}$ by evaluating the s -type integral:

$$R_{000}^{0,\text{Gau}} \left(\frac{p^2}{p+\eta}, \mathbf{R}_{PC} \right) = \frac{p}{p+\eta} \exp \left(\frac{-p\eta}{p+\eta} R_{PC}^2 \right) F_0 \left(\frac{p^2}{p+\eta} R_{PC}^2 \right) \quad (18)$$

Now, incrementing over t , we get

$$R_{100}^{0,\text{Gau}} \left(\frac{p^2}{p+\eta}, \mathbf{R}_{PC} \right) = \frac{\partial}{\partial p_x} R_{000}^{0,\text{Gau}} \left(\frac{p^2}{p+\eta}, \mathbf{R}_{PC} \right) \\ = \frac{p}{p+\eta} F_0 \left(\frac{p^2}{p+\eta} R_{PC}^2 \right) \exp \left(\frac{-p\eta}{p+\eta} R_{PC}^2 \right) \left(\frac{-p\eta}{p+\eta} \right) (2X_{PC}) \\ + \frac{p}{p+\eta} \exp \left(\frac{-p\eta}{p+\eta} R_{PC}^2 \right) \left(-F_1 \left(\frac{p^2}{p+\eta} R_{PC}^2 \right) \right) \frac{p^2}{p+\eta} (2X_{PC}) \\ = \left(\frac{-2p\eta}{p+\eta} \right) R_{000}^{0,\text{Gau}} \left(\frac{p^2}{p+\eta}, \mathbf{R}_{PC} \right) X_{PC} + R_{000}^{1,\text{Gau}} \left(\frac{p^2}{p+\eta}, \mathbf{R}_{PC} \right) X_{PC} \\ = X_{PC} \left[\left(\frac{-2p\eta}{p+\eta} \right) R_{000}^{0,\text{Gau}} \left(\frac{p^2}{p+\eta}, \mathbf{R}_{PC} \right) + R_{000}^{1,\text{Gau}} \left(\frac{p^2}{p+\eta}, \mathbf{R}_{PC} \right) \right] \quad (19)$$

$$R_{t+1,0,0}^{0,\text{Gau}} \left(\frac{p^2}{p+\eta}, \mathbf{R}_{PC} \right) = \left(\frac{\partial}{\partial p_x} \right)^t \frac{\partial}{\partial p_x} R_{000}^{0,\text{Gau}} \left(\frac{p^2}{p+\eta}, \mathbf{R}_{PC} \right) \\ = \left(\frac{\partial}{\partial p_x} \right)^t R_{100}^{0,\text{Gau}} \left(\frac{p^2}{p+\eta}, \mathbf{R}_{PC} \right) \\ = \left(\frac{\partial}{\partial p_x} \right)^t \left[X_{PC} \left(\frac{-2p\eta}{p+\eta} R_{000}^{0,\text{Gau}} + R_{000}^{1,\text{Gau}} \right) \right] \\ = t \left(\frac{\partial}{\partial p_x} \right)^{t-1} \left(\frac{-2p\eta}{p+\eta} R_{000}^{0,\text{Gau}} + R_{000}^{1,\text{Gau}} \right) + X_{PC} \left(\frac{\partial}{\partial p_x} \right)^t \left(\frac{-2p\eta}{p+\eta} R_{000}^{0,\text{Gau}} + R_{000}^{1,\text{Gau}} \right) \\ = t \left(\frac{-2p\eta}{p+\eta} \right) R_{t-1,0,0}^{0,\text{Gau}} + t R_{t-1,0,0}^{1,\text{Gau}} + X_{PC} \left(\frac{-2p\eta}{p+\eta} \right) R_{t,0,0}^{0,\text{Gau}} + X_{PC} R_{t,0,0}^{1,\text{Gau}} \quad (20)$$

Note that we will sometimes skip the arguments of $R_{tuv}^{n,\text{Gau}}$ as they are always the same; later, we will do the same thing for $R_{tuv}^{n,\text{Erf}}$. The resulting angular-momentum recursion relations over t, u, v are

$$R_{t+1,u,v}^{n,\text{Gau}} \left(\frac{p^2}{p+\eta}, \mathbf{R}_{PC} \right) = t \left(\frac{-2p\eta}{p+\eta} \right) R_{t-1,u,v}^{n,\text{Gau}} + t R_{t-1,u,v}^{n+1,\text{Gau}} \\ + X_{PC} \left(\frac{-2p\eta}{p+\eta} \right) R_{tuv}^{n,\text{Gau}} + X_{PC} R_{tuv}^{n+1,\text{Gau}} \quad (21)$$

$$R_{t,u+1,v}^{n,\text{Gau}} \left(\frac{p^2}{p+\eta}, \mathbf{R}_{PC} \right) = u \left(\frac{-2p\eta}{p+\eta} \right) R_{t,u-1,v}^{n,\text{Gau}} + u R_{t,u-1,v}^{n+1,\text{Gau}} \\ + Y_{PC} \left(\frac{-2p\eta}{p+\eta} \right) R_{tuv}^{n,\text{Gau}} + Y_{PC} R_{tuv}^{n+1,\text{Gau}} \quad (22)$$

$$R_{t,u,v+1}^{n,\text{Gau}} \left(\frac{p^2}{p+\eta}, \mathbf{R}_{PC} \right) = v \left(\frac{-2p\eta}{p+\eta} \right) R_{t,u,v-1}^{n,\text{Gau}} + v R_{t,u,v-1}^{n+1,\text{Gau}} \\ + Z_{PC} \left(\frac{-2p\eta}{p+\eta} \right) R_{tuv}^{n,\text{Gau}} + Z_{PC} R_{tuv}^{n+1,\text{Gau}} \quad (23)$$

with the following s -type auxiliary nuclear attraction integrals required:

$$R_{000}^{n,\text{Gau}} \left(\frac{p^2}{p+\eta}, \mathbf{R}_{PC} \right) \\ = \frac{p}{p+\eta} \exp \left(\frac{-p\eta}{p+\eta} R_{PC}^2 \right) \left(\frac{-2p^2}{p+\eta} \right)^n F_n \left(\frac{p^2}{p+\eta} R_{PC}^2 \right) \quad (24)$$

C. The error function range separation

The nuclear attraction integral using the $v_l(r) = \frac{\text{erf}(\omega r)}{r}$ long-range error function potential is

$$V_{tuv}^{\text{Erf}} = \int \Lambda_{tuv}(\mathbf{r}, p, \mathbf{P}) \frac{\text{erf}(\omega|\mathbf{r}-\mathbf{C}|)}{|\mathbf{r}-\mathbf{C}|} d\mathbf{r} \\ = \frac{2\pi}{p} R_{tuv}^{0,\text{Erf}} \left(\frac{p\omega^2}{p+\omega^2}, \mathbf{R}_{PC} \right) \quad (25)$$

where we once again start the derivation of $R_{tuv}^{0,\text{Erf}}$ from the s -type integral:

$$R_{000}^{0,\text{Erf}} \left(\frac{p\omega^2}{p+\omega^2}, \mathbf{R}_{PC} \right) = \left(\frac{\omega^2}{p+\omega^2} \right)^{1/2} F_0 \left(\frac{p\omega^2}{p+\omega^2} R_{PC}^2 \right) \quad (26)$$

Then, incrementing over t ,

$$R_{100}^{0,\text{Erf}} \left(\frac{p\omega^2}{p+\omega^2}, \mathbf{R}_{PC} \right) = \frac{\partial}{\partial p_x} R_{000}^{0,\text{Erf}} \left(\frac{p\omega^2}{p+\omega^2}, \mathbf{R}_{PC} \right) \\ = \left(\frac{\omega^2}{p+\omega^2} \right)^{1/2} \left(-F_1 \left(\frac{p\omega^2}{p+\omega^2} R_{PC}^2 \right) \right) \frac{p\omega^2}{p+\omega^2} (2X_{PC}) \\ = X_{PC} R_{000}^{1,\text{Erf}} \left(\frac{p\omega^2}{p+\omega^2}, \mathbf{R}_{PC} \right) \quad (27)$$

$$R_{t+1,0,0}^{0,\text{Erf}} \left(\frac{p\omega^2}{p+\omega^2}, \mathbf{R}_{PC} \right) = \left(\frac{\partial}{\partial p_x} \right)^t \frac{\partial}{\partial p_x} R_{000}^{0,\text{Erf}} \left(\frac{p\omega^2}{p+\omega^2}, \mathbf{R}_{PC} \right) \\ = \left(\frac{\partial}{\partial p_x} \right)^t R_{100}^{0,\text{Erf}} \left(\frac{p\omega^2}{p+\omega^2}, \mathbf{R}_{PC} \right) \\ = \left(\frac{\partial}{\partial p_x} \right)^t X_{PC} R_{000}^{1,\text{Erf}} \\ = t \left(\frac{\partial}{\partial p_x} \right)^{t-1} R_{000}^{1,\text{Erf}} + X_{PC} \left(\frac{\partial}{\partial p_x} \right)^t R_{000}^{1,\text{Erf}} \\ = t R_{t-1,0,0}^{1,\text{Erf}} + X_{PC} R_{t,0,0}^{1,\text{Erf}} \quad (28)$$

which leads to the following recursion relations over t, u, v :

$$R_{t+1,u,v}^{n,\text{Erf}} \left(\frac{p\omega^2}{p+\omega^2}, \mathbf{R}_{PC} \right) = t R_{t-1,u,v}^{n+1,\text{Erf}} + X_{PC} R_{tuv}^{n+1,\text{Erf}} \quad (29)$$

$$R_{t,u+1,v}^{n,\text{Erf}}\left(\frac{p\omega^2}{p+\omega^2}, \mathbf{R}_{\text{PC}}\right) = uR_{t,u-1,v}^{n+1,\text{Erf}} + Y_{\text{PC}}R_{tuv}^{n+1,\text{Erf}} \quad (30)$$

$$R_{t,u,v+1}^{n,\text{Erf}}\left(\frac{p\omega^2}{p+\omega^2}, \mathbf{R}_{\text{PC}}\right) = vR_{t,u,v-1}^{n+1,\text{Erf}} + Z_{\text{PC}}R_{tuv}^{n+1,\text{Erf}} \quad (31)$$

with the following *s*-type auxiliary integrals:

$$R_{000}^{n,\text{Erf}}\left(\frac{p\omega^2}{p+\omega^2}, \mathbf{R}_{\text{PC}}\right) = \left(\frac{\omega^2}{p+\omega^2}\right)^{1/2} \left(\frac{-2p\omega^2}{p+\omega^2}\right)^n \mathbf{F}_n\left(\frac{p\omega^2}{p+\omega^2} R_{\text{PC}}^2\right) \quad (32)$$

Equations 29–32 are sufficient to recursively compute the integrals $R_{tuv}^{0,\text{Erf}}$ to arbitrary *t*, *u*, *v*, and eqs 21–24 enable doing the same for $R_{tuv}^{0,\text{Gau}}$. Thus, we can now compute all required range-separated nuclear attraction integrals, eqs 17 and 25, in terms of one-center Hermite Gaussians. The transformation of integrals obtained in this way to the Cartesian Gaussian basis proceeds in exactly the same way as for standard non-range separated nuclear attraction integrals computed within the McMurchie-Davidson scheme.

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jctc.4c00608>.

Figures displaying range-separated SAPT0 and ISAPT data for additional systems and basis sets (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Konrad Patkowski – Department of Chemistry and Biochemistry, Auburn University, Auburn, Alabama 36849, United States; orcid.org/0000-0002-4468-207X; Email: patkowsk@auburn.edu

Authors

Du Luu – Department of Chemistry and Biochemistry, Auburn University, Auburn, Alabama 36849, United States
Clemence Corminboeuf – Laboratory for Computational Molecular Design, Institute of Chemical Sciences and Engineering, École Polytechnique Fédérale de Lausanne, Lausanne CH-1015, Switzerland; orcid.org/0000-0001-7993-2879

Complete contact information is available at: <https://pubs.acs.org/10.1021/acs.jctc.4c00608>

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

D.L. and K.P. are supported by the U.S. National Science Foundation award CHE-1955328. C.C. and K.P. thank the Swiss National Science Foundation (Grant IZSEZO_180167) for support.

■ REFERENCES

- (1) Jeziorski, B.; Moszyński, R.; Szalewicz, K. Perturbation Theory Approach to Intermolecular Potential Energy Surfaces of van der Waals Complexes. *Chem. Rev.* **1994**, *94*, 1887–1930.
- (2) Szalewicz, K. Symmetry-adapted perturbation theory of intermolecular forces. *WIREs Comput. Mol. Sci.* **2012**, *2*, 254–272.
- (3) Szalewicz, K.; Patkowski, K.; Jeziorski, B. Intermolecular Interactions via Perturbation Theory: From Diatoms to Biomolecules. *Intermolecular Forces and Clusters II*, Springer, 2005, 116, 43–117.
- (4) Hohenstein, E. G.; Sherrill, C. D. Wavefunction Methods for Noncovalent Interactions. *WIREs Comput. Mol. Sci.* **2012**, *2*, 304–326.
- (5) Patkowski, K. Recent developments in symmetry-adapted perturbation theory. *WIREs Comput. Mol. Sci.* **2020**, *10*, No. e1452.
- (6) Rybak, S.; Jeziorski, B.; Szalewicz, K. Many-body symmetry-adapted perturbation theory of intermolecular interactions. H₂O and HF dimers. *J. Chem. Phys.* **1991**, *95*, 6576–6601.
- (7) Parker, T. M.; Burns, L. A.; Parrish, R. M.; Ryno, A. G.; Sherrill, C. D. Levels of Symmetry Adapted Perturbation Theory (SAPT). I. Efficiency and performance for interaction energies. *J. Chem. Phys.* **2014**, *140*, 094106.
- (8) Hesselmann, A.; Jansen, G.; Schütz, M. Density-functional theory-symmetry-adapted intermolecular perturbation theory with density fitting: A new efficient method to study intermolecular interaction energies. *J. Chem. Phys.* **2005**, *122*, 014103.
- (9) Misquitta, A. J.; Podeszwa, R.; Jeziorski, B.; Szalewicz, K. Intermolecular potentials based on symmetry-adapted perturbation theory with dispersion energies from time-dependent density-functional calculations. *J. Chem. Phys.* **2005**, *123*, 214103.
- (10) Hapka, M.; Przybytek, M.; Pernal, K. Second-Order Exchange-Dispersion Energy Based on a Multireference Description of Monomers. *J. Chem. Theory Comput.* **2019**, *15*, 6712–6723.
- (11) Patkowski, K.; Żuchowski, P. S.; Smith, D. G. A. First-order symmetry-adapted perturbation theory for multiplet splittings. *J. Chem. Phys.* **2018**, *148*, 164110.
- (12) Korona, T.; Hapka, M.; Pernal, K.; Patkowski, K., *Polish Quantum Chemistry from Kolos to Now*, Musiał, M.; Grabowski, I., Eds.; Elsevier, 2023, pp 37–72.
- (13) Gonthier, J. F.; Corminboeuf, C. Exploration of zeroth-order wavefunctions and energies as a first step toward intramolecular symmetry-adapted perturbation theory. *J. Chem. Phys.* **2014**, *140*, 154107.
- (14) Pastorczak, E.; Prlj, A.; Gonthier, J. F.; Corminboeuf, C. Intramolecular symmetry-adapted perturbation theory with a single-determinant wavefunction. *J. Chem. Phys.* **2015**, *143*, 224107.
- (15) Meitei, O. R.; Hesselmann, A. Intramolecular Interactions in Sterically Crowded Hydrocarbon Molecules. *J. Comput. Chem.* **2017**, *38*, 2500–2508.
- (16) Parrish, R. M.; Gonthier, J. F.; Corminboeuf, C.; Sherrill, C. D. Communication: Practical intramolecular symmetry adapted perturbation theory via Hartree-Fock embedding. *J. Chem. Phys.* **2015**, *143*, 051103.
- (17) Luu, D.; Patkowski, K. Overcoming Artificial Multipoles in Intramolecular Symmetry-Adapted Perturbation Theory. *J. Phys. Chem. A* **2023**, *127*, 356–377.
- (18) Schreiner, P. R.; Chernish, L. V.; Gunchenko, P. A.; Tikhonchuk, E. Y.; Hausmann, H.; Serafin, M.; Schlecht, S.; Dahl, J. E. P.; Carlson, R. M. K.; Fokin, A. A. Overcoming lability of extremely long alkane carbon-carbon bonds through dispersion forces. *Nature* **2011**, *477*, 308–311.
- (19) Leininger, T.; Stoll, H.; Werner, H.-J.; Savin, A. Combining Long-Range Configuration Interaction with Short-Range Density Functionals. *Chem. Phys. Lett.* **1997**, *275*, 151–160.
- (20) Iikura, H.; Tsuneda, T.; Yanai, T.; Hirao, K. A long-range correction scheme for generalized-gradient-approximation exchange functionals. *J. Chem. Phys.* **2001**, *115*, 3540–3544.
- (21) Vydrov, O. A.; Scuseria, G. E. Assessment of a Long-Range Corrected Hybrid Functional. *J. Chem. Phys.* **2006**, *125*, 234109.

- (22) Lao, K. U.; Herbert, J. M. Symmetry-adapted perturbation theory with Kohn-Sham orbitals using non-empirically tuned, long-range-corrected density functionals. *J. Chem. Phys.* **2014**, *140*, 044108.
- (23) Gray, M.; Herbert, J. M. Simplified tuning of long-range corrected density functionals for use in symmetry-adapted perturbation theory. *J. Chem. Phys.* **2021**, *155*, 034103.
- (24) Xie, Y.; Smith, D. G. A.; Sherrill, C. D. Implementation of symmetry-adapted perturbation theory based on density functional theory and using hybrid exchange–correlation kernels for dispersion terms. *J. Chem. Phys.* **2022**, *157*, 024801.
- (25) Heyd, J.; Scuseria, G. E.; Ernzerhof, M. Hybrid functionals based on a screened Coulomb potential. *J. Chem. Phys.* **2003**, *118*, 8207–8215.
- (26) Patkowski, K.; Jeziorski, B.; Szalewicz, K. Symmetry-Adapted Perturbation Theory with Regularized Coulomb Potential. *J. Mol. Struct.* **2001**, *547*, 293–307.
- (27) Adams, W. H. Two new symmetry-adapted perturbation theories for the calculation of intermolecular interaction energies. *Theor. Chem. Acc.* **2002**, *108*, 225–231.
- (28) Patkowski, K.; Jeziorski, B.; Szalewicz, K. Unified treatment of chemical and van der Waals forces via symmetry-adapted perturbation expansion. *J. Chem. Phys.* **2004**, *120*, 6849–6862.
- (29) Misquitta, A. J. Charge Transfer from Regularized Symmetry-Adapted Perturbation Theory. *J. Chem. Theory Comput.* **2013**, *9*, 5313–5326.
- (30) Toulouse, J.; Colonna, F.; Savin, A. Long-range–short-range separation of the electron-electron interaction in density-functional theory. *Phys. Rev. A* **2004**, *70*, 062505.
- (31) Chen, Y.-T.; Hui, K.; Chai, J.-D. The van der Waals interactions in rare-gas dimers: The role of interparticle interactions. *Phys. Chem. Chem. Phys.* **2016**, *18*, 3011–3022.
- (32) Smith, D. G. A.; Burns, L. A.; Simmonett, A. C.; Parrish, R. M.; Schieber, M. C.; Galvelis, R.; Kraus, P.; Kruse, H.; Di Remigio, R.; Alenaizan, A.; et al. Psi4 1.4: Open-source software for high-throughput quantum chemistry. *J. Chem. Phys.* **2020**, *152*, 184108.
- (33) Korona, T. Exchange-Dispersion Energy: A Formulation in Terms of Monomer Properties and Coupled Cluster Treatment of Intramonomer Correlation. *J. Chem. Theory Comput.* **2009**, *5*, 2663–2678.
- (34) Adams, W. H. Perturbation theory of intermolecular interactions: Are second-order Rayleigh-Schrödinger energies meaningful. *Int. J. Quantum Chem.* **1991**, *S25*, 165–181.
- (35) Knizia, G. Intrinsic Atomic Orbitals: An Unbiased Bridge between Quantum Theory and Chemical Concepts. *J. Chem. Theory Comput.* **2013**, *9*, 4834–4843.
- (36) Hättig, C.; Klopper, W.; Köhn, A.; Tew, D. P. Explicitly Correlated Electrons in Molecules. *Chem. Rev.* **2012**, *112*, 4–74.
- (37) Kong, L.; Bischoff, F. A.; Valeev, E. F. Explicitly Correlated R12/F12 Methods for Electronic Structure. *Chem. Rev.* **2012**, *112*, 75–107.
- (38) McMurchie, L. E.; Davidson, E. R. Calculation of Integrals over ab initio Pseudopotentials. *J. Comput. Phys.* **1981**, *44*, 289–301.
- (39) Obara, S.; Saika, A. Efficient recursive computation of molecular integrals over Cartesian Gaussian functions. *J. Chem. Phys.* **1986**, *84*, 3963–3974.
- (40) Smith, D. G. A.; Burns, L. A.; Sirianni, D. A.; Nascimento, D. R.; Kumar, A.; James, A. M.; Schriber, J. B.; Zhang, T.; Zhang, B.; Abbott, A. S.; Berquist, E. J.; Lechner, M. H.; Cunha, L. A.; Heide, A. G.; Waldrop, J. M.; Takeshita, T. Y.; Alenaizan, A.; Neuhauser, D.; King, R. A.; Simmonett, A. C.; Turney, J. M.; Schaefer, H. F.; Evangelista, F. A.; DePrince III, A. E.; Crawford, T. D.; Patkowski, K.; Sherrill, C. D. Psi4NUMPy: An Interactive Quantum Chemistry Programming Environment for Reference Implementations and Rapid Development. *J. Chem. Theory Comput.* **2018**, *14*, 3504–3511.
- (41) Jurečka, P.; Šponer, J.; Černý, J.; Hobza, P. Benchmark database of accurate (MP2 and CCSD(T) complete basis set limit) interaction energies of small model complexes, DNA base pairs, and amino acid pairs. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985–1993.
- (42) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory*; Wiley: New York, 2000.

Chapter 5

The Application of Intramolecular-SAPT to Z/E-Azobenzene Isomerization

Azobenzene, composed of two phenyl rings connected by an azo (--N=N--) bond, serves as the foundational molecule for a wide range of aromatic azo compounds. Azobenzene and its derivatives are molecules considered to be prototype photomodulable materials and molecular devices. [8, 18, 36, 37, 41] The photoresponsive property of azobenzene arises from the efficient and reversible $\text{trans} \rightarrow \text{cis}$ photoisomerization in response to the external stimuli, such as light, electrons, and temperature. The photoisomerization occurs without secondary processes, making azobenzene a highly versatile molecular switch. [5, 7, 10, 12, 59] A broad range of molecular design strategies in azobenzene photochemistry, particularly through substitutions on the aromatic rings, enables the fine-tuning of its isomerization mechanism for diverse applications, including light-responsive polymers, photochemical control agents, optical memories and switches, photoactuators, nanotechnology, and molecular machines. [4, 10, 12, 14, 15, 20, 33, 60, 68, 70, 72]

The photochromic behavior of azobenzene has spurred extensive experimental and theoretical research. The Z/E photoisomerization is primarily driven by electronic transitions and steric effects, as the molecule absorbs light at specific wavelengths, leading to distinct absorption bands in the UV-visible spectrum. Upon light irradiation, azobenzene undergoes photoisomerization via two main mechanisms: N=N double bond torsion and N-inversion. These pathways correspond to the electron excitation of the N=N bond that falls into the $\pi \rightarrow \pi^*$ band or the $n \rightarrow \pi^*$ band, respectively. The torsion mechanism is more prominent when the molecule is excited by UV light at around 320 nm, whereas the inversion mechanism typically occurs upon excitation with visible light around 440 nm. [5, 7, 12, 33, 56, 57]

From a molecular modeling perspective, the contribution of London dispersion forces is crucial to determine the overall energetics and relative stabilization of the cis and trans isomers. The dispersion interactions between the phenyl rings significantly influence the energy barriers involved in the photoisomerization and stabilization of the two conformers. [13,24,41,64,71] The advances in the intramolecular dispersion-inclusive computational methods, such as dispersion-corrected density functional theory (DFT), time-dependent DFT (TD-DFT), and ab initio molecular dynamics (AIMD), have provided powerful frameworks for investigating and modeling the isomerization pathways. [3,6,11,13,24,56,71] Equivalently, the quantification of the intramolecular interaction energy between the cis and trans isomers accounts for making precise predictions about the reaction barriers and molecular stability. In this work, the Intramolecular Symmetry-Adapted Perturbation Theory (ISAPT) method [39,49] is employed to study the intramolecular interaction effects between the azobenzene isomers. ISAPT is an ab initio computational method that provides physically meaningful intramolecular interaction energies without requiring geometric pretreatment. By preserving the fully intact molecular structure, ISAPT offers deeper insights into how dispersion interactions, as well as other types of physical interaction, modulate the energy landscapes of the azobenzene's isomerization process.

The molecular structures chosen for the study are inspired by the measurement of the strength of the London dispersion donors on the incremental system of the substituted azobenzene molecule in Ref. [6]. The cis and trans conformers of all-meta alkyl substituted azobenzene geometries, including 10 models varied from Me, Et, iPr up to tBu substitutions, in Figs. 5.1-5.2, are optimized using the aug-cc-pVDZ basis set at the MP2 level of theory. The cis conformation of the iPr-tBu azobenzene system is under optimization. The bulky geometry of the all-meta tBu-substituted azobenzene is optimized using the B3LYP/aug-cc-pVDZ method. The azobenzene isomerization energy, in Table 5.4, is examined by the sum of the ISAPT/SIAO1 conformational energy and the DLPNO-CCSD(T) deformational energy of cis $\rightarrow$ trans isomers. The cis $\rightarrow$ trans isomerization accelerates under the influence of the size of the alkyl substituents. The larger substituents have a faster isomerization rate as a results of less thermal stability that due to the smaller energy gap between the cis and trans configurations.

In the ISAPT/SIAO1 energy decomposition of 10 models of all-meta alkyl substituted azobenzene molecule in Table 5.1, overall, the cis conformations of the azobenzene models provide a smaller attraction of the total interaction energy than the trans ones. The cis structure is mainly bound by the dispersion energy due to the strong repulsion of two phenyl rings that is balanced out by the attractive electrostatic and induction energy terms. In the family of the all-meta alkyl substituted azobenzene molecules, the stabilization of the cis conformations is proportional to the size of the alkyl substituents, increasing from Me to Et to iPr to tBu. The trans structure is stabilized by the electrostatic energy, following by the dispersion energy, because of the far distance between the alkyl substituents of two phenyl rings. The planar structure of the trans conformer provides less repulsive values of the exchange energy corrections, leading to more conformational stabilization compared to the cis conformer. In comparison to the original ISAPT/C method, in Tables 5.1-5.2, for the cis conformation, the magnitude of first-order exchange, second-order induction, and exchange induction energies is significantly reduced in the ISAPT/SIAO1 method. The strongly repulsive net energy in the ISAPT/C is adjusted to be attractive in the ISAPT/SIAO1. The larger alkyl substituents provide the stronger attractive ISAPT/SIAO1 interaction energy. In both ISAPT/C and ISAPT/SIAO1 methods, the energy decomposition of the trans conformation is consistent in all azobenzene models. The ISAPT/C method provides less attraction in total interaction energy, mainly due to the weaker electrostatic energy compared to the energy values in the ISAPT/SIAO1 method.

The deformational energy of the cis $\rightarrow$ trans isomerization can be analyzed by the geometric rearrangement in the rotation of the N-C bond between two phenyl rings. Several approaches are examined by performing the hydrogen-substitution around the -N=N- bond, leading to the intermediate geometries X-Ph-N=N-H, X-Ph-H, and H-N=N-H. Another approach of H-substitution in the -N=N- bond has the intermediate geometry X-Ph-NH₂, where X refers to the alkyl substituents. The intermediate geometries of the cis and trans conformations are reoptimized in the MP2/aug-cc-pVDZ method and computed the single-point energy using the DLPNO-CCSD(T)/aug-cc-pVDZ method. The rotational energy of the cis $\rightarrow$ trans intermediate geometry is evaluated by $\delta(E_{cis}^{CCSD(T)} - E_{trans}^{CCSD(T)})$ in Table 5.3. The predicted relative deformational energy of cis $\rightarrow$ trans azobenzene isomers, in Table 5.4, is measured

by pairing the intermediate geometries, accumulating their rotational energies with respect to the alkyl substituents in two phenyl rings. In the calculation of the deformational energy in the entire substituted azobenzene, the energy of the subsystem H-N=N-H needs to be added (in the X-Ph-H category) or subtracted (in the X-Ph-N=N-H category) to preserve the types of covalent bonds in the azobenzene models.

The ISAPT/SIAO1 method provides meaningful insight into the intramolecular interactions of the cis $\rightarrow$ trans isomerization of an incremental system of the all-meta alkyl substituted azobenzene molecule, including 10 models where the alkyl substituents increase from Me, Et, iPr to tBu. In the cis structure, the larger alkyl substituent exhibits the stronger net energy, whereas in the trans structure, the energy contribution is consistent in all azobenzene models. The stabilization energy of the cis and trans isomers is mainly influenced by the steric effect between the alkyl substituents of two phenyl rings. The cis structure is stabilized by the dispersion term. The strong repulsion of the steric effect is mainly balanced out by the electrostatic and induction terms. The trans structure is dominated by the electrostatic and dispersion terms due to the planar geometry of two phenyl rings. The far distance between the alkyl substituents exhibits weaker repulsion in the exchange term and overall provides a stronger total interaction energy. Consequently, the trans structure provides stronger conformational stabilization than the cis structure. The predicted DLPNO-CCSD(T) deformational energy of the cis $\rightarrow$ trans azobenzene molecules is analyzed in two intermediate models, including X-Ph-N=N-H and X-Ph-H, where X refers to the alkyl substituents. The relative azobenzene isomerization energy is evaluated by a sum of the ISAPT conformational energy and the DLPNO-CCSD(T) deformational energy of the cis $\rightarrow$ trans isomers. According to two intermediate models, the isomerization energy of the available azobenzene molecules exhibits a similar trend. The X-Ph-H model, overall, provides the slightly higher isomerization energy, resulting from the stronger contribution of the deformational energy, than the X-Ph-N=N-H model. In general, a higher isomerization energy is observed in the smaller size of the alkyl substituents. The choice of intermediate system of the cis $\rightarrow$ trans azobenzene isomers needs more examination for the optimal solution of the deformational energy.

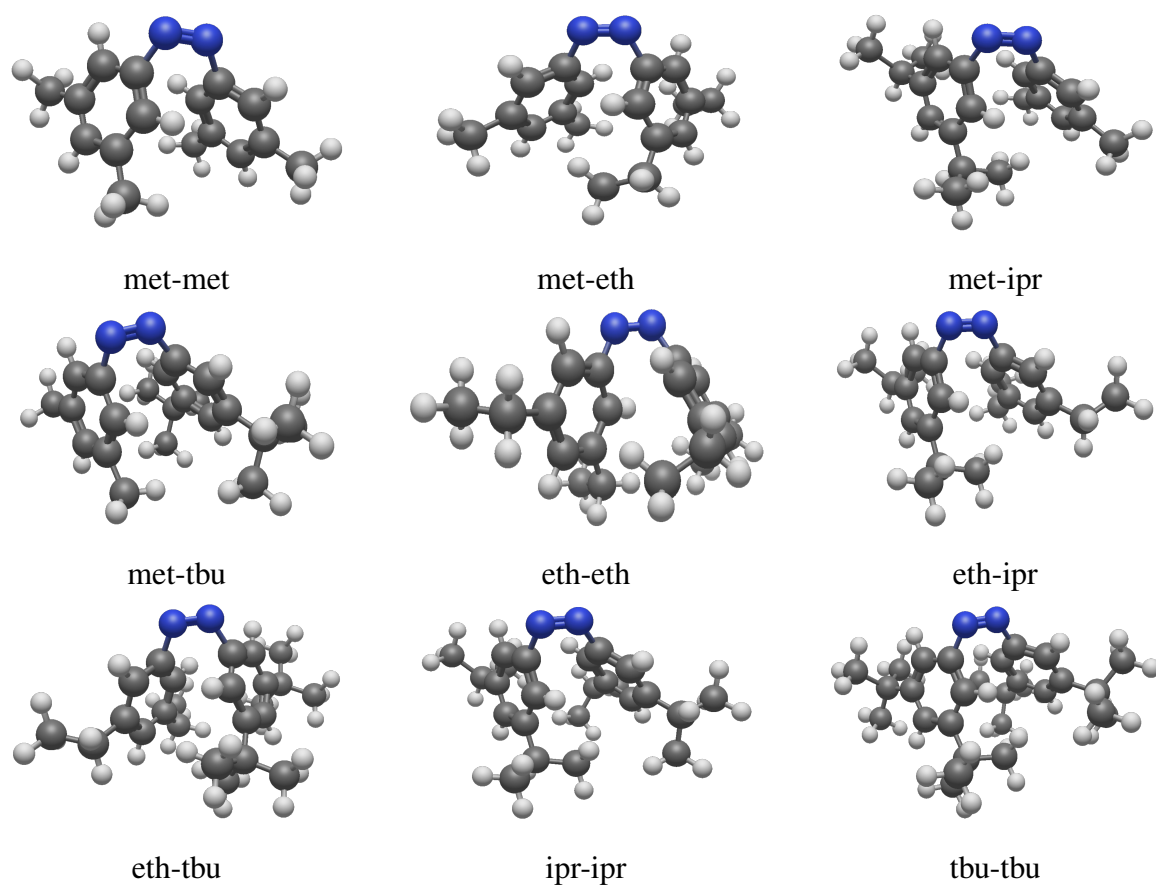


Figure 5.1: The geometric structures of the cis- all-meta alkyl substituted azobenzene molecules. The alkyl groups include methyl (met), ethyl (eth), isopropyl (ipr), and tert-butyl (tbu).

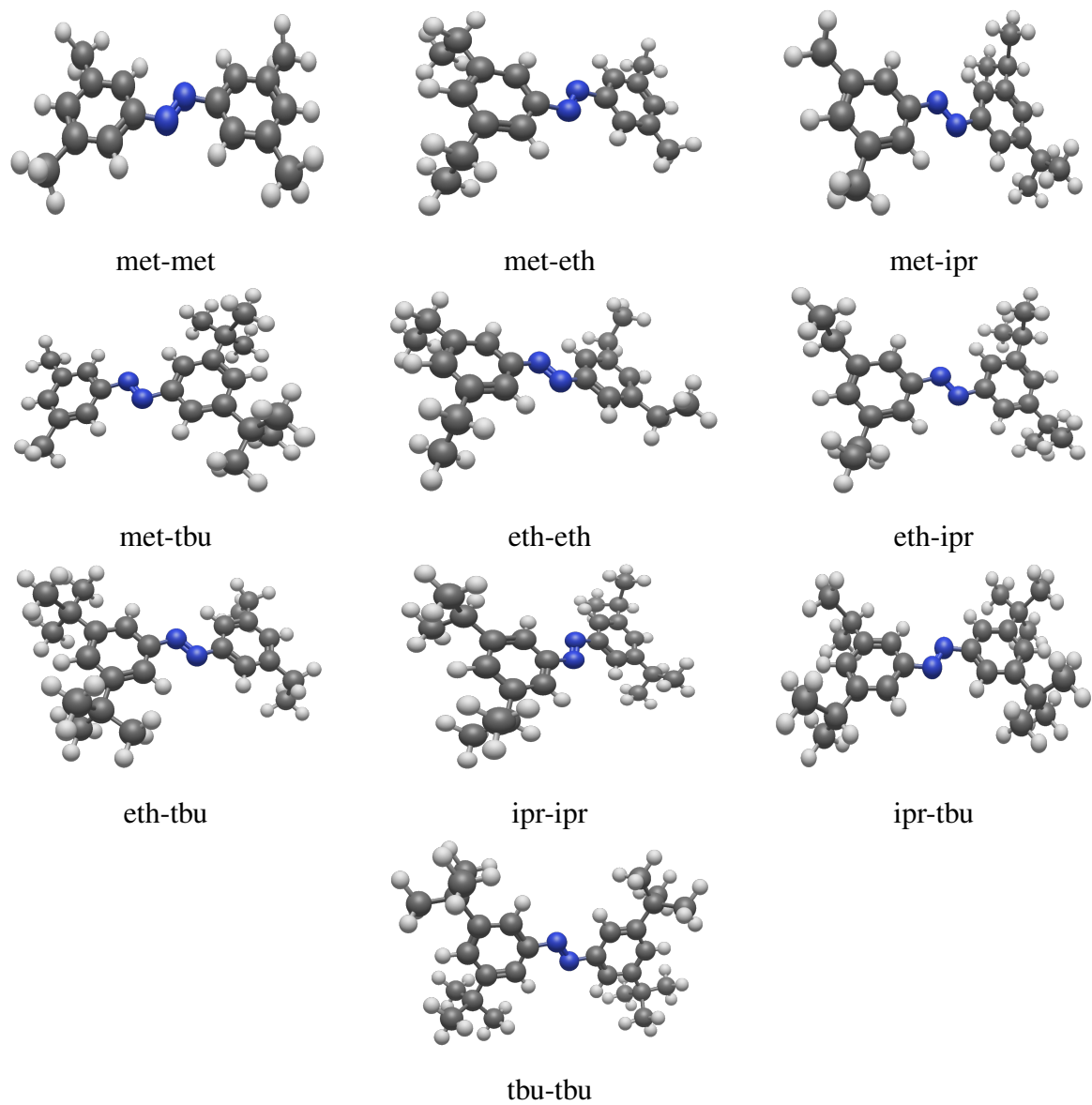


Figure 5.2: The geometric structures of the trans- all-meta alkyl substituted azobenzene molecules. The alkyl groups include methyl (met), ethyl (eth), isopropyl (ipr), and tert-butyl (tbu).

Table 5.1: The ISAPT/SIAO1 decomposition energies (in kcal/mol) of 10 models of the all-meta alkyl substituted azobenzene molecule. The alkyl substituents include Me, Et, iPr, and tBu. The values of δHF are taken from the ISAPT/C linker assignment calculations for the total net energy. [39] The calculations are computed using the aug-cc-pVDZ basis set.

System	Conf.	Elst10	Exch10	Ind20	ExInd20	Disp20	ExDisp20	δHF	Total
met-met	cis	−9.96	21.50	−10.94	9.79	−14.37	2.64	−1.75	−3.27
	trans	−4.95	1.77	−0.87	0.32	−2.31	0.10	−1.53	−7.50
met-eth	cis	−10.94	24.87	−11.96	10.71	−18.64	3.17	−1.81	−4.59
	trans	−4.96	1.78	−0.87	0.32	−2.35	0.10	−1.53	−7.52
met-ipr	cis	−11.28	25.68	−12.29	11.01	−19.34	3.27	−1.84	−4.69
	trans	−4.92	1.77	−0.87	0.32	−2.43	0.10	−1.51	−7.54
met-tbu	cis	−13.60	31.14	−15.21	13.70	−23.46	4.02	−1.93	−5.26
	trans	−4.96	1.81	−0.87	0.32	−2.47	0.10	−1.53	−7.59
eth-eth	cis	−11.25	25.65	−12.29	10.99	−19.37	3.27	−1.85	−4.83
	trans	−4.97	1.79	−0.87	0.32	−2.40	0.10	−1.53	−7.54
eth-ipr	cis	−12.45	28.33	−13.47	12.08	−21.62	3.60	−1.91	−5.60
	trans	−4.93	1.79	−0.87	0.32	−2.48	0.10	−1.52	−7.60
eth-tbu	cis	−14.79	34.49	−16.58	14.95	−26.70	4.49	−2.01	−6.05
	trans	−4.99	1.84	−0.88	0.32	−2.53	0.11	−1.55	−7.66
ipr-ipr	cis	−12.66	30.16	−13.44	12.07	−24.63	3.91	−2.09	−6.61
	trans	−4.96	1.81	−0.88	0.32	−2.51	0.10	−1.53	−7.64
ipr-tbu	cis								
	trans	−5.01	1.86	−0.88	0.32	−2.57	0.11	−1.57	−7.73
tbu-tbu	cis	−9.43	22.29	−9.38	8.30	−20.91	2.94	−1.95	−8.14
	trans	−4.78	1.72	−0.83	0.30	−2.59	0.10	−1.53	−7.59

Table 5.2: The (original) ISAPT/C decomposition energies (in kcal/mol) of 10 models of the all-meta alkyl substituted azobenzene molecule. The alkyl substituents include Me, Et, iPr, and tBu. The calculations are computed using the aug-cc-pVDZ basis set.

System	Conf.	Elst10	Exch10	Ind20	ExInd20	Disp20	ExDisp20	δ HF	Total
met-met	cis	−6.76	30.01	−36.40	33.90	−16.50	4.31	−1.75	6.73
	trans	−2.43	2.53	−2.54	1.08	−2.13	0.09	−1.53	−4.90
met-eth	cis	−7.96	34.28	−40.58	37.95	−21.06	5.08	−1.81	5.61
	trans	−2.46	2.55	−2.54	1.08	−2.17	0.10	−1.53	−4.98
met-ipr	cis	−8.34	34.90	−40.25	37.56	−21.72	5.14	−1.84	5.39
	trans	−2.42	2.55	−2.55	1.09	−2.25	0.10	−1.51	−5.01
met-tbu	cis	−12.06	42.61	−50.83	47.76	−26.48	6.41	−1.93	5.39
	trans	−2.55	2.59	−2.56	1.11	−2.28	0.10	−1.53	−5.12
eth-eth	cis	−8.34	35.10	−41.02	38.31	−21.82	5.20	−1.85	5.62
	trans	−2.49	2.56	−2.55	1.09	−2.21	0.10	−1.53	−5.04
eth-ipr	cis	−9.97	38.07	−43.12	40.26	−24.16	5.60	−1.91	4.94
	trans	−2.45	2.56	−2.56	1.10	−2.29	0.10	−1.52	−5.06
eth-tbu	cis	−13.68	46.78	−54.97	51.71	−29.98	7.09	−2.01	5.02
	trans	−2.61	2.62	−2.58	1.12	−2.35	0.10	−1.55	−5.24
ipr-ipr	cis	−9.93	40.27	−44.47	41.70	−27.31	6.01	−2.09	4.42
	trans	−2.51	2.58	−2.57	1.11	−2.32	0.10	−1.53	−5.15
ipr-tbu	cis								
	trans	−2.67	2.64	−2.59	1.12	−2.38	0.10	−1.57	−5.29
tbu-tbu	cis	−5.40	28.96	−28.89	26.81	−22.68	4.30	−1.95	1.16
	trans	−2.56	2.62	−2.59	1.14	−2.49	0.11	−1.53	−5.35

Table 5.3: The DLPNO-CCSD(T)/aug-cc-pVDZ rotational energy (in kcal/mol) for the H-substituted intermediate geometries of the cis and trans azobenzene isomers. The electron correlation energy E(L-MP2) extracted from the DLPNO-CCSD(T) calculation serves for the reference to the MP2 correlation energy.

System	CCSD(T)	E(L-MP2)
	δ (cis-trans) (kcal/mol)	δ (cis-trans) (kcal/mol)
met-Ph-N=N-H	5.36	6.04
ipr-Ph-N=N-H	5.66	6.41
eth-Ph-N=N-H	5.76	6.42
tbu-Ph-N=N-H	5.69	6.27
met-Ph-H	0.19	0.16
ipr-Ph-H	0.18	0.12
eth-Ph-H	0.43	0.37
tbu-Ph-H	0.42	0.33
met-Ph-NH ₂	-0.30	-0.19
ipr-Ph-NH ₂	-0.10	-0.03
eth-Ph-NH ₂	-0.32	-0.31
tbu-Ph-NH ₂	0.00	0.05
H-N=N-H	5.55	5.97

Table 5.4: The relative isomerization energy (in kcal/mol) of 10 azobenzene models evaluated by the sum of the ISAPT/SIAO1 conformational energy and the predicted DLPNO-CCSD(T)/aug-cc-pVDZ deformational energy in the intermediate systems of X-Ph-N=N-H and X-Ph-H. The calculations are computed in the aug-cc-pVDZ basis set.

Intermediate system	azobenzene	CCSD(T)	E(L-MP2)	ISAPT/SIAO1	Total
		δ (cis-trans)	δ (cis-trans)	δ (cis-trans)	
X-Ph-N=N-H	met-met	5.16	7.29	4.22	9.39
	met-eth	5.56	7.28	2.93	8.49
	met-ipr	5.47	7.14	2.85	8.32
	met-tbu	5.49	7.27	2.33	7.82
	eth-eth	5.96	6.91	2.71	8.67
	eth-ipr	5.87	6.90	2.00	7.87
	eth-tbu	5.89	6.53	1.61	7.50
	ipr-ipr	5.78	6.76	1.03	6.81
	ipr-tbu	5.80	6.99		
	tbu-tbu	5.82	7.13	-0.55	5.27
X-Ph-H	met-met	5.93	6.30	4.22	10.15
	met-eth	6.17	6.05	2.93	9.10
	met-ipr	5.92	6.26	2.85	8.77
	met-tbu	6.16	5.80	2.33	8.49
	eth-eth	6.41	6.08	2.71	9.12
	eth-ipr	6.16	5.83	2.00	8.16
	eth-tbu	6.40	5.87	1.61	8.02
	ipr-ipr	5.91	6.05	1.03	6.94
	ipr-tbu	6.15	6.22		
	tbu-tbu	6.39	6.01	-0.55	5.84

Chapter 6

Summary and Outlook

Symmetry-adapted perturbation theory (SAPT) is a popular method for studying intermolecular and intramolecular interactions. The key strength of SAPT is its ability to decompose the interaction energy into physically meaningful components, including electrostatics, induction, dispersion, and their exchange counterparts. The intramolecular SAPT (ISAPT) method [49] relies on fragmentation of the molecule into two isolated fragments A and B embedded into a covalently bonded linker C. The energy correction of the ISAPT method is only available at the SAPT0 level of perturbation theory. The thesis provides the improvement of the ISAPT method, the potential approach for the development of the non-fragmentation ISAPT algorithm, and the application of the improved ISAPT method to the study of the *Z/E*-azobenzene isomerization.

The refinement of the ISAPT method, in Chapter 3, involves the correction of the quantification of the contributions of interaction energy. In particular, the unphysically positive electrostatic energy has been corrected to strongly negative in the molecular models that exhibit hydrogen bonding. The magnitudes of induction and exchange-induction terms are significantly reduced, yielding a meaningfully attractive net energy. The issue of artificial dipole moments created at the A–C and B–C boundaries is caused by the missing electrons on one of the hybrid orbitals of the linking atom as the entire doubly occupied intrinsic bond orbital (IBO) is assigned to C. The spurious dipoles are overcome by reassigning one electron from C to each of A/B, placing it on the missing hybrid orbital referred to as the intrinsic hybrid orbital (IHO). This reassignment results in updated density matrices for fragments A and B which can be used in standard SAPT0 formulas to compute updated electrostatic, induction, and dispersion energies (it is assumed that the link hybrid orbital does not participate in excitations). For

the corresponding exchange corrections, a choice needs to be made regarding the spin coupling of the unpaired electrons on the A–C and B–C link orbitals, whether they undergo parallel, perpendicular, or averaged spin coupling. The IHO is obtained by the projection of the IBO onto the occupied space of the fragment A/B that is spanned by the AO basis functions (in the ISAPT(SAO) approach) or by the intrinsic atomic orbitals (IAOs) (in the ISAPT(SIAO) approach). The study shows that the ISAPT(SIAO1) approach providing superior consistency between basis sets is the best variant to evaluate intramolecular interactions.

The reliance of the current ISAPT methods of a fragmentation scheme limits the description of the intramolecular interactions arising from the entire molecule. The range separation technique is the potential approach for the design of the non-fragmentation ISAPT protocol. At the first stage, in Chapter 4, the long-ranged Coulomb potential $v_l(r)$ is studied in the SAPT and ISAPT(SIAO1) energy calculations to verify its ability to cover the entire noncovalent interactions. In the interaction potential operator V , the $v_l(r)$, in forms of the Gaussian and error functions, fully replaces $1/r$ potentials (in the RSEP approach) or only replaces the attractive singularity (in the RSEP-eN approach). In the evaluation of the interaction energy, when the separation parameter increases to infinity, the RSEP(gau/erf) and RSEP-eN(gau/erf) methods provide smooth convergence towards the conventional SAPT0 values. RSEP(erf) overperforms other range separation-based SAPT and ISAPT variants and provides the quickest convergence and smallest energy variance between parameter's values. In the design of the fragmentation-free ISAPT algorithm, the short-ranged Coulomb potential $v_s(r)$ could be used to describe the covalent interactions in a molecule and the long-ranged part $v_l(r)$ represents the noncovalent interactions in the energy calculations.

The ISAPT(SIAO1) method is employed to study the intramolecular interaction effects of the cis $\rightarrow$ trans all-meta alkyl substituted azobenzene isomerization in Chapter 5. The stabilization of the cis conformation increases as the size of the alkyl substituents increases from Me up to tBu, while the stabilization of the trans conformation exhibits a good consistency in all azobenzene models. The cis conformer is stabilized by the dispersion energy, followed by the electrostatic and induction energies, mainly due to the steric effects between the alkyl substituents. The trans conformer is mainly bound by the electrostatic and dispersion effects

because of the planar geometry of two phenyl rings. The DLPNO-CCSD(T) deformational energy of cis $\rightarrow$ trans isomers is evaluated in two models of intermediate geometry X-Ph-N=N-H and X-Ph-H, where X includes four alkyl substituents of Me, Et, iPr, and tBu. The predicted cis $\rightarrow$ trans azobenzene isomerization energy has a similar trend in two intermediate systems. The isomerization rates, overall, accelerate when the size of the alkyl substituents increases. The intermediate systems need further examination, probably in the effects and contribution of the dihedral angle in the cis $\rightarrow$ trans geometric rotation, to select the optimal variant for the isomerization energy.

The future work will focus on the development of the non-fragmentation ISAPT algorithm. The range-separation functions need to be carefully examined in the ISAPT energy calculation as they substitute the conventional interaction potentials. The energy corrections could be calculated using the long-ranged separation function with the HF molecular orbitals obtained using the short-ranged separation function. Another possible approach is to accumulate the pairwise interactions of each atom in a molecule with or without the use of the range-separation functions. The current improved-ISAPT method provides physically meaningful intramolecular interaction energy. However, the ISAPT method has limitations in describing the intramolecular effects in the bulky molecules that could or could not be partitioned into fragments. The extension in the capacity of fragments is feasible but problematic in code maintenance. The non-fragmentation ISAPT method would be a breakthrough in the theoretical realm for the quantification of noncovalent intramolecular interactions in any molecular structure of interest.

References

- [1] W. H. Adams. Perturbation theory of intermolecular interactions: are second-order Rayleigh-Schrödinger energies meaningful. *Int. J. Quantum Chem.*, S25:165–181, 1991.
- [2] W. H. Adams. Two new symmetry-adapted perturbation theories for the calculation of intermolecular interaction energies. *Theor. Chem. Acc.*, 108:225–231, 2002.
- [3] Flavia Aleotti, Vasilis Petropoulos, Hannah Van Overeem, Michele Pettini, Michele Mancinelli, Daniel Pecorari, Margherita Maiuri, Riccardo Medri, Andrea Mazzanti, Fabrizio PredaAntonio Perri, Dario Polli, Irene Conti, Giulio Cerullo, and Marco Garavelli. Engineering azobenzene derivatives to control the photoisomerization process. *J. Phys. Chem. A*, 127:10435–10449, 2023.
- [4] R Ballardini, V Balzani, A Credi, M T Gandolfi, and M Venturi. Artificial molecular-level machines: which energy to make them work? *Acc. Chem. Res.*, 34:445–455, 2001.
- [5] H. M. Dhammika Bandara and Shawn C. Burdette. Photoisomerization in different classes of azobenzene. *Chem. Soc. Rev.*, 41:1809–1825, 2012.
- [6] Chiara Di Berardino, Marcel A. Strauss, Dominic Schatz, and Hermann A. Wegner. An incremental system to predict the effect of different london dispersion donors in all-meta-substituted azobenzenes. *Chem. Eur. J.*, 28:e202104284, 2022.
- [7] Alessandro Cembran, Fernando Bernardi, Marco Garavelli, Laura Gagliardi, and Giorgio Orlandi. On the mechanism of the cistrans isomerization in the lowest electronic states of azobenzene: S0, s1, and t1. *J. Am. Chem. Soc.*, 126:3234–3243, 2004.

- [8] Chih-Wei Chang, Ying-Chih Lu, Tsai-Te Wang, and Eric Wei-Guang Diau. Photoisomerization dynamics of azobenzene in solution with s1 excitation: A femtosecond fluorescence anisotropy study. *J. Am. Chem. Soc.*, 126:10109–10118, 2004.
- [9] Y.-T. Chen, K. Hui, and J.-D. Chai. The van der waals interactions in rare-gas dimers: the role of interparticle interactions. *Phys. Chem. Chem. Phys.*, 18:3011–3022, 2016.
- [10] Cristina Cojocariu and P. Rochon. Light-induced motions in azobenzene-containing polymers. *Pure Appl. Chem.*, 76:1479–1497, 2004.
- [11] Christina R. Crecca and Adrian E. Roitberg. Theoretical study of the isomerization mechanism of azobenzene and disubstituted azobenzene derivatives. *J. Phys. Chem. A*, 110:8188–8203, 2006.
- [12] Teresa Cusati, Giovanni Granucci, and Maurizio Persico. Photodynamics and time-resolved fluorescence of azobenzene in solution: A mixed quantum-classical simulation. *J. Am. Chem. Soc.*, 133:5109–5123, 2011.
- [13] Jadranka Dokić, Marcel Gothe, Jonas Wirth, Maike V Peters, Jutta Schwarz, Stefan Hecht, and Peter Saalfrank. Quantum chemical investigation of thermal cis-to-trans isomerization of azobenzene derivatives: substituent effects, solvent effects, and comparison to experimental data. *J. Phys. Chem. A*, 113:6763–6773, 2009.
- [14] Christophe Dugave and Luc Demange. Cistrans isomerization of organic molecules and biomolecules: Implications and applications. *Chem. Rev.*, 103:2475–2532, 2003.
- [15] Francesca Ercole, Thomas P. Davis, , and Richard A. Evans. Photo-responsive systems and biomaterials: photochromic polymers, light-triggered self-assembly, surface modification, fluorescence modulation and beyond. *Polym. Chem.*, 1:37–54, 2010.
- [16] V. A. Fock. Näherungsmethode zur lösung des quantenmechanischen mehrkörperproblems. *Zeitschrift für Physik (in German)*, 61:126–148, 1930.
- [17] V. A. Fock. 'selfconsistent field' mit austausch für natrium. *Zeitschrift für Physik (in German)*, 62:795–805, 1930.

- [18] Tatsuya Fujino, Sergei Yu. Arzhantsev, and Tahei Tahara. Femtosecond time-resolved fluorescence study of photoisomerization of trans-azobenzene. *J. Phys. Chem. A*, 105:8123–8129, 2001.
- [19] J. Garcia, R. Podeszwa, and K. Szalewicz. Sapt codes for calculations of intermolecular interaction energies. *J. Chem. Phys.*, 152:184109, 2020.
- [20] Jaume García-Amorós and Dolores Velasco. Recent advances towards azobenzene-based light-driven real-time information-transmitting materials. *Beilstein J. Org. Chem.*, 8:1003–1017, 2012.
- [21] J. F. Gonthier and C. Corminboeuf. Exploration of zeroth-order wavefunctions and energies as a first step toward intramolecular symmetry-adapted perturbation theory. *J. Chem. Phys.*, 140:154107, 2014.
- [22] Montgomery Gray and John M. Herbert. Simplified tuning of long-range corrected density functionals for use in symmetry-adapted perturbation theory. *J. Chem. Phys.*, 155:034103, 2021.
- [23] Douglas Rayner Hartree and W. Hartree. Self-consistent field, with exchange, for beryllium. *Proc. Roy. Soc.*, A150, 1935.
- [24] Andreas H Heindl, Raffael C Wende, and Hermann A Wegner. London dispersion as important factor for the stabilization of (z)-azobenzenes in the presence of hydrogen bonding. *Beilstein J Org Chem.*, 14:1238–1243, 2018.
- [25] W Heisenberg. The multibody problem and resonance in quantum mechanics-ii. *Z. f. Phys.*, 40:551–576, 1927.
- [26] T. Helgaker, P. Jørgensen, and J. Olsen. *Molecular Electronic-Structure Theory*. Wiley, New York, 2000.
- [27] J. Heyd, G. E. Scuseria, and M. Ernzerhof. Hybrid functionals based on a screened coulomb potential. *J. Chem. Phys.*, 118:8207–8215, 2003.

- [28] E. G. Hohenstein, H. M. Jaeger, E. J. Carrell, G. S. Tschumper, and C. D. Sherrill. Accurate interaction energies for problematic dispersion-bound complexes: Homogeneous dimers of nccn, p₂, and pccp. *J. Chem. Theory Comput.*, 7:2842–2851, 2011.
- [29] E. G. Hohenstein and C. D. Sherrill. Density fitting and cholesky decomposition approximations in symmetry-adapted perturbation theory: Implementation and application to probe the nature of pi-pi interactions in linear acenes. *J. Chem. Phys.*, 132:184111, 2010.
- [30] E. G. Hohenstein and C. D. Sherrill. Wavefunction methods for noncovalent interactions. *WIREs Comput. Mol. Sci.*, 2:304–326, 2012.
- [31] H. Iikura, T. Tsuneda, T. Yanai, and K. Hirao. A long-range correction scheme for generalized-gradient-approximation exchange functionals. *J. Chem. Phys.*, 115:3540–3544, 2001.
- [32] B. Jeziorski, R. Moszyński, and K. Szalewicz. Perturbation theory approach to intermolecular potential energy surfaces of van der waals complexes. *Chem. Rev.*, 94:1887–1930, 1994.
- [33] Rafal Klajn. Immobilized azobenzenes for the construction of photoresponsive materials. *Pure Appl. Chem.*, 82:2247–2279, 2010.
- [34] G. Knizia. Intrinsic atomic orbitals: An unbiased bridge between quantum theory and chemical concepts. *J. Chem. Theory Comput.*, 9:4834–4843, 2013.
- [35] K. U. Lao and J. M. Herbert. Symmetry-adapted perturbation theory with kohn-sham orbitals using non-empirically tuned, long-range-corrected density functionals. *J. Chem. Phys.*, 140:044108, 2014.
- [36] I. K. Lednev, T.-Q. Ye, R. E. Hester, and J. N. Moore. Femtosecond time-resolved uv-visible absorption spectroscopy of trans-azobenzene in solution. *J. Phys. Chem.*, 100:13338, 1996.

- [37] K Lednev, T-Q Ye, P Matousek, M Towrie, Paolo Foggi, FVR Neuwahl, S Umapathy, RE Hester, and John N Moore. Femtosecond time-resolved uv-visible absorption spectroscopy of trans-azobenzene: dependence on excitation wavelength. *Chem. Phys. Lett.*, 290:68–74, 1998.
- [38] T. Leininger, H. Stoll, H.-J. Werner, and A. Savin. Combining long-range configuration interaction with short-range density functionals. *Chem. Phys. Lett.*, 275:151–160, 1997.
- [39] D. Luu and K. Patkowski. Overcoming artificial multipoles in intramolecular symmetry-adapted perturbation theory. *J. Phys. Chem. A*, 127:356–377, 2023.
- [40] Du Luu, Clemence Corminboeuf, and Konrad Patkowski. Range separation of the interaction potential in intermolecular and intramolecular symmetry-adapted perturbation theory. *Chem. Theory Comput.*, 20:7884–7903, 2024.
- [41] Zahid Mahimwalla, Kevin G. Yager, Jun ichi Mamiya, Atsushi Shishido, Arri Priimagi, and Christopher J. Barrett. Azobenzene photomechanics: prospects and potential applications. *Polym. Bull.*, 69:967–1006, 2012.
- [42] I. Mayer. Towards a “chemical” hamiltonian. *Int. J. Quantum Chem.*, 23:341–363, 1983.
- [43] L. E. McMurchie and E. R. Davidson. Calculation of integrals over *ab initio* pseudopotentials. *J. Comput. Phys.*, 44:289–301, 1981.
- [44] Larry E McMurchie and Ernest R Davidson. One- and two-electron integrals over cartesian gaussian functions. *Journal of Computational Physics*, 26:218–231, 1978.
- [45] O. R. Meitei and A. Hesselmann. On the stability of cyclophane derivatives using a molecular fragmentation method. *ChemPhysChem*, 17:3863–3874, 2016.
- [46] O. R. Meitei and A. Hesselmann. Intramolecular interactions in sterically crowded hydrocarbon molecules. *J. Comput. Chem.*, 38:2500–2508, 2017.
- [47] A. J. Misquitta. Charge transfer from regularized symmetry-adapted perturbation theory. *J. Chem. Theory Comput.*, 9:5313–5326, 2013.

- [48] T. M. Parker, L. A. Burns, R. M. Parrish, A. G. Ryno, and C. D. Sherrill. Levels of symmetry adapted perturbation theory (sapt). i. efficiency and performance for interaction energies. *J. Chem. Phys.*, 140:094106, 2014.
- [49] R. M. Parrish, J. F. Gonthier, C. Corminboeuf, and C. D. Sherrill. Communication: Practical intramolecular symmetry adapted perturbation theory via hartree-fock embedding. *J. Chem. Phys.*, 143:051103, 2015.
- [50] E. Pastorczak, A. Prlj, J. F. Gonthier, and C. Corminboeuf. Intramolecular symmetry-adapted perturbation theory with a single-determinant wavefunction. *J. Chem. Phys.*, 143:224107, 2015.
- [51] K. Patkowski. Recent developments in symmetry-adapted perturbation theory. *WIREs Comput. Mol. Sci.*, 10:e1452, 2020.
- [52] K. Patkowski, B. Jeziorski, and K. Szalewicz. Unified treatment of chemical and van der Waals forces via symmetry-adapted perturbation expansion. *J. Chem. Phys.*, 120:6849–6862, 2004.
- [53] K. Patkowski, T. Korona, and B. Jeziorski. Convergence behavior of the symmetry-adapted perturbation theory for states submerged in Pauli forbidden continuum. *J. Chem. Phys.*, 115:1137–1152, 2001.
- [54] Linus. Pauling. The nature of the chemical bond. application of results obtained from the quantum mechanics and from a theory of paramagnetic susceptibility to the structure of molecules. . *Am. Chem. Soc.*, 53:1367–1400, 1931.
- [55] C. C. J. Roothaan. New developments in molecular orbital theory. *Reviews of Modern Physics*, 23:69–89, 1951.
- [56] Anna Laura Sanna, Tatiana Pachova, Alessandra Catellani, Arrigo Calzolari, and Giuseppe Sforazzini. Meta-substituted asymmetric azobenzenes: Insights into structure–property relationship. *Molecules*, 29:1929, 2024.

- [57] H. Satzger, C. Root, and M. Braun. Excited-state dynamics of trans- and cis-azobenzene after uv excitation in the π band. *J. Phys. Chem. A*, 108:6265–6271, 2004.
- [58] E. Schrödinger. An undulatory theory of the mechanics of atoms and molecules. *The Phys. Rev.*, 28, 1926.
- [59] T. Schultz, J. Quenneville, B. Levine, A. Toniolo, T. J. Martínez, S. Lochbrunner, M. Schmitt, J. P. Schaffer, M. Z. Zgierski, and A. Stolow. Mechanism and dynamics of azobenzene photoisomerization. *J. Am. Chem. Soc.*, 125:8098–8099, 2003.
- [60] A N Shipway and I Willner. Electronically transduced molecular mechanical and information functions on surfaces. *Acc. Chem. Res.*, 34:421–432, 2001.
- [61] J. C. Slater. The self consistent field and the structure of atoms. *PHYSICAL REVIEW*, 32:339–348, 1928.
- [62] J. C. Slater. Note on hartree’s method. *Physical Review*, 35:210–211, 1930.
- [63] D. G. A. Smith, L. A. Burns, A. C. Simmonett, R. M. Parrish, M. C. Schieber, R. Galvelis, P. Kraus, H. Kruse, R. Di Remigio, A. Alenaizan, A. M. James, S. Lehtola, J. P. Misiewicz, M. Scheurer, R. A. Shaw, J. B. Schriber, Y. Xie, Z. L. Glick, D. A. Sirianni, J. S. O’Brien, J. M. Waldrop, A. Kumar, E. G. Hohenstein, B. P. Pritchard, B. R. Brooks, H. F. Schaefer III, A. Y. Sokolov, K. Patkowski, A. E. DePrince, U. Bozkaya, R. A. King, F. A. Evangelista, J. M. Turney, T. D. Crawford, and C. D. Sherrill. PSI4 1.4: Open-source software for high-throughput quantum chemistry. *J. Chem. Phys.*, 152:184108, 2020.
- [64] Marcel A. Strauss and Hermann A. Wegner. Exploring london dispersion and solvent interactions at alkyl–alkyl interfaces using azobenzene switches. *Angew Chem Int Ed Engl.*, 58:18552–18556, 2019.
- [65] K. Szalewicz. Symmetry-adapted perturbation theory of intermolecular forces. *WIREs Comput. Mol. Sci.*, 2:254–272, 2012.
- [66] K. Szalewicz, R. Bukowski, and B. Jeziorski. On the importance of many-body forces in clusters and condensed phase. In C. E. Dykstra, G. Frenking, K. S. Kim, and G. E.

- Scuseria, editors, *Theory and Applications of Computational Chemistry: The First 40 Years. A Volume of Technical and Historical Perspectives*, chapter 33, pages 919–962. Elsevier, Amsterdam, 2005.
- [67] O. A. Vydrov and G. E. Scuseria. Assessment of a long-range corrected hybrid functional. *J. Chem. Phys.*, 125:234109, 2006.
- [68] Yanan Wang, Lijie Huang, Xiyue Wang, Xuyang Lu, Beiwei Wang, Yafei Qin, and Chongxing Huang. Photoresponsive triple shape memory polymers with a self-healing function based on poly (lactic acid)/polycaprolactone blends. *Polymer Testing*, 120:107966, 2023.
- [69] Yi Xie, Daniel G. A. Smith, and C. David Sherrill. Implementation of symmetry-adapted perturbation theory based on density functional theory and using hybrid exchange-correlation kernels for dispersion terms. *J. Chem. Phys.*, 157:024801, 2022.
- [70] Kevin G. Yager and Christopher J. Barrett. Novel photo-switching using azobenzene functional materials. *Journal of Photochemistry and Photobiology A: Chemistry*, 182:250–261, 2006.
- [71] Yuanfeng Ye, Juan Pang, Xiaojie Zhou, and Jingwei Huang. Understanding the torsion effects on optical properties of azobenzene derivatives. *Computational and Theoretical Chemistry*, 1076:17–22, 2016.
- [72] Mingxin Zheng and Jinying Yuan. Polymeric nanostructures based on azobenzene and their biomedical applications: synthesis, self-assembly and stimuli-responsiveness. *Org. Biomol. Chem.*, 20:749–767, 2022.
- [73] T. Živković and Z. B. Maksić. Explicit formulas for molecular integrals over hermite–gaussian functions. *J. Chem. Phys.*, 49:3083–3087, 1968.