Title

A Second-Generation Energy Decomposition Analysis for Møller-Plesset Perturbation Theory Using a Global Virtual Space

Author

Ikeda, Kevin K

Publication Date

2023

Peer reviewed|Thesis/dissertation

A Second-Generation Energy Decomposition Analysis for Møller-Plesset Perturbation Theory Using a Global Virtual Space

by

Kevin Kiyoshi Ikeda

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Chemical Engineering

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Martin Head-Gordon, Chair
Associate Professor Ali Mesbah, Co-chair
Associate Professor Kranthi Mandadapu
Associate Professor Eric Neuscamman

Fall 2022

A Second-Generation Energy Decomposition Analysis for Møller-Plesset Perturbation Theory Using a Global Virtual Space

Abstract

A Second-Generation Energy Decomposition Analysis for Møller-Plesset Perturbation Theory Using a Global Virtual Space

by

Kevin Kiyoshi Ikeda

Doctor of Philosophy in Chemical Engineering

University of California, Berkeley

Professor Martin Head-Gordon, Chair

Associate Professor Ali Mesbah, Co-chair

In this thesis, a newly developed second-generation energy decomposition analysis (EDA) based on second order Møller-Plesset perturbation theory (MP2) is presented. EDA's have become widely used to aid in understanding the nature of intermolecular interactions and are commonly based on density functional theory (DFT) and self-consistent field (SCF) calculations. However, using correlated post-SCF methods, such as MP2, for EDA is less common, partly due to the complexity associated with defining suitable approaches. As such, this thesis focuses on presenting a new approach for EDA's for post-SCF methods through the implementation of restricted and unrestricted MP2 EDA calculations. This newly developed MP2 EDA builds upon the SCF-level second-generation absolutely localized molecular orbital EDA approach (ALMO-EDA-II) and provides distinct energy contributions for a frozen interaction energy (containing permanent electrostatics and Pauli repulsions), a polarized energy (yielding induced electrostatics), a dispersion-corrected energy, and the fully relaxed energy (which yields charge transfer). Importantly, the theory has been designed such that there are stable basis set limits for each term, evidenced by basis set stability calculations on a wide variety of systems as well as the S22 and Ionic43 datasets. Additionally, MP2 ALMO-EDA-II has been applied to four non-covalently bonded classes of complexes; a class of conventional hydrogen bonded systems, a class of non-conventionally hydrogen bonded systems, a class of tetrel bonded systems, and a "solvent-resistant" halogen bonded system. Through these systems, the usefulness of MP2 ALMO-EDA-II is explored, especially through its treatment of the correlation component of the interaction energy.

To my parents,
Allan and Serena,
for their constant love and support.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Intermolecular Interactions

Intermolecular interactions are prominent in a variety of different chemical systems, from gas-storage materials[120] to drug design[93]. One way to analyze these chemical systems, including those where intermolecular interactions are relevant, is through the lens of quantum chemistry. Through the use of quantum chemistry, we are able to predict energies and other properties relevant to intermolecular interactions, such as dipole moments and polarizabilities. In quantum chemistry, a molecular wavefunction, Ψ, is able to fully describe the state of a molecular system. With the wavefunction of a system, various observables related to molecular properties can be obtained through the use of Hermitian operators, where the expectation value of an operator, $\hat{O}$, is

$$\langle \Psi | \hat{O} | \Psi \rangle = \int \Psi^* \hat{O} \Psi \tag{1.1}$$

1.2 The Schrödinger Equation

In the field of quantum chemistry, modeling chemical systems usually centers around solving the Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} | \Psi \rangle = \hat{H} | \Psi \rangle \tag{1.2}$$

Commonly, modeling the system focuses on stationary states, thus we focus on the time-independent Schrödinger equation (TISE)

$$\hat{H} | \Psi \rangle = E | \Psi \rangle \tag{1.3}$$

where $\hat{H}$ is the Hamiltonian operator, $|\Psi\rangle$ is the wavefunction that describes a stationary state of the system, and E is the energy of the system. The Hamiltonian operator for a

system of N electrons and M nuclei, in atomic units, is

$$\hat{H} = \left(-\sum_{i=1}^{N} \frac{1}{2}\nabla_i^2\right) + \left(-\sum_{A=1}^{M} \frac{1}{2M_A}\nabla_A^2\right) + \left(-\sum_{i=1}^{N}\sum_{A=1}^{M} \frac{Z_A}{|R_A - r_i|}\right) + \left(\sum_{i=1}^{N}\sum_{j>1}^{N} \frac{1}{|r_i - r_j|}\right) + \left(\sum_{A=1}^{M}\sum_{B>A}^{M} \frac{Z_A Z_B}{|R_A - R_B|}\right)$$

where M_A is the reduced mass of nucleus A, Z_A is the charge of nucleus A, R_A are the coordinates of nucleus A, and r_i are the coordinates of the ith electron.

As the nuclei have a much larger mass than electrons and thus move much more slowly, it is common to approximate the nuclei as stationary, known as the Born-Oppenheimer approximation. With this approximation, the electronic Hamiltonian is defined as

$$\hat{H}_e = \left(-\sum_{i=1}^{N} \frac{1}{2}\nabla_i^2\right) + \left(-\sum_{i=1}^{N}\sum_{M=1}^{A} \frac{Z_A}{|R_A - r_i|}\right) + \left(\sum_{i=1}^{i}\sum_{j>1}^{j} \frac{1}{|r_i - r_j|}\right)$$

and the corresponding electronic TISE is

$$\hat{H}_e \Psi_e = E_e \Psi_e \tag{1.4}$$

For the remainder of this thesis, we will refer to the electronic forms of the equations and terms.

1.3 Hartree-Fock Theory

A common approximation to the TISE is the Hartree-Fock (HF) approximation. The HF approximation is a mean-field approximation to the TISE where each electron moves in the field generated by the other electrons. The variational principle ensures that for a trial wavefunction, $|\Phi\rangle$, that satisfies

$$\langle \Phi | \Phi \rangle = 1 \tag{1.5}$$

it is guaranteed that

$$E_0 = \langle \Phi | \hat{H} | \Phi \rangle \geq \mathcal{E}_0 \tag{1.6}$$

where $\mathcal{E}_0$ is the true ground state energy. A starting guess for the wavefunction would be the product of functions that each depend only on an individual electron

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n) = \chi_1(\mathbf{x}_1)\chi_2(\mathbf{x}_2)...\chi_n(\mathbf{x}_n) \tag{1.7}$$

where χ are orthonormal one-electron spin orbitals which can further be decomposed into a spatial, ϕ, and either an α or β spin component

$$\chi(\mathbf{x}) = \begin{cases} \psi(\mathbf{r})\alpha(\omega) \\ \text{or} \\ \psi(\mathbf{r})\beta(\omega) \end{cases} \tag{1.8}$$

where $\mathbf{r}$ and ω describe the spatial and spin coordinates of the electron respectively. However, due to the fact that electrons are fermions, the wavefunction describing a multi-electron system must be antisymmetric with respect to particle exchange. The simplest antisymmetric wavefunction is a determinant, specifically referred to as a Slater determinant

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n) = (n!)^{-\frac{1}{2}} \begin{vmatrix} \chi_1(\mathbf{x}_1) & \chi_1(\mathbf{x}_2) & \dots & \chi_1(\mathbf{x}_n) \\ \chi_2(\mathbf{x}_1) & \chi_2(\mathbf{x}_2) & \dots & \chi_2(\mathbf{x}_n) \\ \vdots & \vdots & & \vdots \\ \chi_n(\mathbf{x}_1) & \chi_n(\mathbf{x}_2) & \dots & \chi_n(\mathbf{x}_n) \end{vmatrix}$$

Using Slater-Condon rules[114] the energy of a single determinant, Ψ^{HF}, can be written in chemist notation as

$$E^{\mathrm{HF}} = \left\langle \Psi^{\mathrm{HF}} \right| \hat{H} \left| \Psi^{\mathrm{HF}} \right\rangle = \sum_{r}^{N} (r|\hat{h}|r) + \sum_{r}^{N} \sum_{s>r}^{N} ((rr|ss) - (rs|sr)) \tag{1.9}$$

where

$$\langle r| \hat{h} |s\rangle = \int \chi_r^*(\mathbf{x})(-\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M} \frac{Z_A}{|R_A - \mathbf{x}|})\chi_s(\mathbf{x})d\mathbf{x} \tag{1.10}$$

and $(rr|ss)$ and $(rs|sr)$ are two electron integrals defined as

$$(pr|qs) = \int\int \chi_p^*(\mathbf{x})\chi_r(\mathbf{x}) \frac{1}{|\mathbf{x} - \mathbf{x}'|} \chi_q^*(\mathbf{x}')\chi_s(\mathbf{x}')d\mathbf{x}d\mathbf{x}' \tag{1.11}$$

and the N spin orbitals are chosen as the spin orbitals that minimize the energy with the constraint that they are orthonormal.

$$\langle r|s\rangle = \delta_{rs} \tag{1.12}$$

Using the minimization method of Lagrange multipliers[114], the equation for the optimal spin orbitals in canonical form is

$$\hat{f} |r\rangle = \epsilon_r |r\rangle \tag{1.13}$$

where the Fock operator, $\hat{f}$ is composed of effective one electron operators

$$\hat{f}(\mathbf{x}) = \hat{h}(\mathbf{x}) + \sum_{r}^{N} \left(\hat{J}_r(\mathbf{x}) - \hat{K}_r(\mathbf{x}) \right) \tag{1.14}$$

where the coulomb operator, $\hat{J}$, and exchange operator, $\hat{K}$, are defined as

$$\begin{aligned} \hat{J}_r(\mathbf{x})\chi_r(\mathbf{x}) &= \left(\int |\chi_r(\mathbf{x}')|^2 \frac{1}{|\mathbf{x} - \mathbf{x}'|} d\mathbf{x}' \right) \chi_r(\mathbf{x}) \\ \hat{K}_s(x)\chi_r(\mathbf{x}) &\left(\int \chi_r^*(\mathbf{x}') \frac{1}{|\mathbf{x} - \mathbf{x}'|} \chi_s(\mathbf{x}')d\mathbf{x}' \right) \chi_s(\mathbf{x}) \end{aligned} \tag{1.15}$$

We now can define the HF Hamiltonian as

$$\hat{H}_0 = \sum_{i=1}^{N} \hat{f}(i) \tag{1.16}$$

and the HF energy as

$$E^{\mathrm{HF}} = \left\langle \Psi^{\mathrm{HF}} \right| \hat{H}_0 \left| \Psi^{\mathrm{HF}} \right\rangle + \left\langle \Psi^{\mathrm{HF}} \right| \hat{V} \left| \Psi^{\mathrm{HF}} \right\rangle \tag{1.17}$$

where $\hat{V}$ is the difference between the HF Hamiltonian and the true system Hamiltonian.

$$\hat{V} = \hat{H} - \hat{H}_0 = \sum_{i}^{N} \sum_{j>i}^{N} \frac{1}{|r_i - r_j|} - \sum_{i}^{N} \left(\hat{J}_i - \hat{K}_i \right) \tag{1.18}$$

1.4 Perturbation Theory

As HF determines the electronic energy of the system at the mean-field level, it is unable to account for the effect of electron correlation. A commonly used method to estimate the correlation energy is through the use of perturbation theory.

1.4.1 Raleigh Schrödinger Perturbation Theory

In general, we can approximate the solution to the exact Hamiltonian eigenvalue equation

$$\hat{H} \left| \Psi \right\rangle = E \left| \Psi \right\rangle \tag{1.19}$$

by separating the exact Hamiltonian into a Hamiltonian with a known solution, $\hat{H}_0$ and the remaining difference, $\hat{V}$. Assuming we know the solution to

$$\hat{H}_0 \left| \Psi^{(0)} \right\rangle = E^{(0)} \left| \Psi^{(0)} \right\rangle \tag{1.20}$$

we can rewrite Eq. 1.19 as

$$(\hat{H}_0 + \lambda \hat{V}) \left| \Psi \right\rangle = E \left| \Psi \right\rangle \tag{1.21}$$

where λ is an ordering parameter which will be set to 1 later. Expanding E and Ψ in a Taylor series, we obtain

$$\left| \Psi \right\rangle = \left| \Psi^{(0)} \right\rangle + \lambda \left| \Psi^{(1)} \right\rangle + \lambda^2 \left| \Psi^{(2)} \right\rangle + \dots \tag{1.22}$$

$$E = E^{(0)} + \lambda E^{(1)} + \lambda^2 E^{(2)} + \dots \tag{1.23}$$

and again rewrite Eq. 1.19 as

$$\begin{aligned} &(\hat{H}_0 + \lambda \hat{V})(\left| \Psi^{(0)} \right\rangle + \lambda \left| \Psi^{(1)} \right\rangle + \lambda^2 \left| \Psi^{(2)} \right\rangle + \dots) \\ &= (E^{(0)} + \lambda E^{(1)} + \lambda^2 E^{(2)} + \dots)(\left| \Psi^{(0)} \right\rangle + \lambda \left| \Psi^{(1)} \right\rangle + \lambda^2 \left| \Psi^{(2)} \right\rangle + \dots) \end{aligned} \tag{1.24}$$

As this thesis will focus on the second-order correction to the HF solution, as the zeroth and first order energies sum to the HF energy[114], we will concentrate on $E^{(2)}$. Focusing on the λ^2 coefficients, we obtain

$$\hat{H}_0 \left|\Psi^{(2)}\right\rangle + \hat{V} \left|\Psi^{(1)}\right\rangle = E^{(0)} \left|\Psi^{(2)}\right\rangle + E^{(1)} \left|\Psi^{(1)}\right\rangle + E^{(2)} \left|\Psi^{(0)}\right\rangle \tag{1.25}$$

Through careful choice of normalization and a bit of manipulation, the expression for the second-order Raleigh Schrödinger perturbation energy can be shown to be

$$E^{(2)} = \sum_{n=1} \frac{\left|\left\langle \Psi_0^{(0)} \right| \hat{V} \left| \Psi_n^{(0)} \right\rangle\right|^2}{E_0^{(0)} - E_n^{(0)}} \tag{1.26}$$

where $\Psi_n^{(0)}$ and $E_n^{(0)}$ are the eigenfunctions and eigenvalues of $\hat{H}_0$ respectively.

1.4.2 Møller–Plesset Perturbation Theory

Although HF fails to account for correlation, the HF solution can be used as the zeroth order solution to which we can apply perturbation theory to estimate the correlation energy. Commonly, this is done with second order Møller–Plesset perturbation theory (MP2). Using Ψ^{HF} and $\hat{H}^{\mathrm{HF}}$ as $\Psi^{(0)}$ and $\hat{H}_{(0)}$ respectively, we can use Equation 1.26 for the second order correction to the HF energy. Due to Brillouin's theorem, single excitations do not mix with the HF solution while triple and higher excitations do not mix either due to Slater-Condon rules. Applying Slater-Condon rules, the second-order correction, in the canonical basis, becomes

$$E^{(2)} = -\frac{1}{4} \sum_{ij}^{\mathrm{occ}} \sum_{ab}^{\mathrm{virt}} \frac{|(ij||ab)|^2}{\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j} \tag{1.27}$$

where the double bar two electron integral is defined as

$$(ij||ab) = (ia|jb) - (ib|ja) \tag{1.28}$$

Alternatively, the MP2 energy can be expressed as

$$E^{(2)} = -\frac{1}{4} \sum_{ij}^{\mathrm{occ}} \sum_{ab}^{\mathrm{virt}} t_{ij}^{ab} (ij||ab) \tag{1.29}$$

where t_{ij}^{ab} are the so-called t-amplitudes

$$t_{ij}^{ab} = \frac{(ab||ij)}{\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j} \tag{1.30}$$

An alternate form of the MP2 energy is the Hylleraas functional[45]. If we equate coefficients of λ^1 in Eq. 1.4.1, we get

$$\hat{H}_0 \left|\Psi^{(1)}\right\rangle + \hat{V} \left|\Psi^{(0)}\right\rangle = E^{(0)} \left|\Psi^{(1)}\right\rangle + E^{(1)} \left|\Psi^{(0)}\right\rangle \tag{1.31}$$

which can be rewritten as

$$(\hat{H}_0 - E^{(0)}) \left|\Psi^{(1)}\right\rangle = (E^{(1)} - \hat{V}) \left|\Psi^{(0)}\right\rangle \tag{1.32}$$

Multiplying by $\left\langle\Psi^{(1)}\right|$ and equating to 0, we get

$$0 = \left\langle\Psi^{(1)}\right| \hat{H}_0 - E^{(0)} \left|\Psi^{(1)}\right\rangle + \left\langle\Psi^{(1)}\right| \hat{V} - E^{(1)} \left|\Psi^{(0)}\right\rangle \tag{1.33}$$

We then add the equation for the second-order energy,

$$E^{(2)} = \left\langle\Psi^{0}\right| \hat{V} - E^{(1)} \left|\Psi^{(1)}\right\rangle \tag{1.34}$$

which is obtained obtained by multiplying 1.4.1 by $\left\langle\Psi^{(0)}\right|$, and add it to Eq. 1.4.2 to get

$$E^{(2)} = \left\langle\Psi^{(1)}\right| \hat{H}_0 - E^{(0)} \left|\Psi^{(1)}\right\rangle + \left\langle\Psi^{(1)}\right| \hat{V} - E^{(1)} \left|\Psi^{(0)}\right\rangle + \left\langle\Psi^{0}\right| \hat{V} - E^{(1)} \left|\Psi^{(1)}\right\rangle \tag{1.35}$$

invoking $\left\langle\Psi^{(1)}\middle|\Psi^{(0)}\right\rangle = 0$, we obtain the Hylleraas functional expression for the MP2 energy

$$J_H = \left\langle\Psi^{(1)}\right| \hat{H}_0 - E^{(0)} \left|\Psi^{(1)}\right\rangle + \left\langle\Psi^{(1)}\right| \hat{V} \left|\Psi^{(0)}\right\rangle + \left\langle\Psi^{0}\right| \hat{V} \left|\Psi^{(1)}\right\rangle \tag{1.36}$$

in terms of t-amplitudes, this expression becomes

$$J_H = t^{\dagger}\Delta t + t^{\dagger}V + V^{\dagger}t \tag{1.37}$$

where V is the two electron integral $(ij||ab)$ and Δ is an eighth-rank tensor which will be elaborated in more detail in chapter 2. The Hylleraas functional is equal to the MP2 energy when optimized with respect to t-amplitudes

$$E^{(2)} = \min_{t}(J_H) \tag{1.38}$$

This form of the MP2 correlation energy is useful for evaluating an MP2 correlation, as it is valid even when using t-amplitudes that are not optimized for the functional.

1.5 Energy Decomposition Analysis

In electronic structure, it is common to calculate the interaction energy of two or more fragments, ΔE_{INT}, of a system involving intermolecular interactions by calculating the difference in energy between the fragments interacting in the supersystem and the fragments isolated from each other,

$$\Delta E_{\mathrm{INT}} = E(AB) - E(A) - E(B) \tag{1.39}$$

However, when most people think about intermolecular interactions, they do not think solely of the strength of interaction, but also the phenomena, such as electrostatics and dispersion forces, that contribute to the interaction. While calculating the strength of the interaction

is useful, this does not give direct insight to the chemical phenomena that determines the strength of this interaction. An energy decomposition analysis (EDA) aims to decompose the interaction energy into chemically meaningful components such as permanent electrostatics and dispersion. Understanding the underlying contributions to the intermolecular interaction in systems such as gas storage materials is extremely valuable in developing new materials based on their chemical and physical properties.

The first popular EDA was developed by Kitaura and Morokuma for HF[60] which decomposed the interaction energy into electrostatics, exchange, polarization, and charge transfer. However, this EDA did not use properly anti-symmetric wave functions which led to non-physical results at large basis sets[26]. Many other different EDA's have since been developed, and each with differing components and definitions of the components of the interaction energy. Two common approaches to EDA are the variational approach and the perturbative approach. One example of a variational EDA is an EDA based on localized molecular orbitals(LMO-EDA) which decomposes the interaction into an electrostatic, exchange, repulsion, polarization, and dispersion terms[112].

$$\Delta E_{\mathrm{INT}} = \Delta E_{\mathrm{ele}}^{\mathrm{LMO}} + \Delta E_{\mathrm{ex}}^{\mathrm{LMO}} + \Delta E_{\mathrm{rep}}^{\mathrm{LMO}} + \Delta E_{\mathrm{pol}}^{\mathrm{LMO}} + \Delta E_{\mathrm{disp}}^{\mathrm{LMO}} \tag{1.40}$$

The electrostatic term, $E_{\mathrm{ele}}^{\mathrm{LMO}}$ is obtained using the difference between an approximate energy expression for a supersystem consisting of the direct product of the monomer wave function, $\left|\Psi^{\mathrm{DP}}\right\rangle$ and the energy of the isolated monomers.

$$\Delta E_{\mathrm{ele}}^{\mathrm{LMO}} = \left\langle \Psi^{\mathrm{DP}} \right| \hat{F} \left| \Psi^{\mathrm{DP}} \right\rangle - \sum_{M}^{\mathrm{frags}} \left\langle \Psi_{\mathrm{M}} \right| \hat{F} \left| \Psi_{\mathrm{M}} \right\rangle + \Delta V_{\mathrm{NN}} \tag{1.41}$$

Where $\hat{F}$ is $\hat{h} + \frac{1}{2}(\hat{J} + \hat{K})$ and ΔV_{NN} is the nuclear-nuclear repulsion energy between fragments. The exchange term, $\Delta E_{\mathrm{ex}}^{\mathrm{LMO}}$ comes from enforcing antisymmetry on the direct product wavefunction to form an antisymmetric direct product of the monomers' wavefunction $\left|\Psi^{\mathrm{AS}}\right\rangle$

$$\Delta E_{\mathrm{ex}}^{\mathrm{LMO}} = \left\langle \Psi^{\mathrm{AS}} \right| \hat{F} \left| \Psi^{\mathrm{AS}} \right\rangle - \left\langle \Psi^{\mathrm{DP}} \right| \hat{F} \left| \Psi^{\mathrm{DP}} \right\rangle \tag{1.42}$$

while the repulsion term, $E_{\mathrm{rep}}^{\mathrm{LMO}}$, reflects the energy penalty to orthonormalize $\left|\Psi^{\mathrm{AS}}\right\rangle$ to form $\left|\Psi^{\mathrm{ASN}}\right\rangle$ and is defined as

$$\Delta E_{\mathrm{rep}}^{\mathrm{LMO}} = \left\langle \Psi^{\mathrm{ASN}} \right| \hat{F} \left| \Psi^{\mathrm{ASN}} \right\rangle - \left\langle \Psi^{\mathrm{AS}} \right| \hat{F} \left| \Psi^{\mathrm{AS}} \right\rangle \tag{1.43}$$

The remainder of the SCF energy is defined as the polarization energy $\Delta E_{\mathrm{pol}}^{\mathrm{LMO}}$. As such, LMO-EDA does not have a distinction between polarization and charge transfer effects and instead combines both into the polarization term. Additionally, the dispersion term in LMO-EDA, $\Delta E_{\mathrm{disp}}^{\mathrm{LMO}}$, is calculated as the difference between calculations with post-HF methods (MP2, coupled cluster) and HF interaction energies. For instance, for MP2 the LMO dispersion energy would be calculated as

$$\Delta E_{\mathrm{disp}}^{\mathrm{LMO}} = \Delta E_{\mathrm{INT}}^{(2)} \tag{1.44}$$

While dispersion is a purely correlational effect and is not captured in HF, the correlation interaction includes other effects besides dispersion such as electrostatic effects like the correction to dipole moments which HF tends to overestimate[40]. Additionally, dispersion is a purely attractive interaction and previous work has shown that treating the correlation component as just dispersion can lead to nonsensical results [118].

A popular perturbative approach to EDA is through symmetry-adapted perturbation theory (SAPT)[50, 115, 92]. SAPT treats intermolecular interaction as a perturbation to the noninteracting isolated system. The zeroth-order Hamiltonian for an A/B dimer for SAPT is

$$\hat{H}_0 = \hat{F}^A + \hat{F}^B \tag{1.45}$$

the full Hamiltonian is written as

$$\hat{H} = \hat{F}^A + \hat{F}^B + \xi\hat{W}^A + \nu\hat{W}^B + \zeta\hat{V} \tag{1.46}$$

where $\hat{V}$ is the intermolecular perturbation and $\hat{W}^A$ is a Møller-Plesset fluction operator for fragment A. SAPT0 is the simplest form of SAPT and treats the monomers with HF and incldues dispersion terms from second order perturbation.

$$E_{\text{INT}}^{\text{SAPT0}} = E_{\text{ele}}^{(1)} + E_{\text{ex}}^{(1)} + E_{\text{pol}}^{(2)} + E_{\text{disp}}^{(2)} \tag{1.47}$$

SAPT decomposes the interaction energy into electrostatics, exchange, induction, exchange induction, dispersion, and exchange dispersion. Like LMO-EDA, SAPT does not separate polarization and charge transfer effects and instead uses an "induction" term to capture these effects.

As the decomposition of interaction energy is not unique, it is useful to establish some ideal properties of an EDA. Firstly, each term in the EDA should have a stable and meaningful basis set limit. As each term is calculated at the same level of theory as the corresponding interaction energy, we expect that if the total interaction energy is stable with respect to basis set, each term should also be stable with respect to basis set. Additionally, we expect that each EDA term have correct asymptotic behavior. Finally, an ideal EDA should not rely on specific types of basis functions such as atomic orbitals, but instead should also be calculated using other types of basis sets such as plane waves.

An EDA based on absolutely localized molecular orbitals (ALMO) has been developed at Berkeley for HF, DFT, and MP2[59][116] and satisfies many of the idea properties prior stated. ALMO EDA decomposes the interaction energy into a frozen term, which includes contributions from permanent electrostatics and Pauli repulsion, a polarization term, a dispersion term, and a charge transfer term. The first generation ALMO-EDA (ALMO-EDA-I) has been updated (ALMO-EDA-II) to address the issues with meaningful basis set limits for the polarization and charge transfer term for HF and DFT. However, the MP2 ALMO-EDA builds upon ALMO-EDA-I and additionally, basis set stability issues have been discovered for all four terms. As such, an update to MP2 ALMO-EDA-I is much needed both to update it to serve as a correction to HF ALMO-EDA-II as well as provide a post-HF EDA with stable basis set limits and serve as a framework for other post-HF EDA's. A brief summary on each term in the MP2 ALMO EDA is as follows.

1.5.1 Frozen

The frozen term includes contributions from permanent electrostatics and Pauli repulsion. This is calculated by calculating the interaction in a molecular system where each molecule does not polarize or relax due to the presence of other fragments. As the molecules are not allowed to relax, one contribution to this interaction is the effect of the permanent electrostatics of each molecule interacting with others such as dipole-dipole, charge-dipole, and dipole-quadrupole interactions. This, however, is only well-defined when there is no overlap between fragments. As the wave functions of the isolated fragments may overlap, there is an energy penalty associated with enforcing anti-symmetry of the frozen density, which is attributed to Pauli repulsion. As this is an energy penalty, Pauli repulsion is strictly a repulsive interaction. For SCF ALMO-EDA, the frozen energy, ΔE_{FRZ} is defined as the SCF energy change corresponding to bringing infinitely separated fragments into the supersystem geometry and corresponding change to form the frozen wavefunction Ψ_0 which is a properly antisymmetrized wavefunction constructed from unrelaxed nonorthogonal MO's of fragments

$$\Delta E_{\mathrm{FRZ}}^{\mathrm{SCF}} = E^{\mathrm{SCF}}(\Psi_0) - \sum_{F}^{\mathrm{frag}} E^{\mathrm{SCF}}(\Psi_F) \tag{1.48}$$

1.5.2 Polarization

Polarization arises when each molecule's electron density is allowed to relax in the field of other molecules. As this is a relaxation term, polarization is strictly attractive. To separate polarization from charge transfer, it is important to enforce the constraint that electron density remains on the fragment. This is achieved by relaxing the frozen orbitals using SCF for molecular interactions (SCF-MI)[29, 108]. ALMO-EDA-I lacks a useful basis set limit for the charge transfer and polarization at large basis sets. To alleviate this issue, ALMO-EDA-II introduced polarization subspaces in the form of fragment electric field response functions (FERFs)[43]. FERFs are created from the response of the isolated fragments to weak electric fields. To calculate the FERFs, the solution to the coupled perturbed SCF (CPSCF)

$$E^{\Delta\Delta} \cdot \Delta^{\mu} = -E^{\Delta h} \cdot h^{\mu} \tag{1.49}$$

where $E^{\Delta\Delta}$ is the SCF orbital Hessian, Δ^{μ} is is the orbital response of a fragment to the field component, F_{μ}, and $E^{\Delta h} \cdot h^{\mu}$ is related to the multipole moment matrix, M, through

$$E^{\Delta h} \cdot h^{\mu} \equiv \frac{\partial^2 E}{\partial \Delta^{ai} \partial h_{\lambda\sigma}} = 2(M_{\mu})_{ai} \tag{1.50}$$

The fragment orbital response matrix, Δ^{μ} is used to obtain the subspace of the virtual orbitals which participate in the SCF-MI using Singular Value Decomposition,

$$\Delta^{\mu} = L_{\mu} d_{\mu} R_{\mu}^{T} \tag{1.51}$$

L_{μ} is a unitary transformation of virtual orbitals which is used to create the virtual subspace used for SCF-MI.

1.5.3 Dispersion

Dispersion is a purely correlational effect and is not captured by HF due to this. Dispersion is associated with the interaction due to instantaneous multipole moments of each fragment. The standard way for DFT ALMO-EDA to describe dispersion is reliant on a semi-empirical dispersion correction and therefore, MP2 ALMO-EDA is the first wavefunction-based ALMO-EDA to have an ab initio description of dispersion. The MP2 correlation energy of two fragments,

$$E_{AB}^{(2)} = -\frac{1}{4}\sum_{ij}^{\text{occ}}\sum_{ab}^{\text{virt}}\frac{|(ij||ab)|^2}{\epsilon_a+\epsilon_b-\epsilon_i-\epsilon_j} \tag{1.52}$$

can be separated based on the type of two electron integrals, and also noting that,

$$\begin{aligned} E_{AB}^{(2)} &= -\frac{1}{4}\sum_{i_Aj_A}^{\text{occA}}\sum_{a_Ab_A}^{\text{virtA}}\frac{|(i_Aj_A||a_Ab_A)|^2}{\epsilon_{a_A}+\epsilon_{b_A}-\epsilon_{i_A}-\epsilon_{j_A}} - \frac{1}{4}\sum_{i_Bj_B}^{\text{occB}}\sum_{a_Bb_B}^{\text{virtB}}\frac{|(i_Bj_B||a_Bb_B)|^2}{\epsilon_{a_B}+\epsilon_{b_B}-\epsilon_{i_B}-\epsilon_{j_B}} + E_{\text{INT}}^{(2)} \\ &= E_{\text{A}}^{(2)} + E_{\text{B}}^{(2)} + E_{\text{INT}}^{(2)} \end{aligned} \tag{1.53}$$

In the long range, where integrals of type $(i_Aa_B|j_Bb_A) = 0$, the interaction energy can be written as

$$\lim_{R\to\infty} E_{\text{INT}}^{(2)} = -\sum_{i_A}^{\text{occA}}\sum_{j_B}^{\text{occB}}\sum_{a_A}^{\text{virtA}}\sum_{b_B}^{\text{virtB}}\frac{|(i_Aa_A|j_Bb_B)|^2}{\epsilon_{a_A}+\epsilon_{b_B}-\epsilon_{i_A}-\epsilon_{j_B}} \tag{1.54}$$

Using the multipole expansion, as per Stone [110], and truncating at the leading dipole-dipole term, we can rewrite Eq. 1.5.3 as

$$\begin{aligned} \lim_{R\to\infty} E_{\text{INT}}^{(2)} &= -\sum_{i_A}^{\text{occA}}\sum_{j_B}^{\text{occB}}\sum_{a_A}^{\text{virtA}}\sum_{b_B}^{\text{virtB}}\frac{|(i_Aa_A|\hat{\mu}_A^{\alpha}T^{\alpha\beta}\hat{\mu}_B^{\beta}|j_Bb_B)|^2}{\epsilon_{a_A}+\epsilon_{b_B}-\epsilon_{i_A}-\epsilon_{j_B}} \\ &= -(T^{\alpha\beta})^2\sum_{i_A}^{\text{occA}}\sum_{j_B}^{\text{occB}}\sum_{a_A}^{\text{virtA}}\sum_{b_B}^{\text{virtB}}\frac{|(i_A|\hat{\mu}_A^{\alpha}|a_A)|^2|(j_B|\hat{\mu}_B^{\beta}|b_B)|^2}{\epsilon_{a_A}+\epsilon_{b_B}-\epsilon_{i_A}-\epsilon_{j_B}} \end{aligned} \tag{1.55}$$

where α and β denote the Cartesian directions for the dipoles, $\hat{\mu}$ is the dipole operator, and $T^{\alpha\beta}$, contains the dipole R^{-3} dependence on intermolecular distance. Using the identity,

$$\frac{1}{A+B} = \frac{2}{\pi}\int_0^{\infty}\frac{AB}{(A^2+v^2)(B^2+v^2)}dv \tag{1.56}$$

and writing the denominator as $\epsilon_{a_A}-\epsilon_{i_A}+\epsilon_{b_B}-\epsilon_{j_B} = \hbar(\omega_{ai}^A+\omega_{bj}^B)$ we can rewrite Eq. 1.5.3 as

$$\lim_{R\to\infty} E_{\text{INT}}^{(2)} = -\frac{2\hbar}{\pi}(T^{\alpha\beta})^2\int_0^{\infty}\sum_{i_A}^{\text{occA}}\sum_{a_A}^{\text{virtA}}\frac{\omega_{ai}^A|(i_A|\hat{\mu}_A^{\alpha}|a_A)|^2}{\hbar(\omega_{ai}^A+v^2)}\sum_{j_B}^{\text{occB}}\sum_{b_B}^{\text{virtB}}\frac{\omega_{bj}^B|(j_B|\hat{\mu}_B^{\beta}|b_B)|^2}{\hbar(\omega_{bj}^B+v^2)}dv \tag{1.57}$$

by setting $v = \mathrm{i}u$, we identify that the expression in the integral is equivalent to the product of the molecular polarizabilities, α, of molecules A and B at the imaginary frequency $\mathrm{i}u$

$$\lim_{R\to\infty} E_{\mathrm{INT}}^{(2)} = -\frac{2\hbar}{\pi}(T^{\alpha\beta})^2 \int_0^\infty \alpha_\alpha^A(\mathrm{i}u)\alpha_\beta^B(\mathrm{i}u)du \tag{1.58}$$

thus, obtain an expression for the long range Van der Waals interactions to second-order with the leading R^{-6} term.

1.5.4 Charge Transfer

Lastly, charge transfer is associated with the transfer of electron density between fragments. This interaction is purely attractive as it is related to the molecules relaxing due to the donation of electron density with each other. As this relies on the transfer of electron density between fragments, charge transfer is a short range interaction.

1.6 Outline

This thesis starts with the implementation and theory behind MP2 ALMO-EDA-II in Chapter 2. Additionally, Chapter 2 evaluates the performance and basis set stability of MP2 ALMO-EDA-II compared to the original MP2 ALMO-EDA-I as well as DFT ALMO-EDA-II. Chapter 3 then investigates four classes of complexes with varying strengths of intermolecular interactions. This serves to explore MP2 ALMO-EDA-II's usefulness in chemical applications as well as resolving the physical origins of these interactions.

1.6.1 Chapter 2: A second generation energy decomposition analysis of intermolecular interaction energies from second order Møller-Plesset theory: An extensible, orthogonal formulation with useful basis set limits for all terms.

Energy decomposition analysis (EDA) has become widely used to aid in understanding the nature of intermolecular interactions, based on density functional theory (DFT) and self-consistent field (SCF) calculations. There has been less development of EDA methods using correlated post-SCF methods, partly because of the necessarily greater complexity associated with defining suitable approaches. This work reports a new approach to EDA for post-SCF wavefunctions, together with a complete implementation for restricted and unrestricted second-order Møller-Plesset (MP2) calculations. The new approach is an MP2 generalization of the successful SCF-level second-generation absolutely localized molecular orbital EDA approach (ALMO-EDA-II). The new MP2 ALMO-EDA-II provides distinct energy contributions for a frozen interaction energy (containing permanent electrostatics and

Pauli repulsions), a polarized energy (yielding induced electrostatics), a dispersion-corrected energy, and the fully relaxed energy (which yields charge transfer). All terms have useful complete basis set limits due to the design of the theory, and as corroborated by a range of test calculations on model systems, and the S22 and the Ionic43 datasets of weak and strong intermolecular interactions, respectively.

1.6.2 Chapter 3: A second generation energy decomposition analysis of intermolecular interaction energies from second order Møller-Plesset theory: Chemical examples and applications.

A newly developed MP2 ALMO-EDA-II has been applied to four non-covalently bonded classes of complexes; a class of conventional hydrogen bonded systems, a class of non-conventionally hydrogen bonded systems, a class of tetrel bonded systems, and a "solvent-resistant" halogen bonded system. Through these systems, the importance of correctly treating the correlation component of the interaction energy in MP2 ALMO-EDA-II is shown.

Chapter 2

A second generation energy decomposition analysis of intermolecular interaction energies from second order Møller-Plesset theory: An extensible, orthogonal formulation with useful basis set limits for all terms.

2.1 Introduction

Second-order Møller-Plesset theory (MP2) is the simplest wavefunction approach to correcting the mean field Hartree-Fock (HF) method for the effects of electron correlation.[17] Despite its simplicity, MP2 itself yields excellent accuracy[25] for many intermolecular interactions, including water clusters,[23, 127, 94] ion-water interactions,[63, 62] and ionic liquids.[128, 46] Furthermore, from the viewpoint of this work, methods for analyzing intermolecular interactions calculated using MP2 are potentially extensible to more exact theories of electron correlation, particularly coupled cluster methods truncated at the level of single and double substitutions.[7, 6] Finally, there have been quite successful efforts to improve the accuracy of MP2 by semi-empirical modifications.[34, 54, 33, 19, 89, 31, 30, 121] One of the most exciting of those is the recent use of a single regularization parameter[70, 104] which is both physically motivated, and quite transferable, yielding a several-fold reduction in root mean square errors for MP2 intermolecular interaction energies of both small[55] and large[103] molecules. Regularized MP2 itself was inspired by the success of regularization to improve orbital optimized MP2 (OOMP2).[67, 66]

Our goal in this work is to report on the development of a new energy decomposition analysis (EDA) for intermolecular interactions at the MP2 level of theory, as the simplest representative of post-HF methods. As reviewed for density functional theory (DFT) and HF methods recently,[75] an EDA aims to separate the contributions of an intermolecular interaction evaluated via a supermolecule calculation:

$$\Delta E_{\mathrm{INT}} = E_{\mathrm{FULL}} - \sum_{F} E_F \tag{2.1}$$

into a set of additive contributions with a clear physical origin. As a result, an EDA can help to provide some insight for interpretative purposes. In particular, for a post-SCF EDA, we will aim to define SCF and post-SCF contributions to each of the following contributions:

$$\Delta E_{\mathrm{INT}} = \Delta E_{\mathrm{GD}} + \Delta E_{\mathrm{FRZ}} + \Delta E_{\mathrm{POL}} + \Delta E_{\mathrm{DISP}} + \Delta E_{\mathrm{CT}} \tag{2.2}$$

The strategy will be to impose constraints on the correlation amplitudes so that the energy differences defined in Eq. 3.2 can be obtained by subtraction between each consecutive pair of the following 6 separate states of the electrons:

1. Separate fragments (E_F for each fragment F) at their optimal geometry.
2. Separate fragments (E_F for each fragment F) at the geometry of the complex.
3. A "frozen complex" (E_{FRZ}) which uses correlation amplitudes (and orbitals) optimized for the isolated fragments (at the complex geometry).
4. A "polarized complex" (E_{POL}) in which fragment amplitudes (and orbitals) are re-optimized in the environment of the complex but are constrained to have fragment identity.
5. A "dispersion-including complex" (E_{DISP}) in which additional *inter-fragment* amplitudes are included which describe dispersion, still using the polarized orbitals.
6. The fully relaxed complex (E_{FULL}) in which both orbitals and amplitudes are permitted to relax.

This strategy gives a *variational* character to the EDA: each successive energy after E_{FRZ} is less than or equal to the previous one (even if the final energy itself is not guaranteed to be variational with respect to the true energy).

At the SCF level, there has been extensive development of EDAs, beginning with the seminal Kitaura-Morokuma method,[60, 86] and followed by many successors[130, 129, 12, 82, 112, 113] which we cannot adequately review here. However, we should single out the absolutely localized molecular orbital (ALMO)-EDA,[75] which does follow the variational approach advocated for above, by using fragment-blocked self-consistent field (SCF) for molecular interactions (SCF-MI)[108, 29, 88, 58] to treat polarization. The first generation

ALMO-EDA[59, 57, 44] is largely the same as the alternative block-localized wavefunction (BLW)-EDA approach[84, 85, 83]. These methods are successful when used as recommended, but can be criticized for being directly tied to the use of atom-centered atomic orbitals (a good EDA should be representation-independent), as well as lacking a well-defined complete basis set limit for all terms (a good EDA should provide such limits).[4, 41, 64] The second generation ALMO-EDA[43, 77] lifted both of those difficulties,[41] and also introduced a new approach to separate electrostatics, Pauli repulsion and dispersion within ΔE_{FRZ}.[42] While development continues,[74, 122] the second generation ALMO-EDA is a useful approach for analyzing intermolecular interaction energy calculations at the DFT level. The goal of our work is to develop a suitable post-SCF counterpart and prototype it with MP2 theory.

The simplest possible approach to including electron correlation in a post-SCF EDA is to treat the correlation effect as simply an additional term that does not exist at the SCF level.[24, 112] For instance, the only term in Eq. 3.2 that does not exist at the SCF level is the effect of dispersion or London forces, ΔE_{DISP}. However, it has been demonstrated,[118] and is perhaps intuitive, that since charge distributions, polarizabilities, and donor-acceptor interactions all change upon inclusion of electron correlation, *all* terms of Eq. 3.2 are affected by electron correlation. Other post-SCF supermolecular EDAs[107, 102, 5, 47, 101, 2, 28] therefore make efforts to include such contributions. One effort by our group achieves this goal for all terms, in the sense that it directly generalizes the first generation of the ALMO-EDA to MP2 theory.[116, 117, 69] However, just as ALMO-EDA-I is specifically tied to AO basis sets, and lacks well-defined CBS limits for all terms, so too is that first generation ALMO-EDA for MP2 theory (henceforth denoted as MP2 ALMO-EDA-I. This work aims to lift those limitations by designing a post-SCF version of the second generation SCF-level ALMO-EDA, resulting in a method we will denote as MP2 ALMO-EDA-II.

The remainder of the paper is arranged as follows. Section 2.2 (Theory) describes our strategy for defining each of the constrained MP2 energies of an intermolecular complex ($E_{\mathrm{FRZ}}, E_{\mathrm{POL}}, E_{\mathrm{DISP}}$). Together with the isolated fragment energies and the final fully relaxed complex energy, this enables the evaluation of each term of Eq. 3.2. Section 2.3 (Implementation) discusses computationally relevant aspects of MP2 ALMO-EDA-II, including its extension to regularized MP2.[70, 104]. It also provides some characterization of the computational costs associated with each step. A series of model problems and applications of MP2 ALMO-EDA-II to a broad range of intermolecular interactions ranging from very weak to very strong are presented in Section 2.4. Finally, we offer some conclusions and perspectives in Section 2.5.

2.2 Theory

The MP2 ALMO EDA decomposes the correlation part of the interaction energy into four components, frozen (FRZ), polarization (POL), dispersion (DISP), and charge transfer (CT). Each of these terms corrects the corresponding term at the SCF level except for the dispersion term, which is purely a correlated term and does not exist at the mean-field HF level. Like the

previously published MP2 ALMO EDA (MP2 ALMO-EDA-I), the intermediate energies are evaluated as MP2 energies using the intermediate orbitals calculated from the SCF ALMO-EDA allowing the terms to be added as corrections to the corresponding HF component except for dispersion.[118] Unlike MP2 ALMO-EDA-I which extends SCF ALMO-EDA-I, the second generation MP2 ALMO-EDA extends the SCF ALMO-EDA-II.

For the remainder of the paper, the energies $E^{(2)}$ refer to MP2 correlation energies evaluated for either a fragment (F) or the system (FULL), denoted by the first value in the parentheses, and using the basis and Fock matrix denoted by the second value in the parentheses. The individual EDA terms are denoted by ΔE with the subscript indicating the specific EDA term and the optional superscript indicating whether it is the correlation component (2), Hartree Fock component (SCF), or total MP2 energy denoted by the lack of a superscript. Occupied orbitals are denoted by the indexes i, j, k... and virtual orbitals are denoted by the indexes a, b, c...

2.2.1 MP2 Hylleraas functional

The energy components in MP2 ALMO-EDA are evaluated with in various bases which are non-canonical and non-orthogonal; thus, the standard MP2 energy expression for a spin-restricted system is given in terms of spin orbitals as:

$$E^{(2)} = -\frac{1}{4}\sum_{ijab}\frac{|(ij\|ab)|^2}{\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j} \tag{2.3}$$

which requires a (pseudo) diagonal Fock operator. Therefore, the Hylleraas functional formalism is employed, which is a variational formalism to obtain the MP2 energy. The functional J_H[45] is defined as follows:

$$J_H = \left\langle \Psi^{(1)} \middle| \hat{H}_0 - E_0 \middle| \Psi^{(1)} \right\rangle + \left\langle \Psi^{(1)} \middle| \hat{V} \middle| \Psi^{(0)} \right\rangle + \left\langle \Psi^{(0)} \middle| \hat{V} \middle| \Psi^{(1)} \right\rangle \tag{2.4}$$

where the Fock operator ($\hat{F}$) is commonly used as the zeroth-order Hamiltonian. When minimized, the Hylleraas functional yields the t amplitudes that provide the optimal MP2 energy in that same basis (i.e. identical to Eq. 2.3. The Hylleraas functional is an expression for the MP2 energy that is valid for t amplitudes that are not necessarily solved from the same MP2 energy that they are being used to evaluate (i.e. t amplitudes that yield a higher-than-optimal MP2 energy).

2.2.2 Strong Orthogonality

As a general requirement, the virtual space needs to be orthogonal to the entire occupied space. However, this is generally not the case for ALMO orbitals as they are only orthogonal *within* a fragment ($\sigma_{kl}^{FG} \neq 0$, fragment $F \neq G$; σ defines the MO metric tensor).

To ensure that each fragment's virtual space is orthogonal to the occupied space of other fragments, each fragment's virtual space is projected into the orthogonal complement of the entire occupied space to form the projected frozen (pFRZ) basis from the isolated fragment basis and ALMO (pALMO) basis from the SCF-MI fragment basis.

$$V_p = (1 - P)V \tag{2.5}$$

Where V is the original virtual space, P is the projector of the occupied space, and V_p is the projected virtual space (either pFRZ or pALMO). For the pALMO and pFRZ basis, the occupied orbitals are symmetrically orthogonalized between fragments for significant gains in computational simplicity with minimal loss of locality.

Unlike MP2 ALMO-EDA-I, we additionally define the projected isolated (pISO) basis, which is identical to the pFRZ basis, except we do not orthogonalize the occupied space of the isolated fragments. The pISO basis is used for the purposes of calculating t amplitudes to use in the frozen Hylleraas functional and differs from MP2 ALMO-EDA-I which just uses the t amplitudes in the isolated fragment basis (ISO)[118]. This change is necessary to maintain a well-defined basis set limit for the frozen and polarization terms in MP2 ALMO-EDA-II.

2.2.3 Global Virtual Basis

A significant change in the orbital representation of the virtual space relative to MP2 ALMO-EDA-I is the introduction of the orthogonal global virtual (GV) basis. The GV basis is created as the union of the virtual spaces of each fragment. The resulting set of crude virtual functions is then orthogonalized, and any linear dependencies are removed to define the full GV space. This virtual space is spanned by GV's that are orthonormal, and, as the name indicates, global in character, and thus shared between all fragments rather than being fragment assignable. The motivation for defining the GV space in terms of orthonormal virtuals is two-fold: first, complexities associated with using non-orthogonal functions[39] are avoided, and second, their nature is independent of the type of underlying basis function (in contrast to fragment-tagged virtuals that derive specifically from atomic orbitals).

As was evident in Eq. 2.3, the MP2 correction to HF is a linear combination of doubly excited Slater determinants due to the Slater-Condon rules [106]. With a GV basis, the virtual space is no longer fragment assignable, but this poses relatively little problem for generalizing the ALMO-EDA-I definitions of the pair correlation correction to the frozen and polarized energies (or the Hylleraas functional). These correlation corrections involve sums over pairs of occupied orbitals that are only on the same fragment, as schematically illustrated in the first row of Figure 3.1.

Sums over virtual orbitals are over pairs of GV orbitals, which means that the resulting fragment contributions to the frozen and polarized correlation energy now include excitations of the type shown in the third row of Figure 3.1. These additional excitations are best viewed as the correlation part of the basis set superposition error (BSSE) term, where the electrons of one fragment attain some small energy lowering due to using basis functions derived from

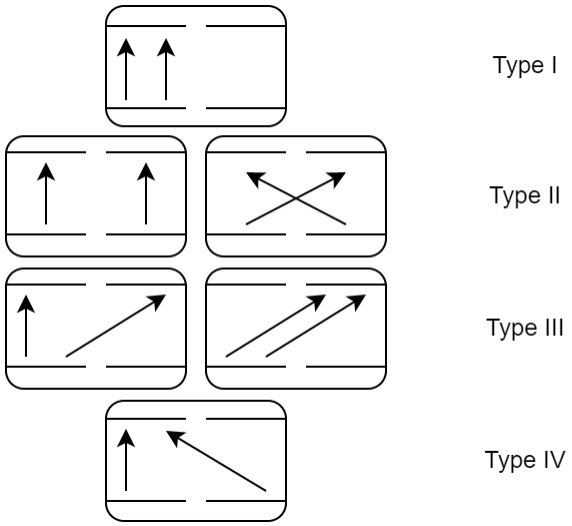


Figure 2.1: The four types of double excitations used to define constraints for each EDA term. Occupied orbitals are fragment-tagged, and this diagram should be read as though the virtuals are also fragment-tagged. The true correlation energy associated with a fragment at either the frozen or polarized level is then represented by the first row of the figure (Type I). The third line (Type III), represent basis set superposition effects, that can be removed by counterpoise correction.The second and fourth line (Type II/IV) are associated with dispersion interactions.

other fragments. Using the standard Boys-Bernardi counterpoise correction CP scheme[13], we see that the correlation part of the CP correction should be applied at the frozen energy level.

As there are many intermediate bases and allowed excitations for each intermediate energy, it is useful to summarize the basis and excitation constraints used for each intermediate level.

2.2.4 Frozen Energy

The first term in the MP2 ALMO-EDA-II is the so-called frozen part of the correlation energy. This is the intra-fragment correlation evaluated with correlation amplitudes obtained from isolated fragment calculations. We thereby associate the $\Delta E^{(2)}_{\mathrm{FRZ}}$ contribution with correlation-derived changes in Pauli repulsion and permanent electrostatics. Specifically:

$$\Delta E^{(2)}_{\mathrm{FRZ}} = \sum_{F}^{\mathrm{frag}} E^{(2)}_{\mathrm{FRZ}}(\xi^{\mathrm{ISO}}, \xi^{\mathrm{ISO}}_{\mathrm{p,GV}}, \xi^{\mathrm{FRZ}}_{\mathrm{p,GV}}) - \sum_{F}^{\mathrm{frag}} E^{(2)}_{\mathrm{F}}(\xi^{\mathrm{ISO}}) \tag{2.6}$$

We calculate $E^{(2)}_{\mathrm{FRZ}}$ using the Hylleraas functional and an additional term involving the relaxed second-order density matrix, P_{ia} to include first-order orbital response effects for the

Symbol	Basis	System Orthogonal Occupied Space?	Projected Virtual Space?	Excitation Types
ξ^{ISO}	ISO	N	N	I
$\xi^{\mathrm{ISO}}_{\mathrm{p,GV}}$	pISO	Y	N	I, III
$\xi^{\mathrm{FRZ}}_{\mathrm{p,GV}}$	pFRZ	Y	Y	I, III
$\xi^{\mathrm{ALMO}}_{\mathrm{p,GV}}$	pALMO	Y	Y	I, III
$\Xi^{\mathrm{CCC}}_{\mathrm{p,GV}}$	pALMO	Y	Y	I, II, III, IV
Ξ^{FULL}	FULL	N/A	N/A	I, II, III, IV

Table 2.1: Overview of the different intermediate basis used through out the scheme; ξ is defined for each fragment, Ξ is defined for the system

frozen term[37]:

$$\begin{aligned} E^{(2)}_{\mathrm{FRZ}} &= J_H[t(\xi^{\mathrm{ISO}}_{\mathrm{p,GV}}), M(\xi^{\mathrm{FRZ}}_{\mathrm{p,GV}})] - 2\sum_{ia} P_{ia}(\xi^{\mathrm{ISO}}) F_{ia}(\xi^{\mathrm{FRZ}}_{\mathrm{p,GV}}) \\ &= \frac{1}{2}\sum_{ijab} t^{ab}_{ij}(ij||ab) + \sum_{ij} P_{ij}F_{ij} + \sum_{ab} P_{ab}F_{ab} - 2\sum_{ia} P_{ia}F_{ia} \end{aligned} \tag{2.7}$$

where M refers to any matrix that is not explicitly a function of t and

$$P_{ij} = -\sum_{abk} t^{ab}_{ik} t^{ab}_{jk} \tag{2.8a}$$

$$P_{ab} = \sum_{ijc} t^{ac}_{ij} t^{bc}_{ij} \tag{2.8b}$$

The second term (i.e. involving the relaxed MP2 density contributions, P_{ia}) is required because the frozen orbitals do not make the SCF energy stationary with respect to orbital rotations (i.e. $F_{ia}(\xi^{\mathrm{FRZ}}_{\mathrm{p,GV}}) \neq 0$). The fragment t amplitudes in equations 2.7-2.8 are calculated in the pISO basis, while the two electron integral and Fock matrix are calculated in the pFRZ basis. The *ov* block of the second-order density matrix, P_{ia} is related to the Z-vector[37] by

$$P_{ia} = -Z_{ia}. \tag{2.9}$$

The Z-vector is defined using the inverse of the HF electronic hessian, $\mathcal{E}^{\theta\theta}$, and the MP2 orbital gradient, L:

$$z^{\dagger} = L\left(\mathcal{E}^{\theta\theta}\right)^{-1} \tag{2.10}$$

$$L_{ia} = \frac{\partial E^{(2)}}{\partial \theta_{ia}} \tag{2.11}$$

The specific expressions for L and $\mathcal{E}^{\theta\theta}$ can be found in the previous work[116]. This equation for a given fragment, F, involves the orbitals that are *specific* to F for the SCF procedure. As this expression is derived from the coupled-perturbed SCF (CPSCF) equation which involves the HF electronic hessian, which is defined in the unmodified basis of the isolated fragments (ISO), we express L_{ia} in this same basis for a given fragment, F

$$L_{ia} = \frac{\partial E_F^{(2)}(\xi^{\mathrm{ISO}})}{\partial \theta_{ia}(\xi^{\mathrm{ISO}})} \tag{2.12}$$

2.2.4.1 Frozen Core Approximation

The frozen core approximation is often used in combination with post-SCF methods, to enable increased compute efficiency as well as basis sets that include only valence correlation. When applied to MP2 ALMO-EDA, the frozen energy requires special attention due to the Hylleraas functional. The frozen energy expression becomes

$$E_{\mathrm{FRZ}}^{(2)} = \frac{1}{2}\sum_{ijab} t_{ij}^{ab}(ij||ab) + \sum_{i''j''} P_{i''j''}F_{i''j''} + \sum_{ab} P_{ab}F_{ab} - 2\sum_{i''a} P_{i''a}F_{i''a} \tag{2.13}$$

Unprimed indexes are active, and the double primes indicate that the index runs over both frozen and active orbitals. The values of the MP2 response density matrix for when the indices are active orbitals do not change. When one or both indices are frozen, the values of the MP2 response density matrix are given by,

$$P_{i'j'} = 0 \tag{2.14a}$$

$$P_{ij'} = P_{j'i} = \frac{(L2)_{ij'}}{\epsilon_i - \epsilon_{j'}} \tag{2.14b}$$

where

$$(L2)_{ip''} = \frac{1}{2}\sum_{jbc} t_{ji}^{bc}(jp''||bc). \tag{2.15}$$

Here the singly prime indexes are core (inactive) orbitals. The ov block is again related to the Z-vector

$$P_{i''a} = -Z_{i''a} \tag{2.16}$$

which can be calculated using Eq. 2.10 with the MP2 orbital gradient L now being defined by Eq. 2.17

$$L_{i''a} = (L1)_{i''a} + (L2)_{i''a} + (L3)_{i''a} \tag{2.17}$$

where $L2$ was defined in Eq. 2.15 and the other two contributions are

$$(L1)_{ap''} = \frac{1}{2}\sum_{ijc} t_{ij}^{ca}(ij''||p''c). \tag{2.18}$$

$$(L3)_{p''q''} = \sum_{i''j''} P_{i''j''}(p''i''||q''j'') + \sum_{a''b''} P_{a''b''}(p''a''||q''b'') \tag{2.19}$$

These equations mirror those of MP2 gradient theory with frozen orbitals.[38]

2.2.5 Polarization energy and the ALMO Basis

Next, we move to the ALMO orbital basis used for the polarization and dispersion energy. The GV projected ALMO (pALMO) basis differs from the original modified MP2 pALMO basis (projected virtual space and symmetrically orthogonalized occupied space) by both having a global virtual basis and also using the polarized ALMO's from the second generation SCF ALMO EDA (SCF ALMO-EDA-II) which utilizes fragment electric field response functions (FERFs) to construct a basis for the ALMO-constrained SCF-MI[41]. This separates the virtual space into a fragment assignable virtual space made from dipolar, quadrupolar, octupolar responses (DQO-FERFs) and the orthogonal complement to the DQO-FERFs. During the SCF-MI, only DQ-FERFs participate by default, as with SCF ALMO-EDA-II. As with the pFRZ basis, the virtual orbitals are projected into the orthogonal complement of the occupied orbitals and the occupied orbitals are symmetrically orthogonalized. Unlike the frozen basis, we symmetrically orthogonalize the virtual space shell by shell, orthogonalizing the dipolar response virtual space first, then quadrupolar, and so on. We do this to preserve as much fragment locality possible while still keeping the virtual space orthogonal.

The polarization contribution to the MP2 interaction energy is defined as

$$\Delta E^{(2)}_{\mathrm{POL}} = \sum_{F}^{\mathrm{frag}} E^{(2)}_{\mathrm{POL}}(\xi^{\mathrm{ALMO}}_{\mathrm{p,GV}}) - \sum_{F}^{\mathrm{frag}} E^{(2)}_{\mathrm{FRZ}}(\xi^{\mathrm{ISO}}, \xi^{\mathrm{ISO}}_{\mathrm{p,GV}}, \xi^{\mathrm{FRZ}}_{\mathrm{p,GV}}) \tag{2.20}$$

As with $E^{(2)}_{\mathrm{FRZ}}$, $E^{(2)}_{\mathrm{POL}}$ is calculated by permitting only double excitations from occupied orbitals on the same fragment to the global virtual space. Unlike the frozen constraint, we no longer need to use t-amplitudes defined in a separate basis, and therefore we can obtain $E^{(2)}_{\mathrm{POL}}$ through the standard MP2 energy expression (Eq. 2.3) in the pALMO basis for each fragment.

2.2.6 Dispersion Energy

With MP2 ALMO-EDA-I, the charge conserving correlations (CCC) constraint is defined to separate out charge transfer-like correlations to define a CT-free correlation energy for MP2 ALMO-EDA-II. For MP2 ALMO-EDA-II, excitations included in the polarization energy (Figure 3.1, Type I/Type III) are included as well as simultaneous excitations to the fragment assignable virtual space (Figure 3.1, Type II). However, it is not straightforward to define the fragments' virtual space in the GV basis. Although the FERF virtuals are fragment assignable, they are not necessarily orthogonal between fragments. As such, the FERFs must be orthogonalized between fragments to maintain a completely orthogonal virtual space. As is shown later in the results section, including only type II excitations leads

to underestimating the dispersion energy. Therefore, in order to maintain a fully orthogonal virtual space, type IV excitations must also be included in the CCC term. The dispersion contribution to the interaction energy is defined as

$$\Delta E^{(2)}_{\mathrm{DISP}} = E^{(2)}_{\mathrm{CCC}}(\Xi^{\mathrm{CCC}}_{\mathrm{p,GV}}) - \sum_{F} E^{(2)}_{\mathrm{POL}}(\xi^{\mathrm{ALMO}}_{\mathrm{p,GV}}) \tag{2.21}$$

As CCC includes all excitation types, $E^{(2)}_{\mathrm{CCC}}$ is calculated using the standard MP2 expression (Eq. 2.3) in the pALMO basis for the whole system.

2.2.7 Charge Transfer

As we have defined the CCC constraint before, it is straightforward to define the contribution to the MP2 interaction energy from charge transfer as

$$\Delta E^{(2)}_{\mathrm{CT}} = E^{(2)}_{\mathrm{FULL}}(\Xi^{\mathrm{FULL}}) - E^{(2)}_{\mathrm{CCC}}(\Xi^{\mathrm{CCC}}_{\mathrm{p,GV}}) \tag{2.22}$$

In MP2 ALMO-EDA-I, the BSSE correction term is subtracted from $\Delta E^{(2)}_{\mathrm{CT}}$. However, as we have introduced the basis functions and virtual orbitals of other fragments in the frozen energy, we assign only the SCF component of the BSSE correction to $\Delta E^{\mathrm{SCF}}_{\mathrm{CT}}$ and assign the correlation component of the BSSE correction to $\Delta E^{(2)}_{\mathrm{FRZ}}$ and do not need to correct $\Delta E^{(2)}_{\mathrm{CT}}$.

2.2.8 MP2 ALMO-EDA-II EDA Summary

Similarly to MP2 ALMO-EDA-I, the required intermediate energies for MP2 ALMO-EDA-II are $E^{(2)}_{\mathrm{CCC}}$ as well as $E^{(2)}_{\mathrm{POL}}$ and $E^{(2)}_{\mathrm{FRZ}}$. Unlike MP2 ALMO-EDA-I, the correlation component of the BSSE correction is assigned to $\Delta E^{(2)}_{\mathrm{FRZ}}$ instead of $\Delta E^{(2)}_{\mathrm{CT}}$.

2.3 Implementation

We have implemented MP2 ALMO-EDA-II in a development version of the Q-Chem electronic structure program[21]. Both restricted and unrestricted orbitals are supported in the code.

2.3.1 Regularized MP2 EDA

Additionally, MP2 ALMO-EDA-II has the option to use kappa regularized MP2 (κMP2).[104] The implementation of κ-MP2 is fairly standard for the isolated, ALMO, and full system calculations as we can use the standard κMP2 equation[67]

$$E^{(2)}_{\kappa\mathrm{MP2}} = -\frac{1}{4}\sum_{ijab}\frac{|(ij\|ab)|^2}{\Delta^{ab}_{ij}}\left(1 - e^{-\kappa\left(\Delta^{ab}_{ij}\right)}\right)^2 \tag{2.23}$$

for the respective correlation energies. As the frozen term does not involve a basis that can be pseudo-canonicalized, we instead have to modify the two-electron integrals used to solve for the t amplitudes using Eq. 2.24

$$(i'j'\|a'b') = \sum_{ijab}(ij\|ab)\left(1 - e^{-\kappa\left(\Delta^{aba'b'}_{iji'j'}\right)}\right) \tag{2.24}$$

where Δ is an 8th rank tensor defined as

$$\Delta^{aba'b'}_{iji'j'} = -f_{ii'}\delta_{aa'}\delta_{jj'}\delta_{bb'} + \delta_{ii'}f_{aa'}\delta_{jj'}\delta_{bb'} - \delta_{ii'}\delta_{aa'}f_{jj'}\delta_{bb'} + \delta_{ii'}\delta_{aa'}\delta_{jj'}f_{bb'} \tag{2.25}$$

where $f_{rr'}$ refers to the MO Fock matrix[39].

2.3.2 Computational Demands

MP2 ALMO-EDA-II maintains the same $\mathcal{O}(\mathrm{n}^5)$ scaling with molecular size and $\mathcal{O}(\mathrm{n}^3)$ scaling with basis set size. However, as the GV basis requires basis functions from all fragments, even during fragment calculations (with the exception of isolated fragment energy calculations), the prefactor is significantly larger. This most significantly increases wall time during the CPSCF step in the frozen energy calculation and increases the memory requirements during the CCC energy calculation step. Additionally, as MP2 ALMO-EDA-II uses SCF ALMO-EDA-II, it inherits the large cost to calculate the FERFs at the SCF level. However, recent work has shown that the use of uncoupled FERFs (uFERFs) have reasonable accuracy with greatly reduced cost compared to regular FERF formation. The greatly reduced cost to compute uFERFs comes from the simplification in the CPSCF step used to calculate the FERFs[41]:

$$\mathcal{E}^{\Delta\Delta} \cdot \Delta^{\mu} = -2(M_{\mu}) \tag{2.26}$$

where $\mathcal{E}^{\Delta\Delta}$ is the HF orbital Hessian, Δ^{μ} is the orbital response to the field component $\mathcal{F}_{\mu}$, and M is the multipole moment matrix. The uFERF approximation uses the diagonal approximation

$$\mathcal{E}^{\Delta\Delta}_{aibj} \approx \frac{1}{\epsilon_a - \epsilon_i}\delta_{ab}\delta_{ij} \tag{2.27}$$

such that Δ^{μ} can be directly solved for.

2.4 Assessment

MP2 ALMO-EDA-II (henceforth sometimes shortened to MP2-EDA-II) is first assessed against similar systems as were used to validate aspects of MP2 ALMO-EDA-I (or, MP2-EDA-I, for short) paper. These are single fragment systems and systems that evaluate the asymptotic behavior of each term[116]. We then assess the extent to which the design goals of achieving useful basis set limits for all terms in MP2-EDA-II have been achieved. This is the aspect in which MP2-EDA-I was least satisfactory. This assessment is performed across the Ionic 43 (I43) [65] and S22 data sets[55].

2.4.1 Single Fragment Cases

For single-fragment cases, MP2-EDA-II is identical to MP2-EDA-I and therefore the results for single-fragment systems will be identical to the ones in the originally published paper[116]. To ensure the correct separation of $\Delta E_{\mathrm{FRZ}}^{\mathrm{MP2}}$ and $\Delta E_{\mathrm{POL}}^{\mathrm{MP2}}$ while not being trivially identical to MP2 ALMO-EDA-I, we investigate a single fragment in the presence of an external electric field with the addition of a weakly interacting fragment far away. This should ensure that MP2 ALMO-EDA-II frozen term matches the correct behavior of MP2 ALMO-EDA-I.

To illustrate this, we investigate a hydrogen fluoride molecule with a helium atom 20 Åaway in a uniform external electric field along the axis of the hydrogen fluoride. As with the MP2 ALMO-EDA-I, the slope corresponding with how the MP2 frozen correlation energy varies with respect to electric field strength should match the MP2 correction to the dipole moment of the system. As seen in Figure 2.2, the frozen energy for the MP2 ALMO-EDA-II closely matches MP2 ALMO-EDA-I frozen energy, which both match the expected dipole moment correction of 0.0468 au Thus, like MP2 ALMO-EDA-I, it is crucial to include the relaxed density in the frozen term.

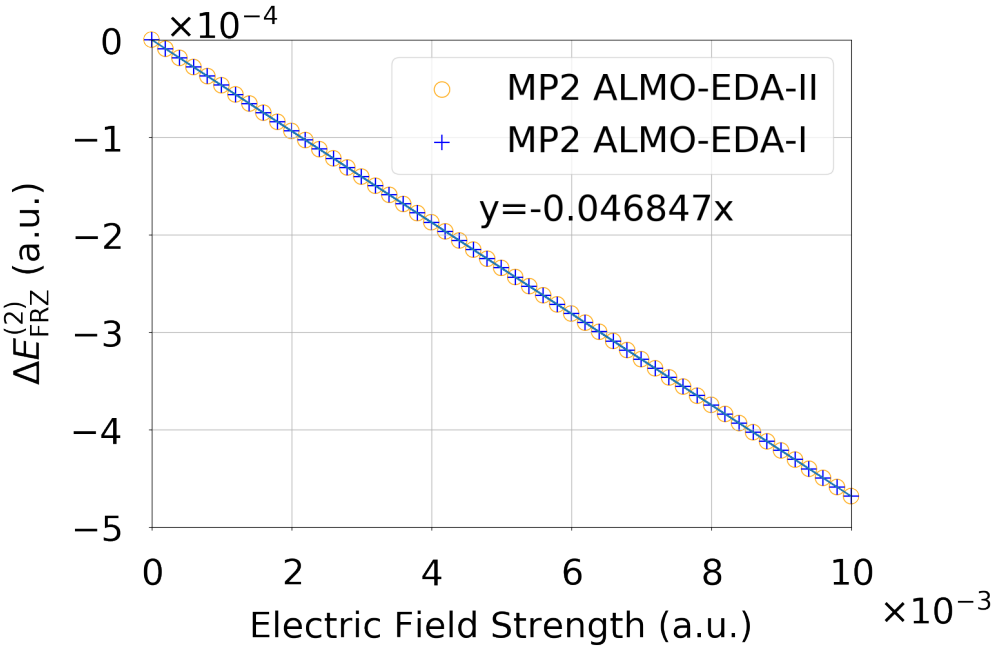


Figure 2.2: $\Delta E_{\mathrm{FRZ}}^{(2)}$ of hydrogen fluoride with a helium atom 20 Å away with an electric field along the hydrogen fluoride bond axis at varying strengths.

2.4.2 Asymptotics

As the overlap between fragment virtual spaces decreases as fragments separate, MP2-EDA-II becomes identical to MP2-EDA-I in the long range. We will therefore concentrate on

focusing on the dispersion contribtution.

For the dispersion term, there are multiple choices for defining the fragment assignable virtual space. A natural choice for a fragment assignable virtual space is the FERFs which are fragment assignable. However, as FERFs are not orthogonal between fragments, they must be orthogonalized to maintain an orthogonal virtual space. This introduces an energy penalty as each fragments' FERF space encompasses less Hilbert space after the orthogonalization. A consequence of this energy penalty is the overestimation of charge transfer as the CCC term underestimates dispersion with the orthogonalization. Figure 2.3 shows that this energy penalty is large enough that it predicts a significant amount of CT for the methane dimer near equilibrium, even when using DQO-FERFs. Therefore, in order to fully capture dispersion while keeping an orthogonal virtual space, the entire GV space must be used.

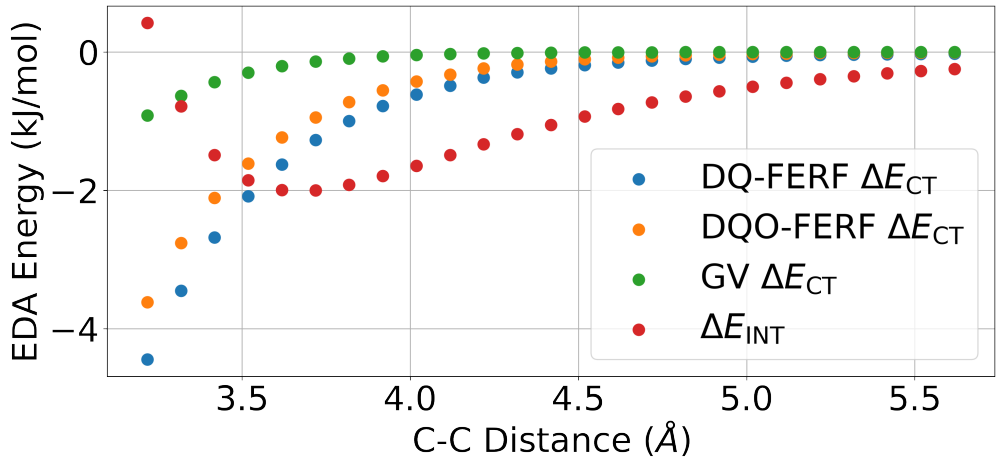


Figure 2.3: Comparison of different fragment virtual spaces for asymptotic behavior of ΔE_{CT} of methane dimer at aug-cc-pVQZ.

2.4.3 Strong interactions: The I43 dataset

As MP2 ALMO-EDA-I builds upon the first-generation SCF-level ALMO EDA (henceforth sometimes shortened to SCF-EDA-I), it inherits the known issue of ill-defined charge transfer from SCF-EDA-I[4]. MP2-EDA-II, as described in the previous sections, builds upon SCF-EDA-II which addresses the ill-defined charge transfer issue by utilizing FERFs to construct the polarization subspace[41]. To evaluate the basis set convergence for EDA, we will evaluate the difference between aug-cc-pVTZ and aug-cc-pVQZ basis sets for each EDA term. Because MP2 is known to converge to the basis set limit more slowly than HF, we should evaluate the convergence of each term alongside the convergence of the MP2 interaction energy, as we would like each EDA term to converge similarly to the overall MP2 interaction energy.

The first systems we investigate are those in the I43 data set[63] which contains 3 subsets. The first includes 21 hydrogen-bonded anion-neutral dimers (AHB21), another that contains

3 cation-water and 3 cation-benzene complexes (CHB6), and the last contains 16 anion-cation systems commonly encountered in ionic liquids (IL16). The systems contained in this data set have relatively large binding energies compared to interactions between uncharged molecules, and as we shall see, MP2-EDA-I has significant problems, which makes this a good starting point for our analysis. We have omitted 3 proton-bridged systems from AHB21, as they do not have well-defined fragments for EDA. The graphs in Figure 2.4 show the maximum difference of each term calculated between the aug-cc-pVTZ and aug-cc-pVQZ basis sets for AHB21 and IL16, and def2-TZVPPD and def2-QZVPPD for CHB6.

As the basis set difference for each EDA term sums up to the basis set difference for the interaction energy, it is useful to focus on the largest positive basis set difference term for each system. It is clear that MP2-EDA-II exhibits significant improvement in basis convergence compared to MP2-EDA-I. Specifically, MP2 ALMO-EDA-I fails to show stable results for the polarization and frozen terms of IL16 and AHB21. For instance, $\Delta\Delta E_{\mathrm{POL}}$ for MP2-EDA-I changes by almost 20 kJ/mol for two systems while the interaction energy changes by roughly 4 kJ/mol. The polarization term can partially be attributed to the ill-defined polarization term in SCF-EDA-I, however the frozen term is the same at the SCF level and the MP2 EDAs only differ for the correlation part.

One reason for this poor behavior of MP2-EDA-I is due to the strong orthogonality requirement for MP2 in the frozen energy. For systems with diffuse electron density for a fragment, such as the anionic fragments in IL16 and AHB21, there is a larger overlap of the anionic fragment occupied orbitals with the virtual orbitals of the other fragment. Thus, enforcing the strong orthogonality constraint for the pFRZ basis while using the ISO basis t amplitudes in MP2-EDA-I leads to poor basis set stability. However, as MP2-EDA-II uses a pISO basis for the t amplitudes in the frozen energy, MP2 ALMO-EDA-II's frozen and polarization term has a much improved basis set stability.

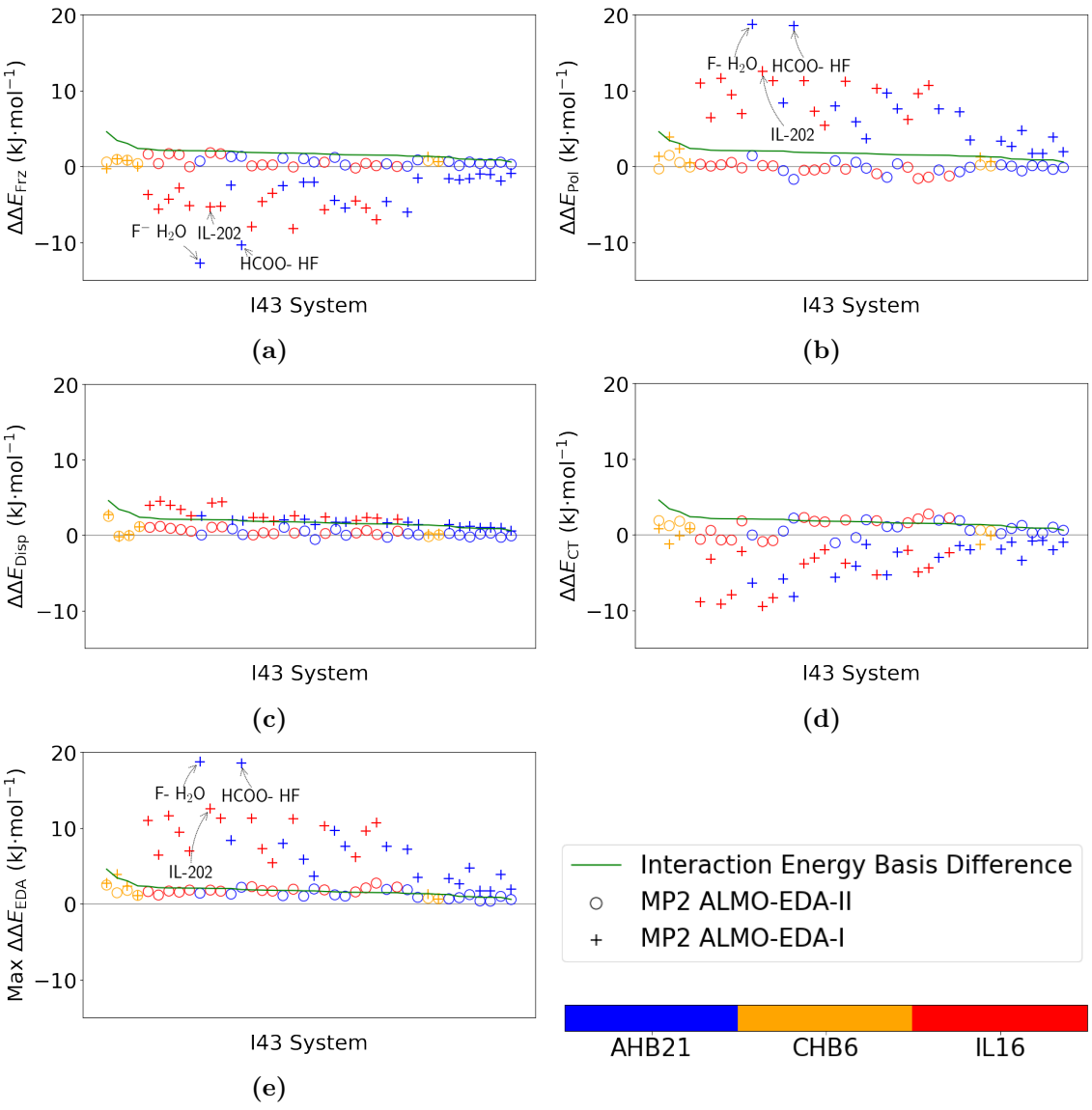


Figure 2.4: Basis set differences of frozen (a), polarization (b), dispersion (c), and charge transfer (d) energy for MP2 ALMO-EDA-II and MP2 ALMO-EDA-I. The maximum difference of each term (e) is plotted as well. Basis set difference was calculated by subtracting the EDA term energy at def2-TZVPPD from def2-QZVPPD for CHB6 and aug-cc-pVTZ from aug-cc-pVQZ for IL16 and AHB21. Shared proton systems omitted. Basis set difference of interaction energy plotted for reference.

Figure 2.5 shows basis set convergence for each term of the $F^- \cdots HOH$ complex from AHB21, one of the two worst-behaved examples for MP2-EDA-I in Figure 2.4. It is evident from Figure 2.5 that the MP2-EDA-I frozen interaction energy lacks a useful basis set limit. By contrast, ΔE_{FRZ} for MP2-EDA-II exhibits excellent basis set stability, and is nearly con-

verged at aug-cc-pVTZ. As the two MP2 EDA's share the identical SCF frozen energy, this difference illustrates the dramatic improvement of MP2-EDA-II over MP2-EDA-I's frozen term through the introduction of the GV basis.

The polarization interaction energy, ΔE_{POL} differs between MP2-EDA-I and MP2-EDA-II at both the SCF and the correlated level. ΔE_{POL} of SCF-EDA-I behaves poorly by becoming more attractive (by capturing some CT) as the basis set is improved towards completeness[4, 41]. Additionally, the correlation component of ΔE_{POL} for MP2-EDA-I also does not have a well-defined basis set limit as shown in Figure 2.6b. In comparison, the polarization energy for the MP2-EDA-II seems to be sufficiently converged at aug-cc-pVTZ. Thus, both the addition of SCF-EDA-II and the GV basis to the correlation components of MP2-EDA-II greatly improve the basis set stability for both ΔE_{FRZ} and ΔE_{POL} relative to MP2-EDA-I.

Looking at the charge transfer term, ΔE_{CT}, in Figure 2.5d, there is also a significant difference between MP2-EDA-I amnd MP2-EDA-II. For MP2-EDA-I, ΔE_{CT} trends towards zero as the basis set in increased, a trivial basis set limit, as can be anticipated based on previous studies[41]. By contrast, MP2-EDA-II has a stable, non-trivial basis set limit for the charge transfer term.

In terms of electron correlation, the poor behavior of MP2-EDA-I for ΔE_{CT} is arising from its treatment of the dispersion term. Figure 2.5c illustrates that ΔE_{Disp} in MP2-EDA-I also lacks a useful basis set limit, as corroborated by the fact that it is not fully converged even at aug-cc-pV6Z. This is a result of the dependence on atomic-tagging used in MP2-EDA-I, where functions on one fragment can nearly represent functions on the other fragment as the basis set limit is approached.

On the other hand, MP2-EDA-II dispersion seems to be converged as early as aug-cc-pVTZ. This is a direct result of using the FERF-DQO space to describe dispersive correlations in MP2-EDA-II; another of the significant design changes introduced here. This is fully consistent with the known fact that correlation effects associated with long-range dispersion converge much more rapidly with basis set size than short-range correlations.[32, 71]

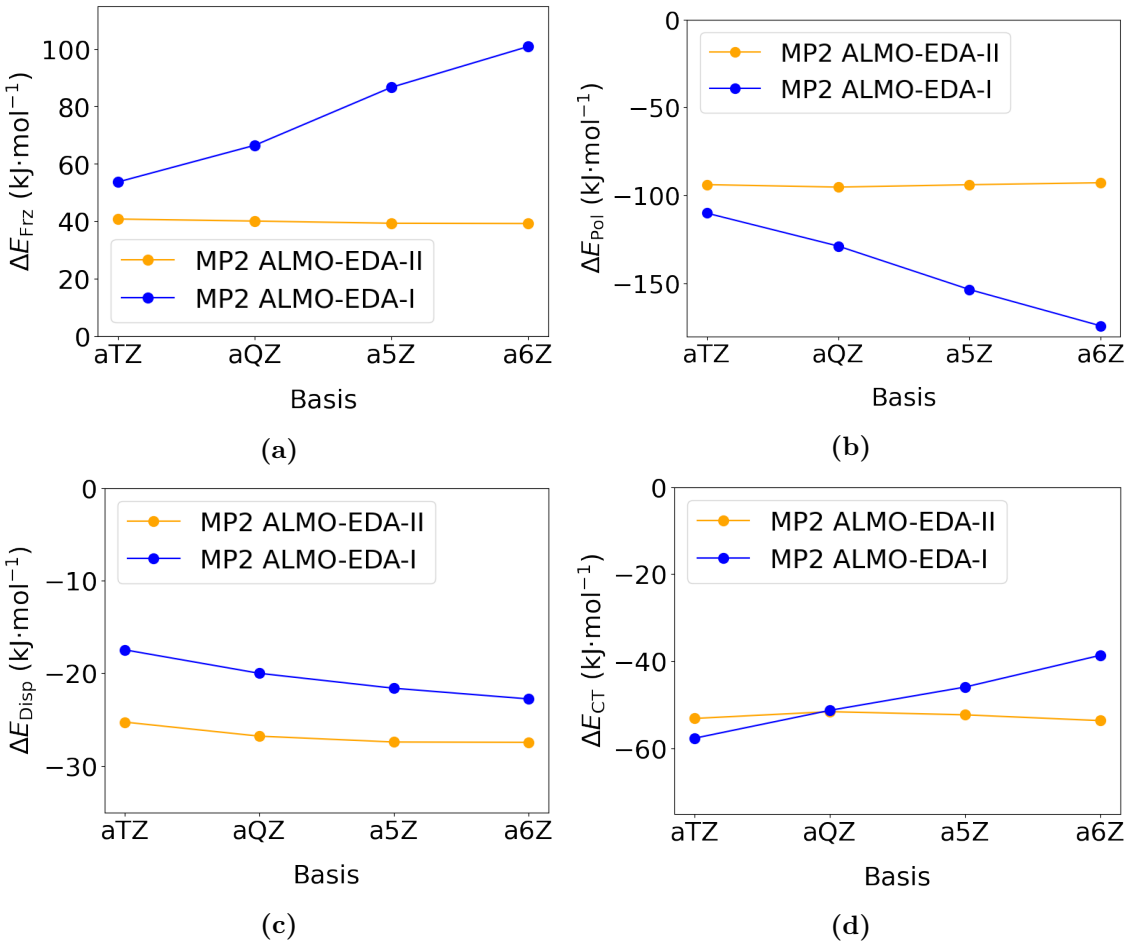


Figure 2.5: Basis set convergence of the frozen (a), polarization, (c) dispersion, and charge transfer (d) energies of the AHB21 fluoride water system for both MP2 EDA's.

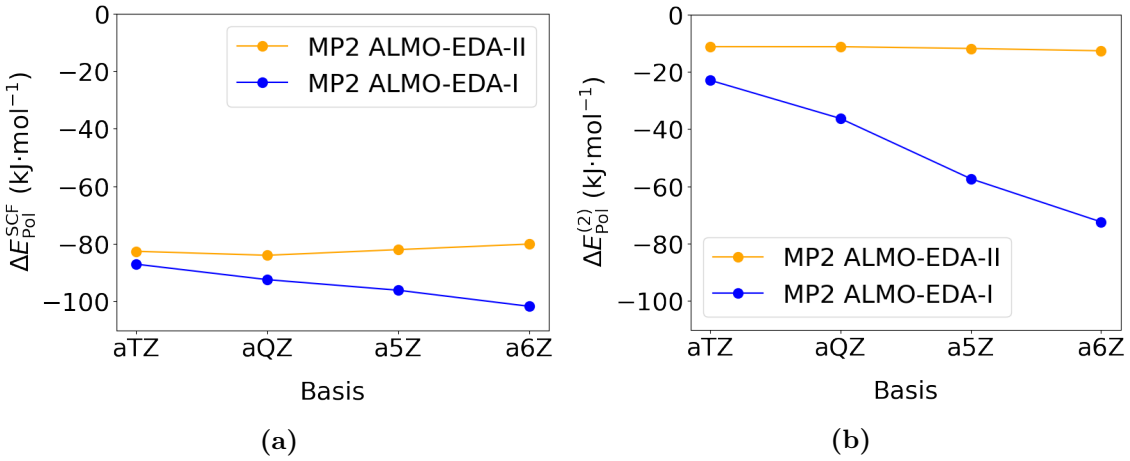


Figure 2.6: Basis set convergence of the (a) SCF component ($\Delta E_{\mathrm{POL}}^{\mathrm{SCF}}$) and (b) correlation component ($\Delta E_{\mathrm{POL}}^{(2)}$) of the polarization energy for the AHB21 fluoride water system for both MP2 EDA's.

2.4.4 Weaker interactions: The S22 data set

We next move to S22, a popular data set containing 22 hydrogen bonded and dispersion bonded systems[55]. As seen in Figure 2.7, the S22 results are similar to the results from I43, although the difference between MP2 EDA's are much smaller for most systems. One explanation may be that the interactions for the systems in S22 are in general much weaker than the systems in I43.

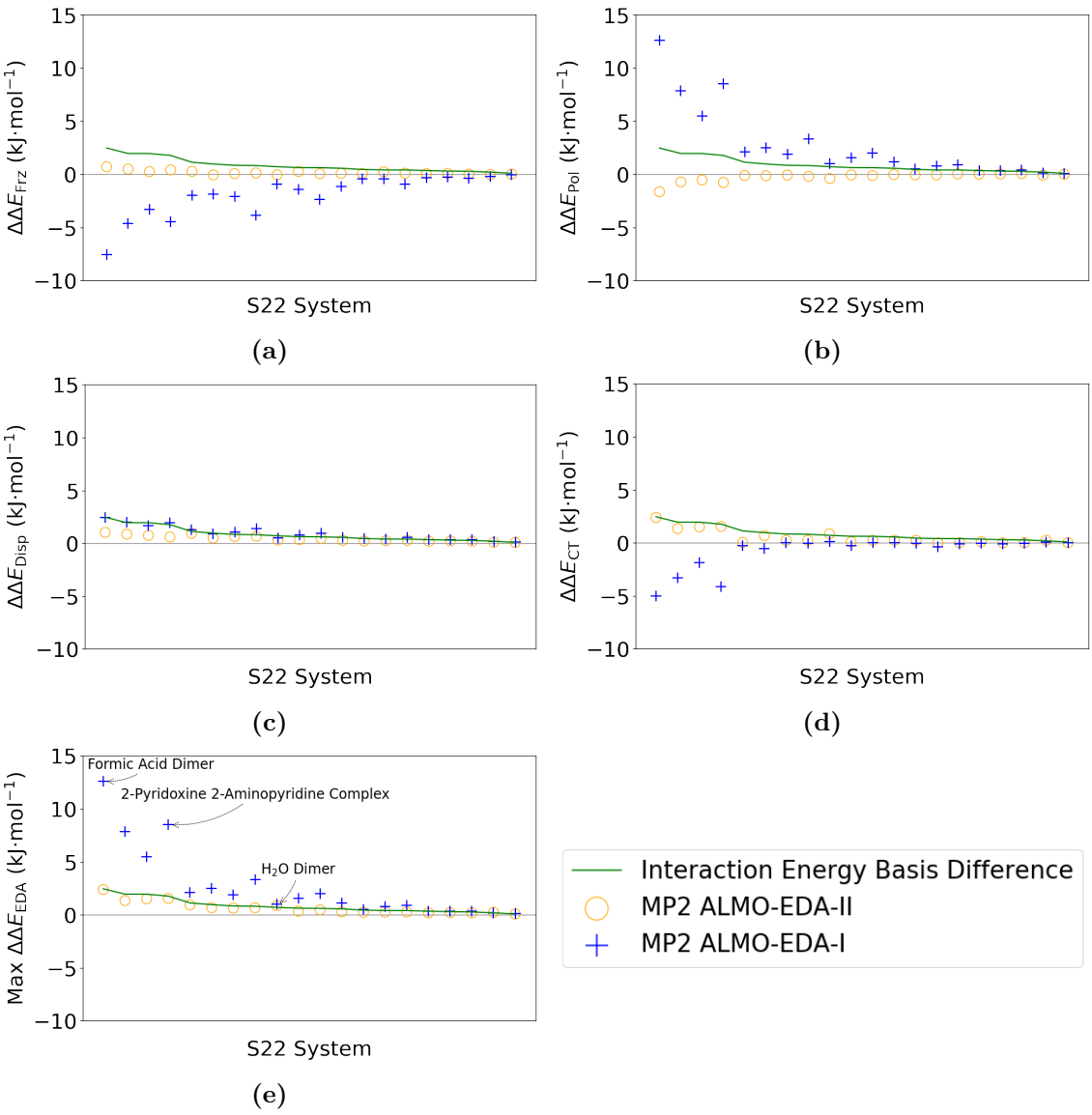


Figure 2.7: Basis set difference of (a) frozen, (b) polarization, (c) dispersion, and (d) charge transfer energies for MP2 ALMO-EDA-II and published MP2 EDA. Basis set difference was calculated by subtracting the EDA term energy at aug-cc-pVTZ from aug-cc-pVQZ. Adenine Thymine omitted due to size. Basis set difference of interaction energy plotted for reference.

It is important to point out that MP2 ALMO-EDA-II performs similarly to MP2 ALMO-EDA-I for the water dimer case when looking at the basis set difference between aug-cc-pVTZ and aug-cc-pVQZ where both have a slightly larger max EDA term basis set difference compared to the interaction energy basis set difference. The largest basis difference term for the water dimer comes from the CT term for MP2 ALMO-EDA-II and from the POL term for MP2 ALMO-EDA-I. To investigate this behavior, the results are expanded to larger

basis sets, shown in Figure 2.8. The CT and POL terms both seem to be relatively stable for MP2 ALMO-EDA-II at higher basis sets, while MP2 ALMO-EDA-I does not seem to converge to a stable basis set limit even at aug-cc-pV6Z. It is also worth noting that the improvement in basis set limit convergence is not only due to the improvements at the SCF level due to SCF ALMO-EDA-II. Similarly to the fluoride water case, Figure 2.9 shows that both the correlation and SCF components of CT and Pol for the MP2 ALMO-EDA-II has a more well defined basis set limit compared to MP2 ALMO-EDA-I's.

Although the systems in S22 are overall generally more well-behaved compared to I43 for MP2 ALMO-EDA-I, we still see a significant improvement to the basis set convergence for many of the S22 systems using the MP2 ALMO-EDA-II at the aug-cc-pVTZ to aug-cc-pVQZ basis sets. In addition, when looking at larger basis sets, the MP2 ALMO-EDA-II have a stable basis set limit unlike MP2 ALMO-EDA-I.

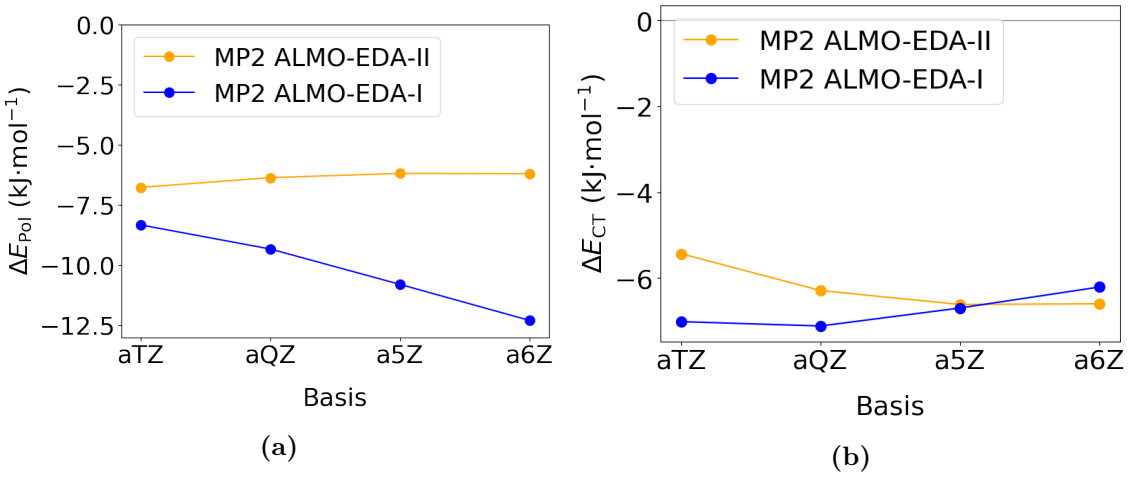


Figure 2.8: Basis set convergence of the (a) polarization and (b) charge transfer energies of the S22 water dimer system for both MP2 EDA's. The first panel, (a), shows that the MP2 ALMO-EDA-II polarization energies converge to a well-defined basis set limit while MP2 ALMO-EDA-I polarization continues to increase in strength even at large basis sets. Similarly, the second panel, (b), shows that the charge transfer energies have a well-defined basis set limit for the MP2 ALMO-EDA-II at higher basis sets while MP2 ALMO-EDA-I shows that at higher basis sets, the charge transfer term does not have a non-trivial basis set limit like SCF ALMO-EDA-I. [41]

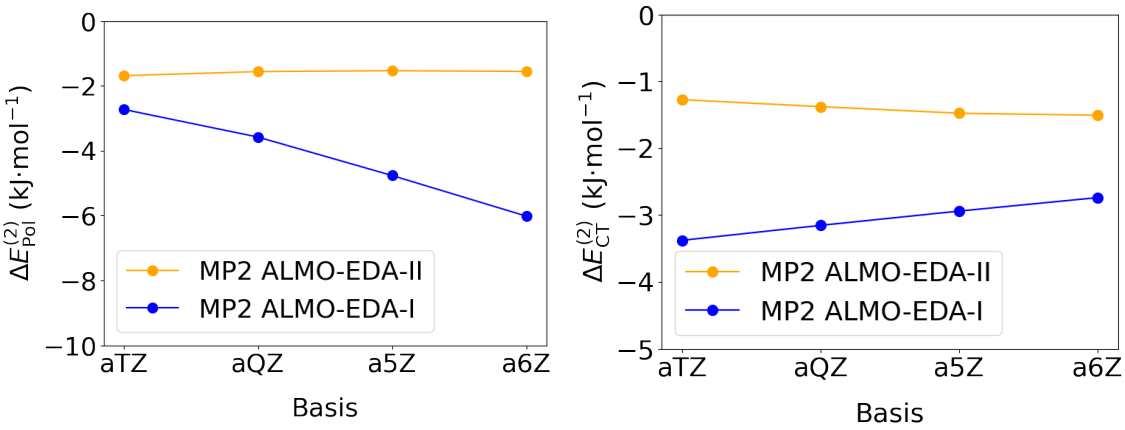


Figure 2.9: Basis set convergence of the correlation component of (a) polarization and (b) charge transfer energies of the S22 water dimer system for both MP2 EDA's. Both graphs show that the lack of a non-trivial basis set limit for MP2 ALMO-EDA-I exists not just at the SCF level, but the correlation level as well.

2.4.5 Extent of compatibility with DFT-EDA-II

The DFT-based SCF-EDA-II is well-established[43, 77, 75] and becoming quite widely used. It is therefore useful to examine the behavior of the terms identified in MP2-EDA-II relative to those separated in DFT-EDA-II for some representative cases where the total interaction energy is similar when evaluated by both methods. Here we fix the AO basis set at aug-cc-pVQZ and examine the distance dependence of each term via MP2-EDA-II, and by ωB97M-V/aug-cc-pVQZ[78, 53] for three systems, the water dimer, the complex between fluoride anion and water, and the methane dimer. The results are shown in Figure 2.10. As expected, the long range behavior of each term (e.g. for separations $R > 3.5$Å for water dimer and fluoride-water) is almost identical. Near equilibrium however, the behavior of each EDA differs slightly and will be discussed separately below.

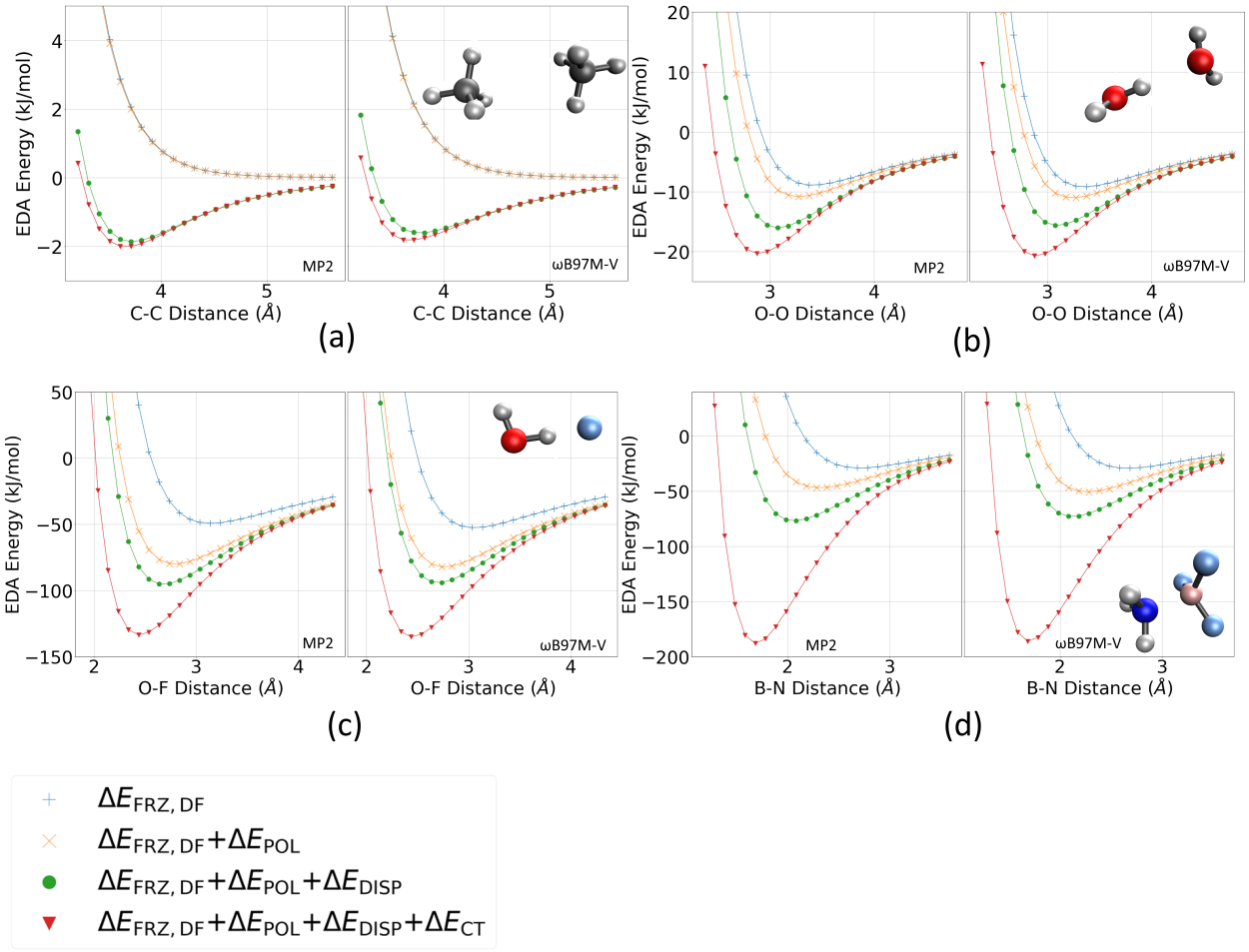


Figure 2.10: Distance-dependence of the (cumulative) EDA contributions for (left) MP2 ALMO-EDA-II vs. (right) ωB97M-V ALMO-EDA-II for (a) methane dimer, (b) water dimer, (c) fluoride water, (d) BF_3 NH_3 calculated at aug-cc-pVQZ level of theory.

For the methane dimer, the binding largely reflects competition between Pauli repulsion and dispersion. Good qualitative agreement is evident between MP2-EDA-II and DFT-EDA-II using ωB97M-V.

The water dimer interaction is far stronger than the methane dimer, and has driving forces for binding from permanent and induced electrostatics and CT.[56, 72, 76] For the water dimer at equilibrium (2.87 Å), MP2 EDA predicts the (dispersion-free) frozen term to be slightly repulsive while the ωB97M-V EDA calculates the dispersion-free frozen term to be very slightly attractive. The corresponding equilibrium geometries, which reflect the tradeoff between Pauli repulsion and attractive permanent electrostatics, are encouragingly similar. The contribution due to polarization is also quite similar, as is the final interaction energy.

The fluoride-water interaction is roughly 6 times stronger than the water dimer interaction, and is dominated by permanent and induced electrostatics and charge-transfer contributions. There is pleasing qualitative consistency between MP2-EDA-II and DFT-EDA-II for this system. We conclude that the partition of these 3 dominant contributions is quite compatible between the two approaches for this system.

BF_3 NH_3 is a strongly interacting system, comparable to some covalently bonded systems, dominated by charge transfer. The contribution to electrostatics and Pauli repulsion as well as polarization are similar between MP2-EDA-II and DFT-EDA-II. The slightly larger contribution of the dispersion term in MP2-EDA-II relative to DFT-EDA-II is visually evident in Figure 2.10(d). We reiterate that the EDA "dispersion" term reflects that contribution from electron correlation which *becomes* pure dispersion in the non-overlapping limit. It may better be referred to as "dispersion-like" correlation in the shorter range, and we should not necessarily expect very close agreement between two quite independent approaches. Even so, dispersion is the least significant contributor to binding, and therefore both EDAs report the same physical origins of strong binding.

2.5 Conclusions

We have developed a second generation energy decomposition analysis (EDA) based on absolutely localized molecular orbitals (ALMOs) for calculations of intermolecular interactions based on MP2 theory. Like its predecessor,[116, 117, 69] MP2 ALMO-EDA-I, the new MP2 ALMO-EDA-II yields correlation corrections to each of the three contributions to an interaction energy at the mean-field level: the frozen interaction, the polarization interaction, and the charge-transfer (or electron delocalization) interaction. In addition, one extra term describing dispersive effects (the correlation contributions that become dispersion in the non-overlapping regime) is separated out. The approach of correcting each mean-field EDA contribution is essential for a viable correlated EDA: simply adding the correlation binding energy as a distinct term[112] is known to not be physically correct.[118] We also view separation of polarization[41] and charge-transfer[76] effects as essential in a useful EDA.

MP2 ALMO-EDA-II improves upon its predecessor, MP2 ALMO-EDA-I, by design choices that achieve the following properties:

- Greatly improved basis set convergence. Each EDA term has a stable basis set limit for both its SCF contribution (as a result of using SCF ALMO-EDA-II), and its correlation correction (as a consequence of design improvements introduced here). MP2-ALMO-EDA-I does not have satisfactory basis set convergence, by comparison.

- No reliance upon atom orbitals as the choice of representation. Every term in MP2 ALMO-EDA-II can be evaluated without assuming that the single particle space is represented via atomic orbitals. This permits implementations that use different basis representations. By contrast, MP2 ALMO-EDA-I relied upon an AO representation.

- No use of non-orthogonal orbitals to represent either the occupied or virtual space in any intermediate energy. This design choice should be useful when extending MP2 ALMO-EDA-II to other post-SCF methods such as OOMP2 EDA or a coupled cluster EDA. By contrast MP2 ALMO-EDA-I relied upon nonorthogonal virtual orbitals, which introduces greater algebraic complexity.

Alongside these improvements, MP2 ALMO-EDA-II maintains the correct asymptotic behavior of MP2 ALMO-EDA-I. Additionally, MP2-ALMO-EDA-II includes the option to use κ regularized MP2 which has shown improvement over the standard MP2 in calculating a wide variety of noncovalent interactions[70, 104]. Additionally, we hope that this approach will be extended to the development of other recently developed methods based on MP2 such as a regularized OOMP2[67, 66, 95] EDA or an EDA based on double hybrid functionals[79]. It also seems possible to extend MP2 ALMO-EDA-II to coupled cluster methods, directly at the singles + doubles (CCSD) level, and with further exploration, perhaps the CCSD(T) level.

Chapter 3

A second generation energy decomposition analysis of intermolecular interaction energies from second order Møller-Plesset theory: Chemical examples and applications.

3.1 Introduction

As numerical experiments, quantum chemical calculations of intermolecular interaction energies are highly successful[96]. Wavefunction-based methods based on high-level coupled cluster theory[7] achieve an accuracy that is often significantly better than 1 kJ/mol, provided that an adequately large basis set[49] is employed to approach the complete basis set (CBS) limit. Thus methods such as CCSD(T)/CBS have become de facto benchmarks for intermolecular interaction energies.[96] More approximate methods including suitable modifications[103, 11] of MP3, and even adaptations[34, 54, 33, 19, 31, 30, 104] of MP2 are also very useful for such calculations at reduced computational cost. MP2 itself also performs very well for some intermolecular interactions such as hydrogen-bonding[23, 127], ion-water complexes and ionic liquid interactions[128, 46], and remains widely used, despite well-recognized deficiencies for cases such as collective dispersion interactions.[90]

To give chemical insight beyond simply a numerical experiment, it is desirable to disentangle the contribution of individual physical driving forces, both stabilizing and destabilizing, that collectively account for a total intermolecular interaction energy. This is the objective of energy decomposition analysis (EDA).[75] A successful EDA for noncovalent interactions should identify the energetic contributions from as many as possible of the following well-

recognized terms: (i) permanent electrostatics, (ii) Pauli repulsion, (iii) dispersion, (iv) polarization, (v) charge-transfer or dative interactions. While all these terms have textbook definitions[110] (or are zero) in the limit of non-overlapping fragments, it is crucial to connect them smoothly to the chemically relevant overlapping regime. Towards that goal, the objective of this paper is to report on the development of a new EDA framework for wavefunction-based methods, and its concrete realization as a method for MP2, and the recently proposed improvement; regularized MP2.[104]

There are two general approaches to EDA. This work follows the supermolecule approach, where the interaction energy of a collection of fragments (typically molecules) is defined as the difference between the supermolecule and its individual fragments:

$$\Delta E = E - \sum_F E_F \tag{3.1}$$

ΔE should include the counterpoise correction[13] for basis set superposition error (BSSE) if atom-centered basis sets are used that are not close to the complete basis set limit. A well-posed EDA (suitable criteria have been specified[75]) begins from the energy of isolated fragments, and employs a sequence of intermediate states with progressively weaker constraints that isolate the different physical contributions to ΔE until the unconstrained energy is obtained. An advantage is that any model energy can be employed (in principle). A disadvantage is that the design of the constraints is critical to obtaining physically valid results, and these may have to be method-dependent (as well shall discuss later). Furthermore interactions such as dispersion, polarization, and permanent electrostatics that have textbook definitions in the long-range non-overlapping limit are necessarily ambiguous in the chemically relevant overlapping regime.

Differences between energies of intermediate states reveal the supermolecular EDA contributions to ΔE, such as:

$$\Delta E_{\mathrm{INT}} = \Delta E_{\mathrm{GD}} + \Delta E_{\mathrm{FRZ}} + \Delta E_{\mathrm{POL}} + \Delta E_{\mathrm{DISP}} + \Delta E_{\mathrm{CT}} \tag{3.2}$$

Here ΔE_{GD} is the "geometric distortion" energy needed to deform the optimal fragment geometries (used to evaluate E_F) to their geometries in the complex. The first physical contribution, ΔE_{FRZ} is associated with electronic degrees of freedom such as orbitals frozen at their fragment geometries, and physically includes permanent electrostatics and Pauli repulsions (since it omits dispersion in Eq. 3.2, it also sometimes referred to being "dispersion-free"; FRZ-DF). The second physical contribution, ΔE_{POL} is the energy lowering due to on-fragment relaxation (or polarization) of the electronic degrees of freedom, whilst prohibiting charge flow. This is followed by the energy lowering associated with dispersion, ΔE_{DISP} and finally the energy lowering associated with charge transfer or dative interactions, between fragments, ΔE_{CT}.

The principal alternative to the supermolecule approach is to directly correct the fragment energies for the interactions using valid quantum mechanical approaches such as perturbation

theory. Symmetry-adapted perturbation theory (SAPT)[50, 115, 92] is the paradigmatic example of this approach to compute the interaction energy and get its components simultaneously. With great improvements in accuracy vs cost with the development of SAPT-DFT,[48] perhaps the main remaining limitation is that some physical concepts do not emerge naturally in SAPT (e.g. polarization and charge transfer are naturally combined as "induction", but are not (yet) naturally separated[109, 81, 76]).

Space limits preclude detailed discussion of the considerable effort that has been devoted to DFT-based EDA development for intermolecular interactions, apart from a brief historical perspective and summary of present status. Modern EDA development began with the seminal Kitaura-Morokuma method,[60, 86] which has inspired many successors[130, 129, 12, 84, 85, 59, 57, 83, 82, 112, 44, 113]. Another noteworthy approach to understanding inter (and intra) molecular interactions is the Natural Bond Orbital (NBO) framework.[124] While NBO is very widely used to aid in analysis of quantum chemistry calculations, it is important to note that it substantially overestimates charge-transfer effects in an unphysical way,[56, 111] Therefore the CT energies from NBO should at most be used as descriptors. Separate from EDA, many other useful descriptors of non-covalent interactions have been developed, including the electron localization function (ELF),[99] the σ and π hole models,[16] the non-covalent interactions (NCI) index,[52], etc.

At the self-consistent field (SCF) level of density functional theory (DFT) or mean-field Hartree-Fock (HF) theory, one well-developed and (apparently) satisfactory supermolecular EDA is the second generation absolutely localized MO (ALMO)-EDA[43, 77, 75]. The SCF ALMO-EDA-II (henceforth denoted as SCF-EDA-II) provides useful complete basis set limits for polarization and charge-transfer,[41], and is not tied to any specific basis representation. While development continues,[74, 122] the second generation ALMO-EDA is a useful approach for analyzing intermolecular interaction energy calculations at the DFT level. It is therefore natural to ask whether or not this can be adequately generalized from the SCF level to treat post-SCF methods such as MP2 in the first instance, and then perhaps double hybrid DFT and higher order MP and CC methods. That will be the purpose of the research described here.

There are already a number of notable efforts to develop supermolecular post-SCF EDAs. In the pair interaction energy decomposition analysis (PIEDA)[24] and the localized MO (LMO)-EDA framework,[112] the entire correlation binding energy is assigned as a dispersion contribution to binding (i.e. $\Delta E_{\mathrm{DISP}}^{\mathrm{LMO\text{-}EDA}} = E^{(\mathrm{corr})} - \sum_F E_F^{(\mathrm{corr})}$). This assumption is not formally correct;[118] in principle *all* physical contributions to an intermolecular interaction are affected by electron correlations. Other post-SCF supermolecular EDAs[107, 102, 5, 47, 101, 2, 28] therefore make efforts to include such contributions. Most relevant to this work is the development of an EDA at the MP2 level which generalizes the first-generation ALMO-EDA by providing correlation corrections to each term.[116, 117, 69] Unfortunately, this MP2-EDA-I does not provide useful basis set limits for each term. While it can be useful for chemical applications if applied carefully, an improved approach is clearly desirable.

Very recently, new theory and an associated implementation have been reported to de-

fine a second generation MP2-based EDA, MP2 ALMO-EDA-II (henceforth referred to as MP2-EDA-II). This theory, which is reviewed in Section 3.2 was designed with the goal of achieving useful basis set limits for each contribution in Eq. 3.2. The purpose of this work is to apply MP2-EDA-II to four classes of complexes with varying strengths of intermolecular interactions to explore its usefulness for chemistry, as well as helping to resolve the physical origin of those interactions. First, we will investigate a water-xylose system involving hydrogen bonding, one of the most familiar intermolecular interactions. We next will look at a class of systems that participate in unconventional hydrogen bonding involving metal-hydrides as electron acceptors. We will also investigate a group of complexes involving tetrel bonding, which have recently garnered a lot of attention in biology, chemistry, and catalysis [126, 10, 9, 27]. Finally, we will investigate a "solvent-resistant halogen bond"[97] in which a halogen bonded complex showed remarkable stability with respect to solvent type.

3.2 Theory

The MP2 ALMO-EDA-II decomposes both the SCF interaction energy, $\Delta E_{\mathrm{INT}}^{\mathrm{SCF}}$, and the correlation component of the MP2 interaction energy, $\Delta E_{\mathrm{INT}}^{(2)}$, into separate components. The SCF interaction energy is separated into a frozen interaction, $\Delta E_{\mathrm{FRZ}}^{\mathrm{SCF}}$, which includes electrostatic effects and Pauli repulsion, a polarization interaction, $\Delta E_{\mathrm{POL}}^{\mathrm{SCF}}$, and a charge transfer interaction $\Delta E_{\mathrm{CT}}^{\mathrm{SCF}}$. The SCF component of MP2 ALMO-EDA-II is identical to SCF ALMO-EDA-II and therefore we will focus on the decomposition of the correlation component of the interaction energy. Like $\Delta E_{\mathrm{INT}}^{\mathrm{SCF}}$, $\Delta E_{\mathrm{INT}}^{(2)}$ contains a frozen, polarization, and charge transfer component, however, there is an additional dispersion term, $\Delta E_{\mathrm{DISP}}^{(2)}$ as dispersion is a purely correlational effect. MP2 ALMO-EDA-II treats each decomposed term of $\Delta E_{\mathrm{INT}}^{(2)}$ as a correction to the equivalent SCF ALMO-EDA-II term (including $\Delta E_{\mathrm{DISP}}^{(2)}$ as a correction to the lack of a dispersion term in SCF ALMO-EDA-II).

SCF ALMO-EDA-II computes intermediate energies of the complex with different constraints on the electron density, fragment isolated densities, $\mathbf{P}_{\mathrm{F}}$ a frozen density, $\mathbf{P}_{\mathrm{FRZ}}$, a relaxed/polarized density, $\mathbf{P}_{\mathrm{ALMO}}$, and the fully relaxed system density, $\mathbf{P}_{\mathrm{SYS}}$. Each SCF ALMO-EDA-II term is computed as the difference between these intermediate energies,

$$\Delta E_{\mathrm{FRZ}}^{\mathrm{SCF}} = E^{\mathrm{SCF}}(\mathbf{P}_{\mathrm{FRZ}}) - \sum_{\mathrm{F}}^{\mathrm{frag}} E_{\mathrm{F}}^{\mathrm{SCF}}(\mathbf{P}_{\mathrm{F}}) \tag{3.3a}$$

$$\Delta E_{\mathrm{POL}}^{\mathrm{SCF}} = E^{\mathrm{SCF}}(\mathbf{P}_{\mathrm{ALMO}}) - E^{\mathrm{SCF}}(\mathbf{P}_{\mathrm{FRZ}}) \tag{3.3b}$$

$$\Delta E_{\mathrm{CT}}^{\mathrm{SCF}} = E^{\mathrm{SCF}}(\mathbf{P}_{\mathrm{SYS}}) - E^{\mathrm{SCF}}(\mathbf{P}_{\mathrm{ALMO}}) \tag{3.3c}$$

Each of the MP2 intermediate energies are calculated with the equivalent density, with the addition of $\Delta E_{\mathrm{DISP}}^{(2)}$. With MP2, the correction to the HF energy is a function of doubly excited Slater determinants. As such, the MP2 intermediate energies are also evaluated with different constraints on the doubly excited Slater determinants. The different types of double

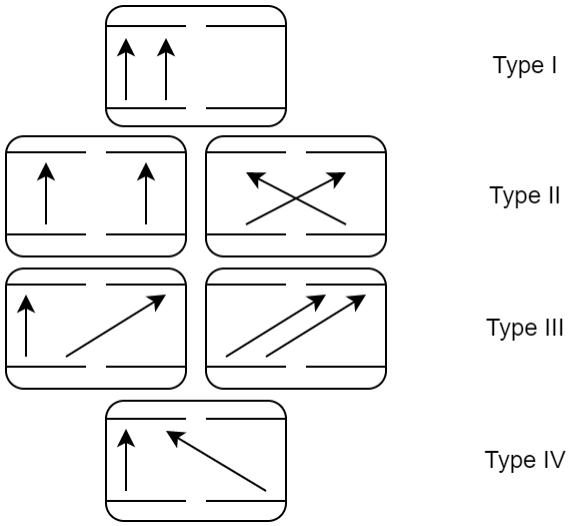


Figure 3.1: The four types of double excitations used to define constraints for each EDA term. Occupied orbitals are fragment-tagged, and this diagram should be read as though the virtuals are also fragment-tagged. The true correlation energy associated with a fragment at either the frozen or polarized level is then represented by the first row of the figure (Type I). The third line (Type III), represent basis set superposition effects, that can be removed by counterpoise correction. The second and fourth line (Type II/IV) are associated with dispersion interactions.

excitations can be categorized into four types, which are summarized in Figure 3.1 and the summary of MP2 ALMO-EDA-II terms is summarized in Table 3.1. With the inclusion of the dispersion term, $\Delta E^{(2)}_{\mathrm{DISP}}$, the corresponding intermediate energy, $E^{(2)}_{\mathrm{CCC}}(\mathbf{P}_{\mathrm{ALMO}})$ contains both type II and IV excitations. It is important to note that type III excitations are considered a type of basis set superposition error (BSSE) for the purposes of MP2 ALMO-EDA-II and therefore the standard counterpoise correction for the correlation component of BSSE is applied to $\Delta E^{(2)}_{\mathrm{FRZ}}$ as this is the first term that contains type III excitations.

Intermediate Term	Corresponding EDA Component	Excitation Types Included
$E^{(2)}_{\mathrm{F}}(\mathbf{P}_{\mathrm{F}})$		Type I
$E^{(2)}_{\mathrm{FRZ}}(\mathbf{P}_{\mathrm{FRZ}})$	$\Delta E^{(2)}_{\mathrm{FRZ}} = E^{(2)}_{\mathrm{FRZ}}(\mathbf{P}_{\mathrm{FRZ}}) - \sum_{\mathrm{F}}^{\mathrm{frag}} E^{(2)}_{\mathrm{F}}(\mathbf{P}_{\mathrm{F}})$	Type I, III
$E^{(2)}_{\mathrm{ALMO}}(\mathbf{P}_{\mathrm{ALMO}})$	$\Delta E^{(2)}_{\mathrm{POL}} = E^{(2)}_{\mathrm{ALMO}}(\mathbf{P}_{\mathrm{ALMO}}) - E^{(2)}_{\mathrm{FRZ}}(\mathbf{P}_{\mathrm{FRZ}})$	Type I, III
$E^{(2)}_{\mathrm{CCC}}(\mathbf{P}_{\mathrm{ALMO}})$	$\Delta E^{(2)}_{\mathrm{DISP}} = E^{(2)}_{\mathrm{CCC}}(\mathbf{P}_{\mathrm{ALMO}}) - E^{(2)}_{\mathrm{ALMO}}(\mathbf{P}_{\mathrm{ALMO}})$	Type I, II, III, IV
$E^{(2)}_{\mathrm{SYS}}(\mathbf{P}_{\mathrm{SYS}})$	$\Delta E^{(2)}_{\mathrm{CT}} = E^{(2)}_{\mathrm{SYS}}(\mathbf{P}_{\mathrm{SYS}}) - E^{(2)}_{\mathrm{CCC}}(\mathbf{P}_{\mathrm{ALMO}})$	Type I, II, III, IV

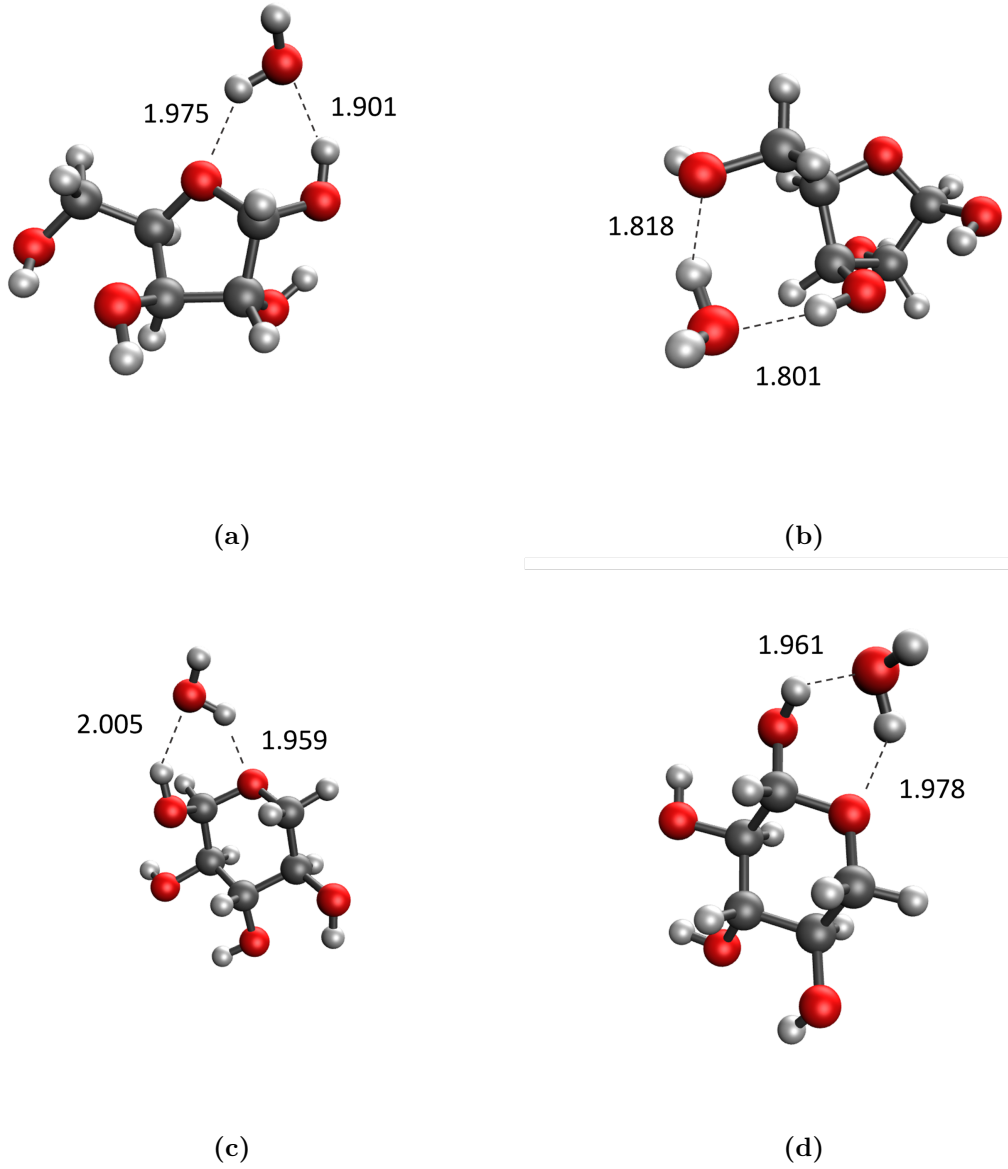
Table 3.1: Summary of Correlation Intermediate Terms Used in MP2 ALMO-EDA-II

3.3 Chemical applications

3.3.1 Hydrogen bonding in model xylose-water complexes

Hydrogen bonding (H-bonding) is a non-covalent interaction of the form $-\mathrm{X}-\mathrm{H}\cdots\mathrm{Y}$ between a proton donor on one molecule and an electronegative proton acceptor site on another molecule. More precisely, according to IUPAC:[3] The hydrogen bond is an attractive interaction between a hydrogen atom from a molecule or a molecular fragment X–H in which X is more electronegative than H, and an atom or a group of atoms in the same or a different molecule, in which there is evidence of bond formation. H-bonds involving water are important to a wide variety of inorganic and biological systems, from water itself, to protein solvation to DNA water interactions.[51, 91] Debate on the driving forces behind the water-water hydrogen bond has been on-going until even recently, with extremes of opinion ranging from an entirely resonance-based picture[123, 125, 124] to an entirely electrostatic picture.[87] Use of reliable EDA methods has tended to provide a more nuanced intermediate picture,[56, 76] where a majority of the binding energy is associated with permanent and induced electrostatics (i.e. dipole-dipole and dipole-induced dipole, etc), but there is a non-trivial contribution from dative effects corresponding to $n(\mathrm{O}) \rightarrow \sigma^*(\mathrm{H}-\mathrm{O})$ interactions.

A classic case of conventional H-bonding is the solvation of carbohydrates in water; since H-bonds between carbohydrates and water are dominated by their hydroxyl groups, they are of comparable strength to water-water H-bonds. This accounts for their often high solubility in water, as well as the significant impact of carbohydrate solutes on the structure and dynamics of water.[20, 68] At the level of hydration of carbohydrates by individual water molecules, there have been a range of experimental studies with complementary computational characterizations of the H-bonded structures with small numbers of water molecules.[105, 80, 14] From the perspective of characterizing such H-bonds via EDA, Koli et al. studied the interactions of four xylose-water complexes[61] using the LMO-EDA[112]. The structures were optimized at the RIMP2/aug-cc-pVDZ level of theory and are shown in Figure 3.2.



(a) **(b)**

(c) **(d)**

Figure 3.2: Optimized geometries for (a) α-Xylofuranose, (b) β-Xylofuranose, (c) α-Xylopyranose, (d) β-Xylopyranose interacting with a water molecule calculated with RIMP2 at aug-cc-pVDZ level of theory

The MP2 ALMO-EDA-II contributions for the four xylose-water complexes are displayed in Figure 3.3a. The interaction strengths range between 33 kJ/mol and 49 kJ/mol relative to the H-bond strength of roughly 20 kJ in the water dimer. For every complex, the interactions were stabilized primarily through charge-transfer and polarization interactions, with minor contributions from dispersion. Overall, the contributions of electrostatics and Pauli repulsion

is repulsive, however, this is not true at the SCF level for ALMO-EDA-II shown in Figure 3.3b. This can partially be attributed to the tendency of Hartree Fock to overestimate dipole moments and underestimate polarizability[40, 36, 35] and therefore overestimate electrostatic attraction as seen in similar H-bonded systems[118]. The β-xylofuranose-water complex is the most stable structure of the four. Of the four conformations, only β-xylofuranose-water has two H-bonds involving the water molecule with two hydroxyl groups on the xylose. The other three systems have an H-bond to one hydroxyl group on the xylose and one H-bond with the ether group on the xylose. This is supported by the shorter water-xylose distance for β-xylofuranose-water and is accompanied by the slight increase in charge transfer, polarization, and dispersion, as well as an increase in a repulsive frozen term due to the increased overlap and therefore increased Pauli repulsion compared to the other three complexes.

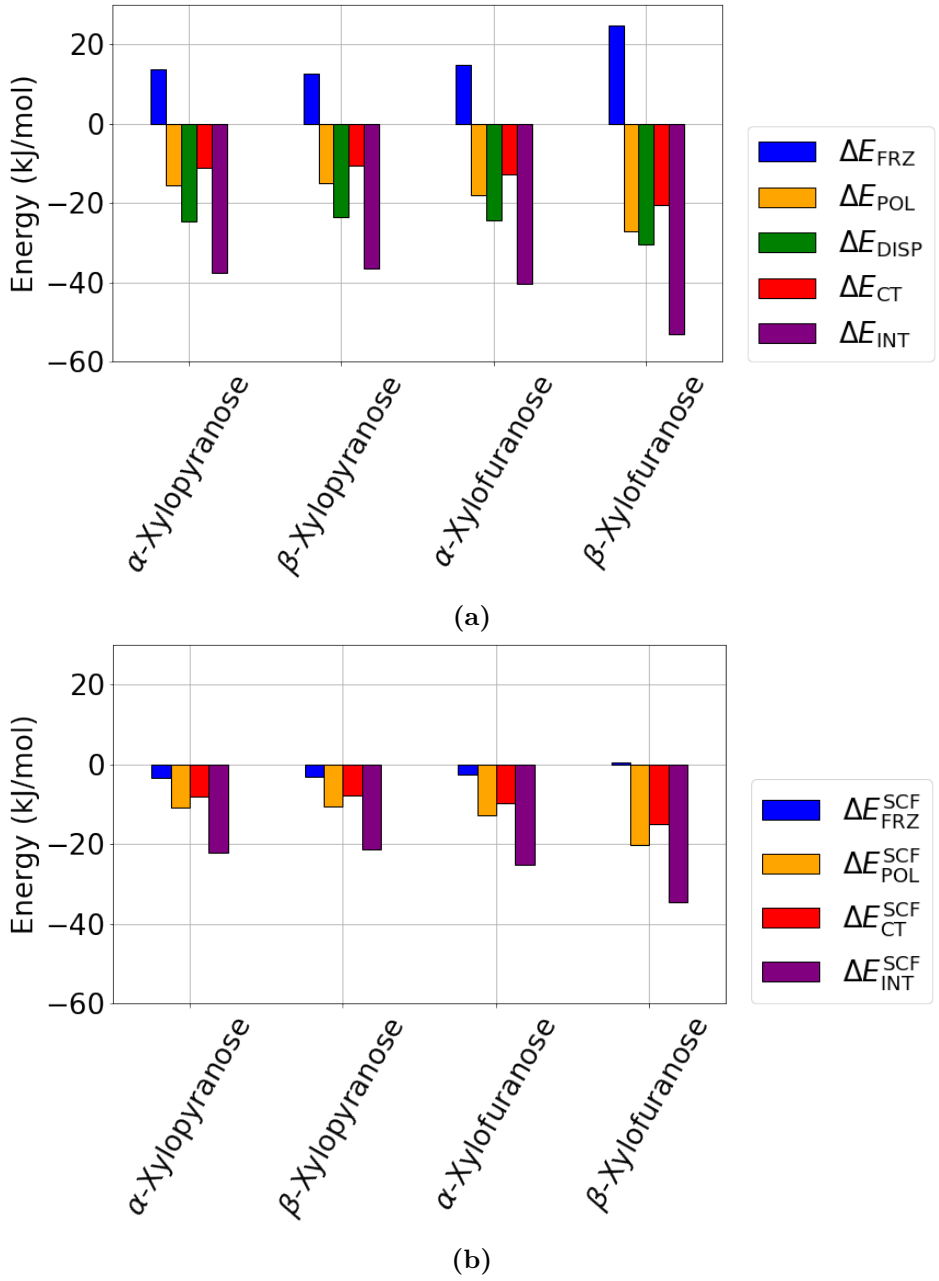


Figure 3.3: (a)MP2 ALMO-EDA-II and (b) SCF ALMO-EDA-II results for xylose-water systems calculated at aug-cc-pVDZ level of theory

Compared to the LMO-EDA results shown in Figure 3.4, SCF ALMO-EDA-II shows similar results. LMO-EDA computes the the effects of electrostatics ($\Delta E_{\mathrm{Es}}^{\mathrm{LMO}}$), exchange ($\Delta E_{\mathrm{Ex}}^{\mathrm{LMO}}$), and Pauli repulsion ($\Delta E_{\mathrm{Rep}}^{\mathrm{LMO}}$) separately[112] instead of a single frozen term as in ALMO-EDA, although it can be further decomposed into permanent electrostatics, dis-

persion, and Pauli repulsion contributions[42, 73]. However, LMO-EDA does not separate the charge transfer and polarization effects, but instead combines them into a single $\Delta E_{\mathrm{POL}}^{\mathrm{LMO}}$ term[112]. In order to reduce confusion with MP2 EDA II's polarization term, this paper will refer to this LMO-EDA term as $\Delta E_{\mathrm{ORB}}^{\mathrm{LMO}}$ (following e.g. the ETS-NOCV EDA[12, 82]). Additionally, LMO-EDA does not decompose the MP2 correlation component of the interaction energy and instead uses $\Delta E_{\mathrm{DISP}}^{\mathrm{LMO}}$ to describe the entire correlation component of the interaction. Therefore, LMO-EDA fails to account for the MP2 correction to the electrostatics and polarizability and therefore underestimates the increased repulsion in the frozen term.

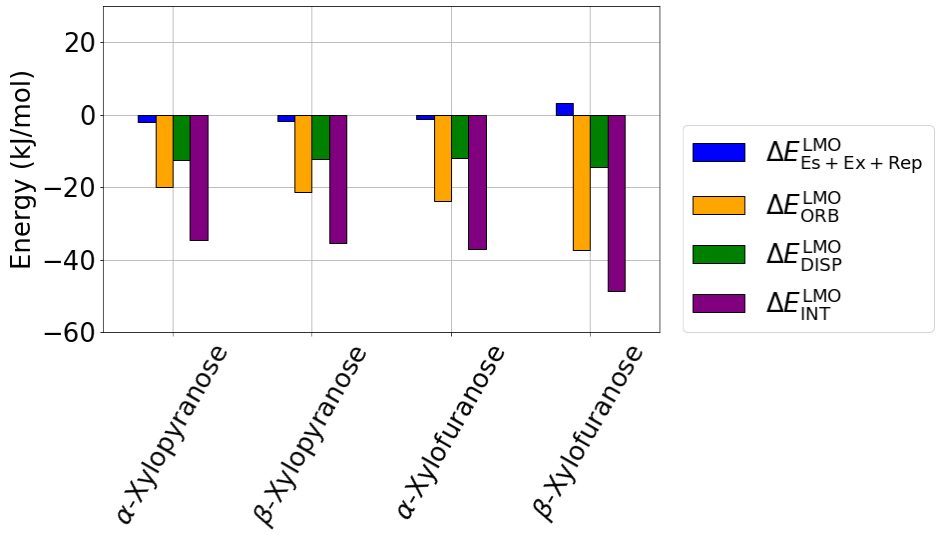


Figure 3.4: LMO-EDA results for xylose-water systems calculated at aug-cc-pVDZ level of theory[61]

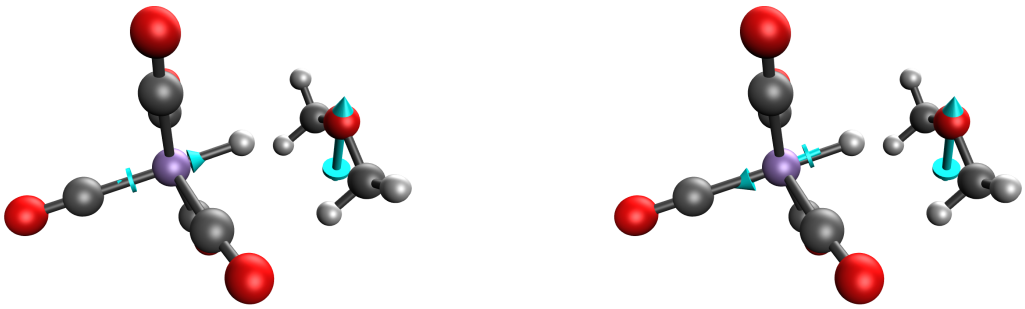
3.3.2 M-H...Y H-bonds

As exemplified in the mode carbohydrate-water cases discussed above, the conventional H-bond involves a hydrogen atom attached to a more electronegative atom interacting with an electronegative electron donor. However, there have been multiple examples of systems that do not follow the conventional H-bond definition in organometallic chemistry, where metal hydride hydrogens are sometimes capable of being proton donors, despite their formal negative charge.[1, 22] Sahoo et al. investigated the nature of some less conventional H-bonded systems involving M−H$\cdots$Y (M = Mn, Fe, Co; Y = O, S, Se) H-bonds[98]. Sahoo studied the interaction between three metal carbonyl hydrides ($Mn(CO)_5H$, $Fe(CO)_4H_2$, and $Co(CO)_4H$)) interacting with either dimethyl ether (DME), dimethylsulfide (DMS), or dimethylselenide (DMSe). They observed that the M−H$\cdots$Y H-bond was similar in strength to conventional H-bonds, and suggested that it was "dispersive" in origin.

Comparing the results from MP2-EDA-II in Figure 3.6, and LMO-EDA in Figure 3.8, there is a large discrepancy in the dispersion term and frozen term. As the correlation component of the interaction is represented entirely by $\Delta_{\mathrm{DISP}}^{\mathrm{LMO}}$, any effect of correlation on the frozen term is omitted. For these systems, the differences in electrostatics between SCF and MP2 are significant. Table 3.2 shows the dipole moments along the M-H axis for the metal center fragments calculated with MP2 and SCF at aug-cc-pVDZ in their M-Y...DME geometries. As shown in Figure 3.5, using Mn-H...DME as an example, the dipole moment of the metal carbonyl hydrides is aligned to be repulsive with the dipole moment of the H-bond acceptors (DME, DMS, DMSe) when calculated using HF, while the MP2 dipole moment is aligned to be attractive. This result supports MP2 ALMO-EDA-II as the correlation reduces the repulsive SCF frozen term. Therefore, it is vital to include the effect of correlation in the frozen term, and it is disingenuous to attribute the correlation energy to dispersion alone.

Species	HF Dipole moment(a.u)	MP2 Dipole moment(a.u)	HF Dipole-Dipole angle (deg)	MP2 Dipole-Dipole angle (deg.)
$Mn(CO)_5H$	-0.5953	0.4677	91.32	-95.4
$Co(CO)_4H$	-0.7487	0.4049	107.6	-76.1
$Fe(CO)_4H_2$	-0.7446	0.5365	46.6	-134.3

Table 3.2: Dipole moments for metal carbonyl hydrides along M-H (M = Mn, Co, Fe) bond axis in a.u. calculated with HF and MP2 at aug-cc-pVDZ. Angle is for the angle between the dipole moment vectors of the fragments



(a) **(b)**

Figure 3.5: Dipole moment orientation for isolated fragments of Mn-H...DME complex calculated with (a) HF and (B) MP2 at aug-cc-pVDZ

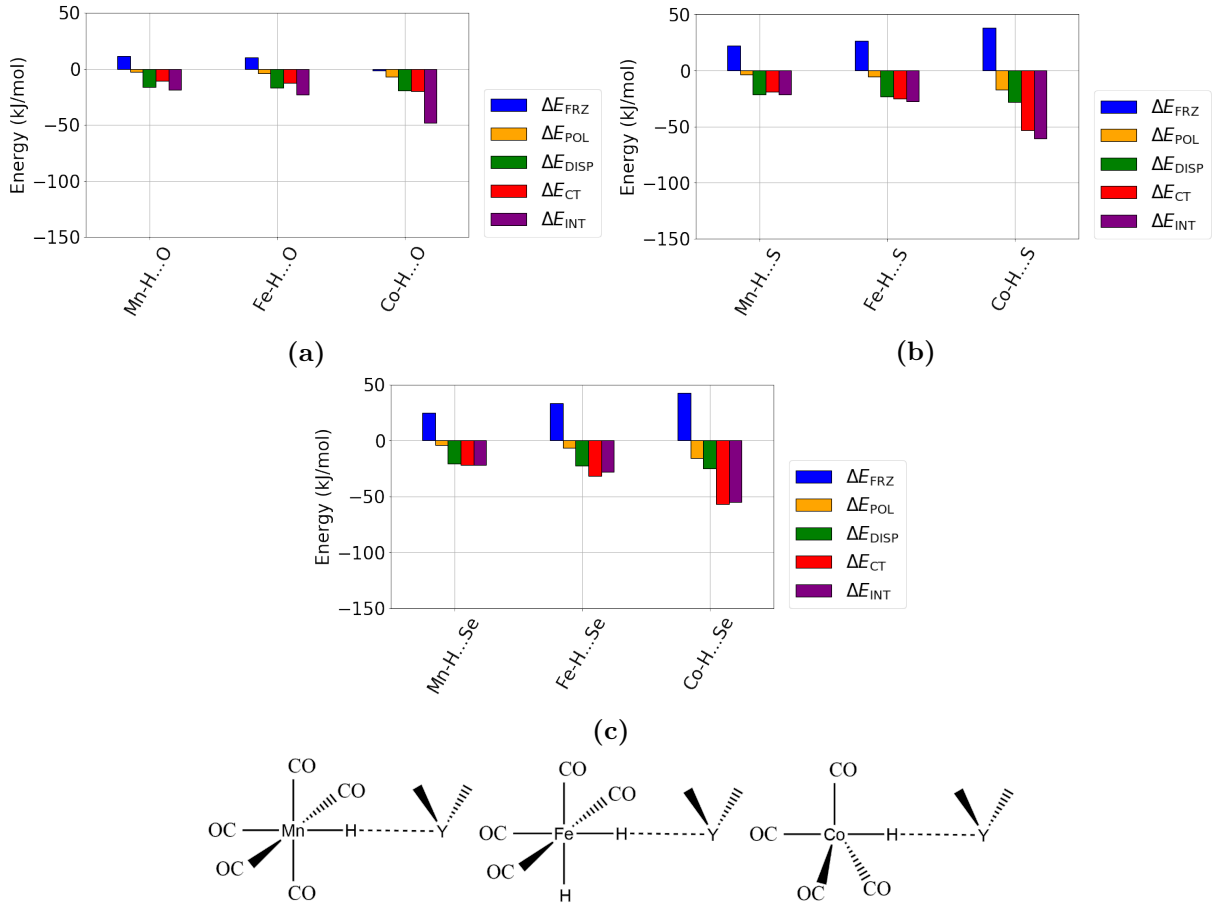


Figure 3.6: MP2 ALMO-EDA-II results for (a) DME (Y = O), (b) DMS (Y = S), and (c) DMSe (Y = Se) interacting with $Mn(CO)_5H$, $Fe(CO)_4H_2$, and $Co(CO)_4H$ calculated at aug-cc-pVDZ level of theory

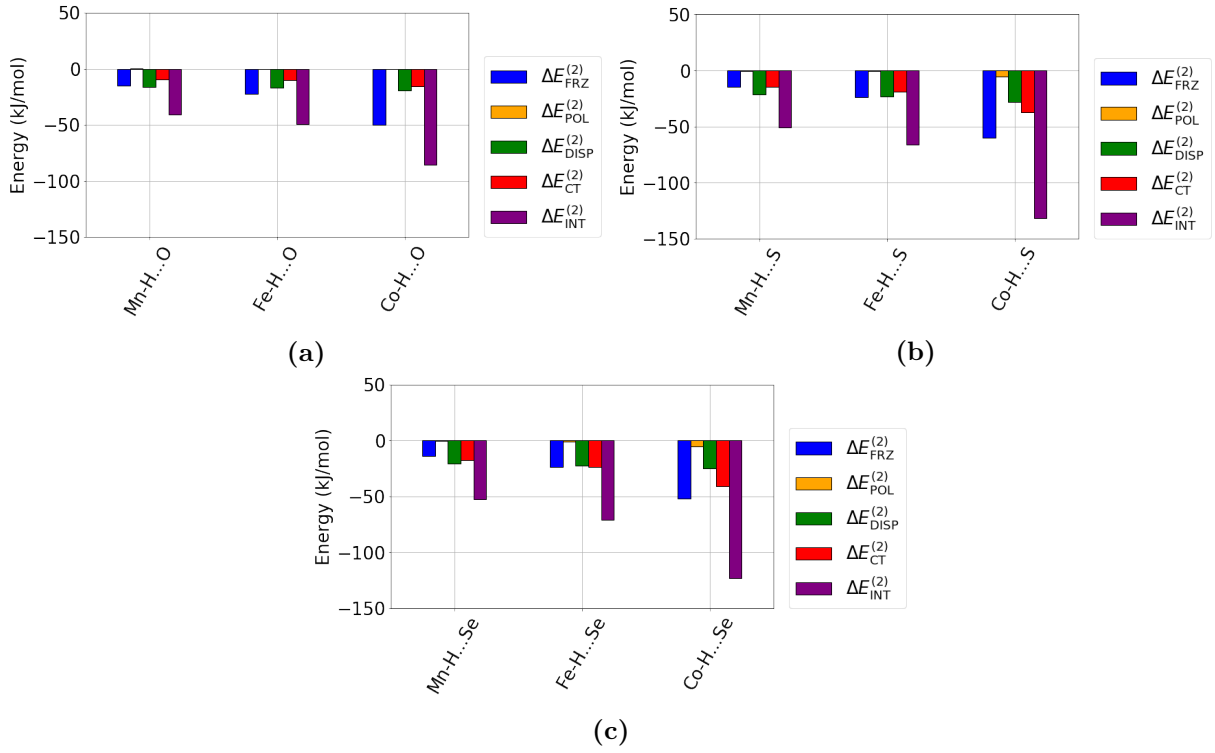


Figure 3.7: MP2 ALMO-EDA-II correlation results for (a) DME (Y = O), (b) DMS (Y = S), and (c) DMSe (Y = Se) interacting with $Mn(CO)_5H$, $Fe(CO)_4H_2$, and $Co(CO)_4H$ calculated at aug-cc-pVDZ level of theory

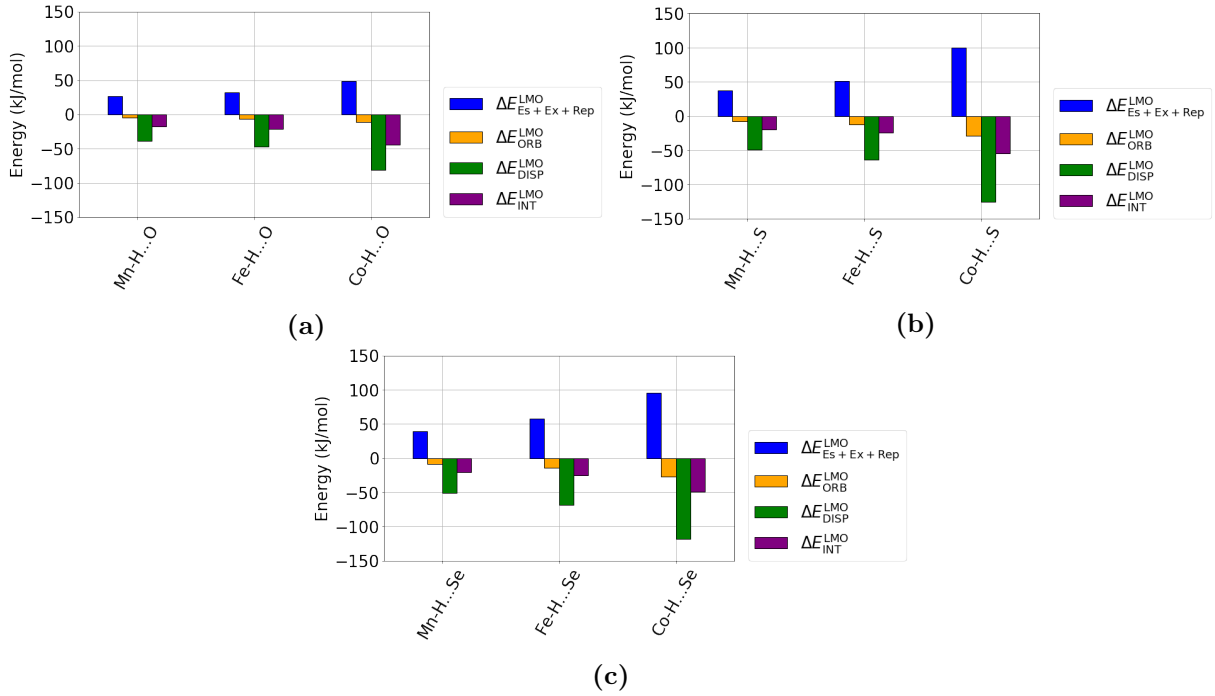
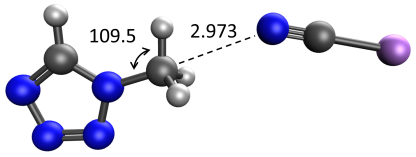


Figure 3.8: LMO-EDA results for (a) dimethyl ether, (b) dimethylsulfide, and (c) dimethylselenide interacting with $Mn(CO)_5H$, $Fe(CO)_4H_2$, and $Co(CO)_4H$ calculated at aug-cc-pVDZ level of theory[98]

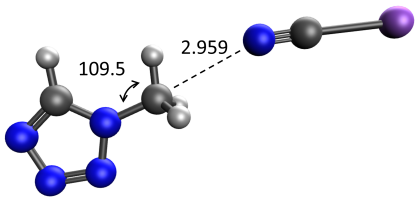
3.3.3 Tetrel Bonding

Tetrel bonding (Tt-bond) is another type of attractive non-covalent interaction, similar to H-bonding, with a group 14 atom (C, Si, Ge, etc) on one fragment as an electron acceptor and an electron donor on another fragment[9]. Tetrel bonding has attracted significant attention recently due to its promising applications in catalysis and other fields.[10, 27] Similar to the more well known H-bonding and halogen bonding, Tt-bonding has been categorized as a σ-hole type interaction. [8] However, previous work on halogen bonding[117], has shown that the σ-hole picture may not be sufficient in fully characterizing halogen bonds, which motivates a closer inspection of the origins of Tt-bonding as well. Recently, Wu et al. investigated Tt-bonding involving TA-TtX_3 (Tt = C, Si; X = H, F) interacting with CNM (M = Na, Li)[126]. Figure 3.9 displays the eight complexes studied. Figure 3.10 shows the results of MP2 ALMO-EDA-II for each of the Tt-bonded systems along with the interaction energy with the geometry distortion term included as it is significant for the TA-SiX_3 systems. For the TA-CH_3 systems, the interaction is primarily electrostatically driven. However, there is a significant contribution from polarization and dispersion as well. Comparatively, TA-CF_3 systems however show smaller frozen terms and lower overall interactions. This is likely due to the increased distance between the fragments for the TA-CF_3 systems compared to

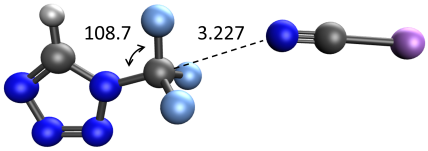
the TA-CH_3 systems, due to the increased Pauli repulsion associated with CF_3 compared to CH_3.



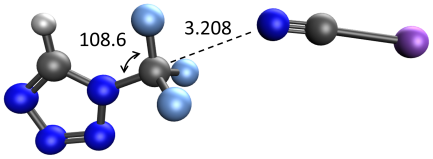
(a) TA-CH_3 ... NCLi



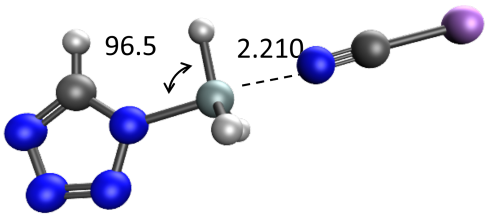
(b) TA-CH_3 ... NCNa



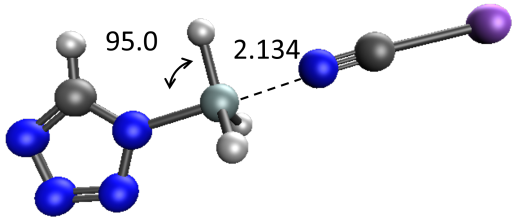
(c) TA-CF_3 ... NCLi



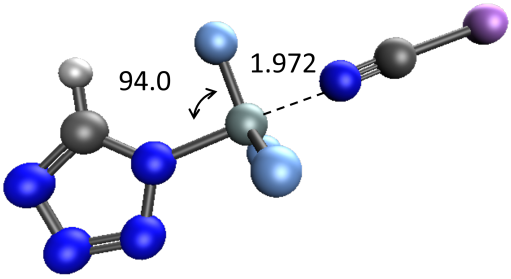
(d) TA-CF_3 ... NCNa



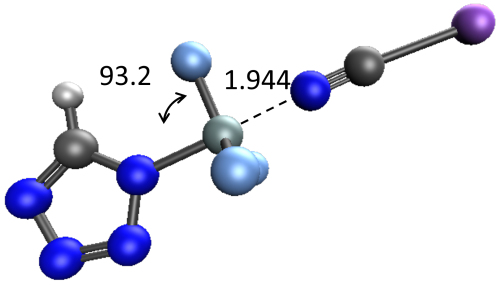
(e) TA-SiH_3 ... NCLi



(f) TA-SiH_3 ... NCNa



(g) TA-SiF_3 ... NCLi



(h) TA-SiF_3 ... NCNa

Figure 3.9: Optimized structures of N-TX_3...NCM (T = C, Si; X = H, F; M = Li, Na) tetrel bonded complexes with T-N distance ($\AA$) and N-T-X angle (deg).[126]

For the TA-SiH_3 and TA-SiF_3 systems, Figure 3.10b, there is a similar trend in frozen energy, although the frozen interaction is strongly repulsive due to the shorter distance between fragments compared to the TA-CX_3 systems as well as larger atomic radius of Si compared to C. The shorter distance is likely due to the large attractive charge transfer and polarization terms for these systems. This shorter distance however is coupled with a larger geometry distortion term as the fluoride and hydrogen atoms on the silicon on the TA-SiX_3 fragment "open up" as indicated by the lower N-Si-X angles in Figure 3.9. The ability for the TA-SiX_3 complexes to "open up" is due to greater ability for silicon to hybridize compared to carbon, and therefore is subjected to a relatively lower repulsive geometric distortion penalty[100].

Looking at the correlation components of the MP2 ALMO-EDA-II terms in Figure 3.11, although there is a very small correction to the interaction energy, there is a large correlation correction to the frozen term for the TA-SiF_3 systems. Therefore, as with the previous examples, it is important to consider the effect of correlation on each individual term.

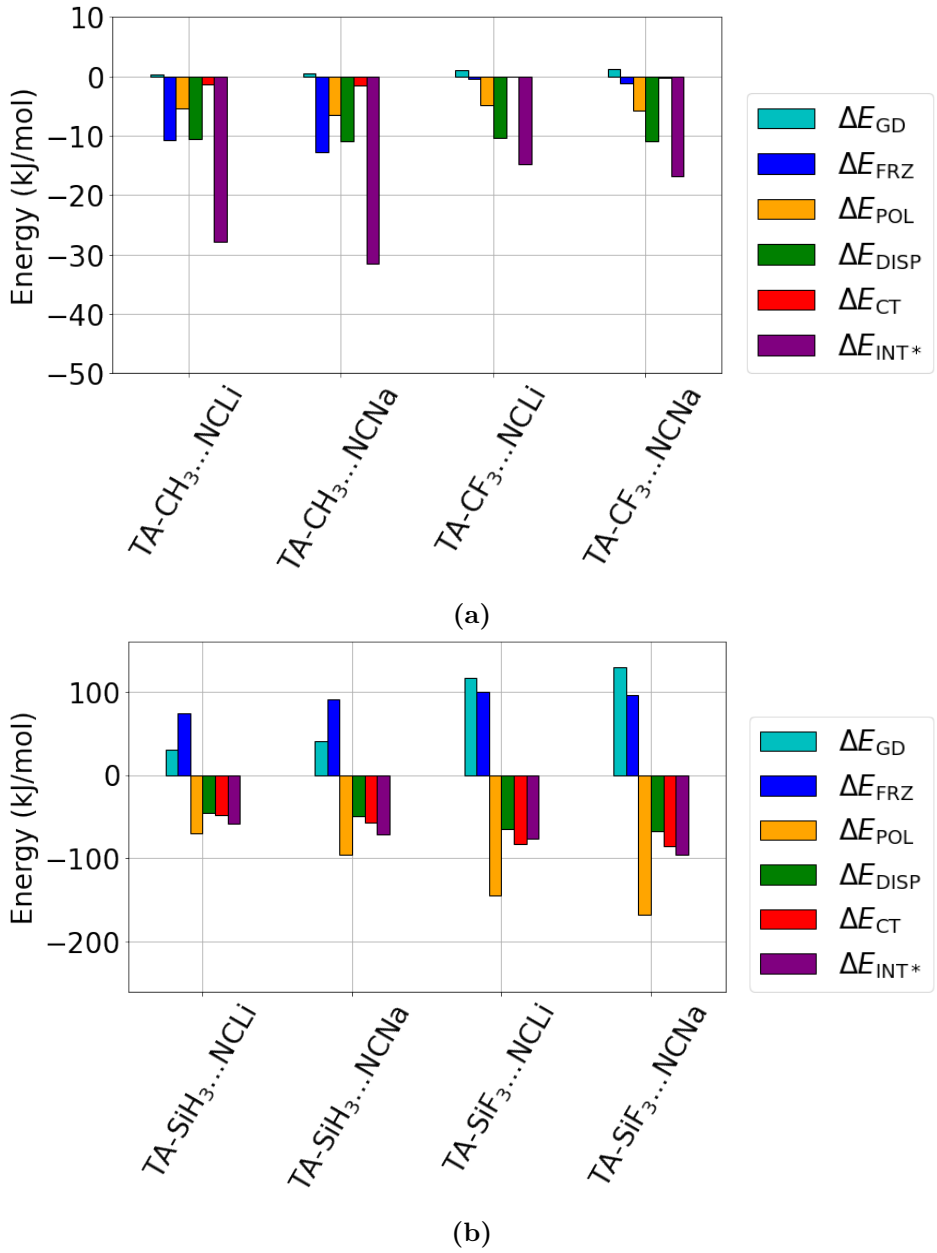


Figure 3.10: MP2 ALMO-EDA-II results for (a) TA-CX_3...CNM and (b) TA-SiX_3...CNM (X = H, F; M = Li, Na) calculated at aug-cc-pVTZ level of theory. $\Delta E_{\mathrm{INT}^*}$ includes contributions due to geometry distortion

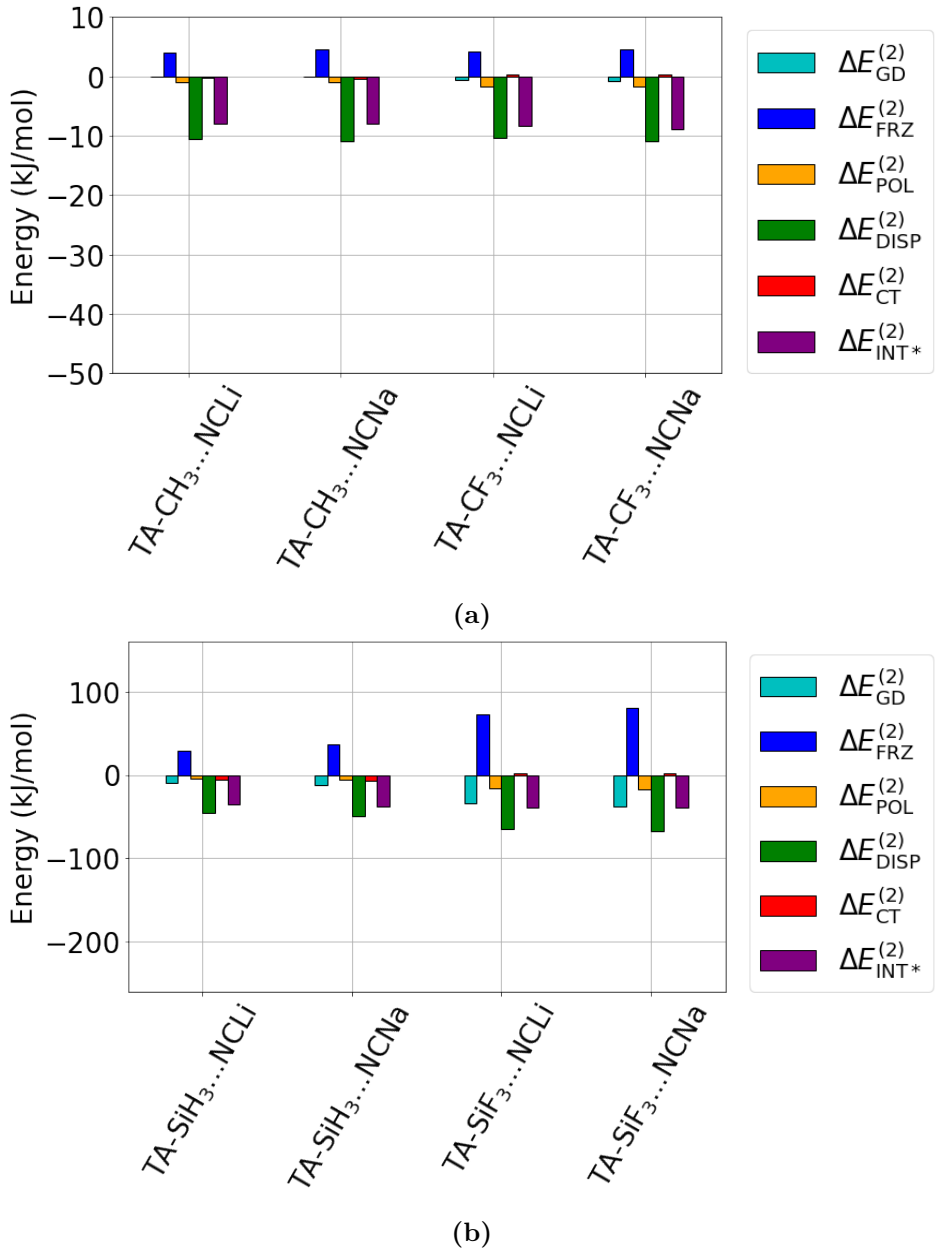


Figure 3.11: Correlation components for MP2 ALMO-EDA-II for (a) TA-CX_3...CNM and (b) TA-SiX_3...CNM (X = H, F; M = Li, Na) calculated at aug-cc-pVTZ level of theory$\Delta E^{(2)}_{\mathrm{INT}^*}$ includes contributions due to geometry distortion

3.3.4 Halogen Bonding

Conventionally, halogen bonding (X-bond) is described as the interaction between an electrophilic region of a halogen atom on one fragment and a nucleophilic region on another

fragment.[119, 18, 97] Halogen bonding is a strong intermolecular interaction (relative to H-bonding) and is found in many applications such as crystal engineering, organocatalysis, and drug design[15]. Robertson et al. discovered an X-bonded complex, tetramethylthiourea-iodine, whose stability is insensitive to the nature of the solvent[97]. Robertson's results suggested that charge transfer is a major contribution to the stability of this complex. To investigate these claims, MP2 ALMO-EDA-II calculations were performed and the results for the interaction of iodine with tetramethylthiourea (R=S...I_2) versus tetramethylurea (R=O...I_2), R = $(Me_2N)_2C$ are shown in Figure 3.12. The geometries of the complexes are shown in Figure 3.13 If the interaction was electrostatically driven, one would expect that the R=O...I_2 complex would have a stronger interaction due to the more electronegative oxygen. However, the R=S...I_2 has a stronger interaction energy that is dominated by charge transfer and dispersion. As R=S...I_2 is stabilized through both dispersion and charge transfer dispersive interactions, this supports the findings of Robertson et al. of the stability of this complex with respect to solvent.

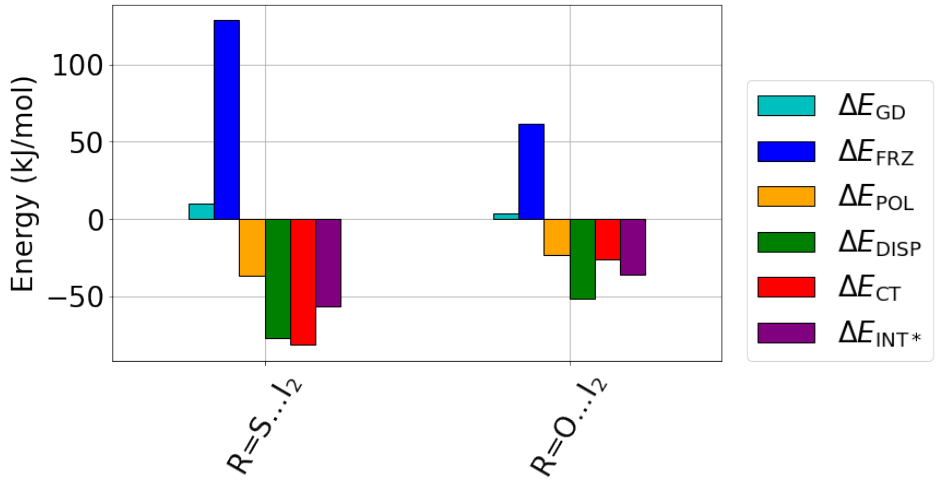


Figure 3.12: MP2 ALMO-EDA-II results for R=S...I_2 and R=O...I_2 (R = $(Me_2N)_2C$) complexes calculated at def2-TZVPPD level of theory[61]

(a) **(b)**

Figure 3.13: Optimized structures of (a) tetramethylthiourea-iodine and (b) tetramethylurea-iodine optimized with MP2 at def2-SVPD level of theory, marked with S...I / O...I distances ($\AA$) and C=S...I / C=O...I angles (deg)

3.4 Conclusions

In this paper, we have investigated a variety of intermolecular interactions, including hydrogen bonding, tetrel bonding, and halogen bonding, using the newly developed MP2 ALMO-EDA-II. These systems have shown the usefulness of MP2 ALMO-EDA-II and the importance of an accurate treatment of the correlation component of the interaction energy. Future work includes the extension of MP2 ALMO-EDA-II to other post-HF ALMO EDA's such as an EDA based on orbital optimized MP2, double hybrid DFT, or coupled cluster theory.

Bibliography

[1] Ibon Alkorta, Isabel Rozas, and José Elguero. "Non-conventional hydrogen bonds". In: *Chem. Soc. Rev.* 27.2 (1998), pp. 163–170.

[2] Ahmet Altun, Frank Neese, and Giovanni Bistoni. "Effect of electron correlation on intermolecular interactions: A pair natural orbitals coupled cluster based local energy decomposition study". In: *J. Chem. Theory Comput.* 15.1 (2018), pp. 215–228.

[3] Elangannan Arunan et al. "Definition of the hydrogen bond (IUPAC Recommendations 2011)". In: *Pure Appl. Chem.* 83.8 (2011), pp. 1637–1641.

[4] R. J. Azar et al. "Useful lower limits to polarization contributions to intermolecular interactions using a minimal basis of localized orthogonal orbitals: theory and analysis of the water dimer". In: *J. Chem. Phys.* 138 (2013), p. 084102. DOI: 10.1063/1.4792434.

[5] R. Julian Azar and Martin Head-Gordon. "An energy decomposition analysis for intermolecular interactions from an absolutely localized molecular orbital reference at the coupled-cluster singles and doubles level". In: *J. Chem. Phys.* 136.2 (2012), p. 024103. ISSN: 0021-9606. DOI: 10.1063/1.3674992. URL: http://dx.doi.org/10.1063/1.3674992.

[6] R. J. Bartlett. "Coupled-cluster theory and its equation-of-motion extensions". In: *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* 2 (2012), pp. 126–138. DOI: 10.1002/wcms.76.

[7] R. J. Bartlett and M Musial. "Coupled-cluster theory in quantum chemistry". In: *Rev. Mod. Phys.* 79.1 (2007), pp. 291–352.

[8] Antonio Bauzá, Tiddo J. Mooibroek, and Antonio Frontera. "Tetrel Bonding Interactions". In: *The Chemical Record* 16.1 (2016), pp. 473–487. DOI: https://doi.org/10.1002/tcr.201500256. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/tcr.201500256. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/tcr.201500256.

[9] Antonio Bauzá, Tiddo J. Mooibroek, and Antonio Frontera. "Tetrel-bonding interaction: Rediscovered supramolecular force?" In: 52.47 (2013), pp. 12317–12321. ISSN: 14337851. DOI: 10.1002/anie.201306501.

[10] Deepak Behera et al. "Ruthenium(II)-Catalyzed Regioselective 1,2-Hydrosilylation of N-Heteroarenes and Tetrel Bonding Mechanism". In: *ACS Catalysis* 11.10 (2021), pp. 5885–5893. ISSN: 21555435. DOI: 10.1021/acscatal.1c01148.

[11] Luke W Bertels, Joonho Lee, and Martin Head-Gordon. "Third-Order Møller–Plesset Perturbation Theory Made Useful? Choice of Orbitals and Scaling Greatly Improves Accuracy for Thermochemistry, Kinetics, and Intermolecular Interactions". In: *J. Phys. Chem. Lett.* 10.15 (2019), pp. 4170–4176.

[12] F Matthias Bickelhaupt and Evert Jan Baerends. "Kohn-Sham density functional theory: predicting and understanding chemistry". In: (2000), pp. 1–86.

[13] S. F. Boys and F. Bernardi. In: *Mol. Phys.* 19 (1970), p. 553. DOI: 10.1080/00268977000101561.

[14] Pierre Çarçabal, Emilio J Cocinero, and John P Simons. "Binding energies of micro-hydrated carbohydrates: measurements and interpretation". In: *Chem. Sci.* 4.4 (2013), pp. 1830–1836.

[15] Gabriella Cavallo et al. "The halogen bond". In: *Chem. Rev.* 116.4 (2016), pp. 2478–2601. ISSN: 15206890. DOI: 10.1021/acs.chemrev.5b00484.

[16] Timothy Clark et al. "Halogen bonding: the σ-hole". In: *Journal of molecular modeling* 13.2 (2007), pp. 291–296.

[17] D. Cremer. "Moller-Plesset perturbation theory: from small molecule methods to methods for thousands of atoms". In: *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* 1.4 (2011), pp. 509–530. ISSN: 1759-0876. DOI: 10.1002/wcms.58. URL: %3CGo%20to%20ISI%3E://WOS:000296005000003.

[18] Gautam R. Desiraju et al. In: *Pure Appl. Chem.* 85.8 (2013), pp. 1711–1713. DOI: doi:10.1351/PAC-REC-12-05-10. URL: https://doi.org/10.1351/PAC-REC-12-05-10.

[19] R. A. DiStasio, Jr. and M. Head-Gordon. "Optimized spin-component scaled second-order Møller-Plesset perturbation theory for intermolecular interaction energies". In: *Mol. Phys.* 105 (2007), p. 1073. DOI: 10.1080/00268970701283781.

[20] Søren B Engelsen et al. "The diluted aqueous solvation of carbohydrates as inferred from molecular dynamics simulations and NMR spectroscopy". In: *Biophys. Chem.* 93.2-3 (2001), pp. 103–127.

[21] Evgeny Epifanovsky et al. "Software for the frontiers of quantum chemistry: An overview of developments in the Q-Chem 5 package". In: *J. Chem. Phys.* 155.8 (2021), p. 084801.

[22] Lina M. Epstein and Elena S. Shubina. "New types of hydrogen bonding in organometallic chemistry". In: *Coord. Chem. Rev.* 231.1-2 (2002), pp. 165–181. ISSN: 00108545. DOI: 10.1016/S0010-8545(02)00118-2.

[23] George S Fanourgakis, Edoardo Apra, and Sotiris S Xantheas. "High-level ab initio calculations for the four low-lying families of minima of (H 2 O) 20. I. Estimates of MP2/CBS binding energies and comparison with empirical potentials". In: *J. Chem. Phys.* 121.6 (2004), pp. 2655–2663.

[24] Dmitri G Fedorov and Kazuo Kitaura. "Pair interaction energy decomposition analysis". In: *J. Comput. Chem.* 28.1 (2007), pp. 222–237.

[25] Reinhold F Fink. "Why does MP2 work?" In: *J. Chem. Phys.* 145.18 (2016), p. 184101.

[26] Regina F. Frey and Ernest R. Davidson. "Energy partitioning of the self-consistent field interaction energy of ScCO". In: *J. Chem. Phys.* 90.10 (1989), pp. 5555–5562. DOI: 10.1063/1.456408. eprint: https://doi.org/10.1063/1.456408. URL: https://doi.org/10.1063/1.456408.

[27] Xavier García-Llinás et al. "Importance of R-CF3O Tetrel Bonding Interactions in Biological Systems". In: *J. Phys. Chem. A* 121.28 (2017), pp. 5371–5376. ISSN: 15205215. DOI: 10.1021/acs.jpca.7b06052.

[28] Soumen Ghosh et al. "Fragment-Based Local Coupled Cluster Embedding Approach for the Quantification and Analysis of Noncovalent Interactions: Exploring the Many-Body Expansion of the Local Coupled Cluster Energy". In: *J. Chem. Theory Comput.* 17.6 (2021), pp. 3348–3359.

[29] E. Gianinetti, M. Raimondi, and E. Tornaghi. "Modification of the Roothaan Equations to Exclude BSSE from Molecular Interaction Calculations". In: *Int. J. Quantum Chem.* 60 (1996), p. 157. DOI: 10.1002/(SICI)1097-461X(1996)60:1<157::AID-QUA17>3.0.CO;2-C.

[30] M. Goldey, A. Dutoi, and M. Head-Gordon. "Attenuated second order Moller-Plesset theory: Assessment and performance in the aug-cc-pVTZ basis". In: *Phys. Chem. Chem. Phys.* 15 (2013), p. 15869. DOI: 10.1039/c3cp51826d.

[31] M. Goldey and M. Head-Gordon. "Attenuating away the errors in inter- and intramolecular interactions from second-order Møller–Plesset calculations in the small aug-cc-pVDZ basis set". In: *J. Phys. Chem. Lett.* 3 (2012), pp. 3592–3598. DOI: 10.1021/jz301694b.

[32] Jérôme F Gonthier and Martin Head-Gordon. "Compressed representation of dispersion interactions and long-range electronic correlations". In: *J. Chem. Phys.* 147.14 (2017), p. 144110.

[33] S. Grimme. "Accurate Calculation of the Heats of Formation for Large Main Group Compounds with Spin-Component Scaled MP2 Methods". In: *J. Phys. Chem. A* 109 (2005), p. 3067. DOI: 10.1021/jp050036j.

[34] Stefan Grimme. "Improved second-order Møller–Plesset perturbation theory by separate scaling of parallel-and antiparallel-spin pair correlation energies". In: *J. Chem. Phys.* 118.20 (2003), pp. 9095–9102.

[35] D. Hait and M. Head-Gordon. "How accurate are static polarizability predictions from density functional theory? An assessment over 132 species at equilibrium geometry". In: *Phys. Chem. Chem. Phys.* 20 (2018), pp. 19800–19810. DOI: 10.1039/c8cp03569e.

[36] D. Hait and M. Head-Gordon. "How accurate is density functional theory at predicting dipole moments? An assessment using a new database of 200 benchmark values". In: *J. Chem. Theory Comput.* 14 (2018), pp. 1969–1981. DOI: 10.1021/acs.jctc.7b01252.

[37] Nicholas C. Handy and Henry F. Schaefer. "On the evaluation of analytic energy derivatives for correlated wave functions". In: *J. Chem. Phys.* 81.11 (1984), pp. 5031–5033. ISSN: 00219606. DOI: 10.1063/1.447489.

[38] M. Head-Gordon. In: *Mol. Phys.* 96 (1999), p. 673. DOI: 10.1080/00268979909483003.

[39] Martin Head-Gordon, Paul E Maslen, and Christopher A White. "A tensor formulation of many-electron theory in a nonorthogonal single-particle basis". In: *The Journal of chemical physics* 108.2 (1998), pp. 616–625.

[40] A Leif Hickey and Christopher N Rowley. "Benchmarking quantum chemical methods for the calculation of molecular dipole moments and polarizabilities". In: *J. Phys. Chem. A* 118.20 (2014), pp. 3678–3687.

[41] P. R. Horn and M. Head-Gordon. "Polarization contributions to intermolecular interactions revisited with fragment electric-field response functions". In: *J. Chem. Phys.* 143 (2015), p. 114111. DOI: 10.1063/1.4930534.

[42] P. R. Horn, Y. Mao, and M. Head-Gordon. "Defining the contributions of permanent electrostatics, Pauli repulsion, and dispersion in density functional theory calculations of intermolecular interaction energies". In: *J. Chem. Phys.* 144 (2016), p. 114107. DOI: 10.1063/1.4942921.

[43] P. R. Horn, Y. Mao, and M. Head-Gordon. "Probing non-covalent interactions with a second generation energy decomposition analysis using absolutely localized molecular orbitals". In: *Phys. Chem. Chem. Phys.* 18 (2016), pp. 23067–23079. DOI: 10.1039/C6CP03784D.

[44] P. R. Horn et al. "Unrestricted absolutely localized molecular orbitals for energy decomposition analysis: theory and applications to intermolecular interactions involving radicals". In: *J. Chem. Phys.* 138 (2013), p. 134119. DOI: 10.1063/1.4798224.

[45] E. Hylleraas. "Über den Grundterm der Zweielektronenprobleme von H, He, Li+, Be++ usw." In: *Zeitschrift fur Physik* September (1930).

[46] Ekaterina I Izgorodina et al. "Quantum chemical methods for the prediction of energetic, physical, and spectroscopic properties of ionic liquids". In: *Chem. Rev.* 117.10 (2017), pp. 6696–6754.

[47] Mirosław Jabłoński and W Andrzej Sokalski. "Physical nature of interactions in charge-inverted hydrogen bonds". In: *Chem. Phys. Lett.* 552 (2012), pp. 156–161.

[48] Georg Jansen. "Symmetry-adapted perturbation theory based on density functional theory for noncovalent interactions". In: *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* 4.2 (2014), pp. 127–144.

[49] Frank Jensen. "Atomic orbital basis sets". In: *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* 3.3 (2013), pp. 273–295.

[50] B. Jeziorski, R. Moszynski, and K. Szalewicz. In: *Chem. Rev.* 94 (1994), p. 1887. DOI: 10.1021/cr00031a008.

[51] Lin Jiang et al. "A "solvated rotamer" approach to modeling water-mediated hydrogen bonds at protein-protein interfaces". In: *Proteins: Structure, Function and Genetics* 58.4 (2005), pp. 893–904. ISSN: 08873585. DOI: 10.1002/prot.20347.

[52] Erin R Johnson et al. "Revealing noncovalent interactions". In: *Journal of the American Chemical Society* 132.18 (2010), pp. 6498–6506.

[53] T. H. Dunning Jr. In: *J. Chem. Phys.* 90 (1989), p. 1007. DOI: 10.1063/1.456153.

[54] Y. Jung et al. "Scaled opposite-spin second order Møller–Plesset correlation energy: an economical electronic structure method". In: *J. Chem. Phys.* 121 (2004), p. 9793. DOI: 10.1063/1.1809602.

[55] Petr Jurečka et al. "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". In: *Phys. Chem. Chem. Phys.* 8.17 (2006), pp. 1985–1993.

[56] R. Z. Khaliullin, A. T. Bell, and M. Head-Gordon. In: *Chem. Eur. J* 15 (2009), p. 851. DOI: 10.1002/chem.200802107.

[57] R. Z. Khaliullin, A. T. Bell, and M. Head-Gordon. "Analysis of charge transfer effects in molecular complexes based on absolutely localized molecular orbitals". In: *J. Chem. Phys.* 128 (2008), p. 184112. DOI: 10.1063/1.2912041.

[58] R. Z. Khaliullin, M. Head-Gordon, and A. T. Bell. "An efficient self-consistent field method for large systems of weakly interacting components". In: *J. Chem. Phys.* 124 (2006), p. 204105. DOI: 10.1063/1.2191500.

[59] R. Z. Khaliullin et al. "Unravelling the origin of intermolecular interactions using absolutely localized molecular orbitals". In: *J. Phys. Chem. A* 111 (2007), pp. 8753–8765. DOI: 10.1021/jp073685z.

[60] Kazuo Kitaura and Keiji Morokuma. "A New Energy Decomposition Scheme for Molecular Interactions within the Hartree-Fock Approximation". In: *Int. J. Quantum Chem.* 10.2 (1976), pp. 325–340.

[61] Amol R. Koli and Sachin D. Yeole. "Understanding the interactions between hydrogen-bonded complexes of xylose and water: Quantum Chemical Investigation". In: *J. Chem. Sci.* 132.1 (2020), pp. 1–7. ISSN: 09737103. DOI: 10.1007/s12039-020-1741-3. URL: https://doi.org/10.1007/s12039-020-1741-3.

[62] K. U. Lao and J. M. Herbert. "A simple correction for nonadditive dispersion within extended symmetry-adapted perturbation theory (XSAPT)". In: *J. Chem. Theory Comput.* 14 (2018), pp. 5128–5142. DOI: 10.1021/acs.jctc.8b00527.

[63] K. U. Lao and J. M. Herbert. "Accurate and efficient quantum chemistry calculations of noncovalent interactions in many-body systems: The XSAPT family of methods". In: *J. Phys. Chem. A* 119 (2015), pp. 235–253. DOI: 10.1021/jp5098603.

[64] K. U. Lao and J. M. Herbert. "Energy decomposition analysis with a stable charge-transfer term for interpreting intermolecular interactions". In: *J. Chem. Theory Comput.* 12 (2016), p. 2569. DOI: 10.1021/acs.jctc.6b00155.

[65] Ka Un Lao et al. "Accurate Description of Intermolecular Interactions Involving Ions Using Symmetry-Adapted Perturbation Theory". In: *Journal of Chemical Theory and Computation* 11 (2015), pp. 2473–2486. ISSN: 15499626. DOI: 10.1021/ct5010593.

[66] J. Lee and M. Head-Gordon. "Distinguishing artificial and essential symmetry breaking in a single determinant: Approach and application to the C_{60}, C_{36}, and C_{20} fullerenes". In: *Phys. Chem. Chem. Phys.* 21 (2019), pp. 4763–4768. DOI: 10.1039/C8CP07613H.

[67] J. Lee and M. Head-Gordon. "Regularized orbital-optimized second-order Møller–Plesset perturbation theory: A reliable fifth-order-scaling electron correlation model with orbital energy dependent regularizers". In: *J. Chem. Theory Comput.* 14 (2018), pp. 5203–5219. DOI: 10.1021/acs.jctc.8b00731.

[68] Sau Lawrence Lee, Pablo G Debenedetti, and Jeffrey R Errington. "A computational study of hydration, solution structure, and dynamics in dilute carbohydrate solutions". In: *J. Chem. Phys.* 122.20 (2005), p. 204511.

[69] Matthias Loipersberger et al. "Energy Decomposition Analysis for Interactions of Radicals: Theory and Implementation at the MP2 Level with Application to Hydration of Halogenated Benzene Cations and Complexes between CO2–· and Pyridine and Imidazole". In: *J. Phys. Chem. A* 123.44 (2019), pp. 9621–9633.

[70] Matthias Loipersberger et al. "Exploring the Limits of Second-and Third-Order Møller–Plesset Perturbation Theories for Noncovalent Interactions: Revisiting MP2. 5 and Assessing the Importance of Regularization and Reference Orbitals". In: *J. Chem. Theory Comput.* 17.9 (2021), pp. 5582–5599.

[71] Cameron J Mackie, Jérôme F Gonthier, and Martin Head-Gordon. "Compressed intramolecular dispersion interactions". In: *J. Chem. Phys.* 152.2 (2020), p. 024112.

[72] Y. Mao, P. R. Horn, and M. Head-Gordon. "Energy decomposition analysis in an adiabatic picture". In: *Phys. Chem. Chem. Phys.* 19 (2017), pp. 5944–5958. DOI: 10.1039/C6CP08039A.

[73] Y. Mao et al. "Assessing ion–water interactions in the AMOEBA force field using energy decomposition analysis of electronic structure calculations". In: *J. Chem. Theory Comput.* 12 (2016), pp. 5422–5437. DOI: 10.1021/acs.jctc.6b00764.

[74] Y. Mao et al. "Consistent inclusion of continuum solvation in energy decomposition analysis: Theory and application to molecular CO_2 reduction catalysts". In: *Chem. Sci.* 12 (2021), pp. 1398–1414. DOI: 10.1039/D0SC05327A.

[75] Y. Mao et al. "From intermolecular interaction energies and observable shifts to component contributions and back again: A tale of variational energy decomposition analysis". In: *Annu. Rev. Phys. Chem.* 72 (2021), pp. 641–666. DOI: 10.1146/annurev-physchem-090419-115149.

[76] Y. Mao et al. "On the computational characterization of charge-transfeer effects in noncovalently bound molecular complexes". In: *J. Chem. Theory Comput.* 14 (2018), pp. 2401–2417. DOI: 10.1021/acs.jctc.7b01256.

[77] Y. Mao et al. "Probing radical–molecule interactions with a second generation energy decomposition analysis of DFT calculations using absolutely localized molecular orbitals". In: *Phys. Chem. Chem. Phys.* 22 (2020), pp. 12867–12885. DOI: 10.1039/D0CP01933J.

[78] N. Mardirossian and M. Head-Gordon. "ωB97M-V: A combinatorially optimized, range-separated hybrid, meta-GGA density functional with VV10 nonlocal correlation". In: *J. Chem. Phys.* 144 (2016), p. 214110. DOI: 10.1063/1.4952647.

[79] N. Mardirossian and M. Head-Gordon. "Survival of the most transferable at the top of Jacob's ladder: Defining and testing the ωB97M(2) double hybrid density functional". In: *J. Chem. Phys.* 148 (2018), p. 241736. DOI: 10.1063/1.5025226.

[80] Nitzan Mayorkas et al. "Carbohydrate hydration: heavy water complexes of α and β anomers of glucose, galactose, fucose and xylose". In: *Phys. Chem. Chem. Phys.* 13.41 (2011), pp. 18671–18678.

[81] Alston J Misquitta. "Charge transfer from regularized symmetry-adapted perturbation theory". In: *J. Chem. Theory Comput.* 9.12 (2013), pp. 5313–5326.

[82] M. P. Mitoraj, A. Michalak, and T. Ziegler. "A Combined Charge and Energy Decomposition Scheme for Bond Analysis". In: *J. Chem. Theory Comput.* 5 (2008), p. 962. DOI: 10.1021/ct800503d.

[83] Y. R. Mo, P Bao, and J. L. Gao. "Energy decomposition analysis based on a block-localized wavefunction and multistate density functional theory". In: *Phys. Chem. Chem. Phys.* 13.15 (2011), pp. 6760–6775.

[84] Yirong Mo, Jiali Gao, and Sigrid D. Peyerimhoff. "Energy decomposition analysis of intermolecular interactions using a block-localized wave function approach". In: *J. Chem. Phys.* 112.13 (2000), pp. 5530–5538. DOI: 10.1063/1.481185. URL: http://link.aip.org/link/JCPSA6/v112/i13/p5530/s1%5C&Agg=doi.

[85] Yirong Mo, Lingchun Song, and Yuchun Lin. "Block-Localized Wavefunction (BLW) Method at the Density Functional Theory (DFT) Level". In: *J. Phys. Chem. A* 111.34 (2007), pp. 8291–8301. URL: http://pubs.acs.org/doi/abs/10.1021/jp0724065.

[86] Keiji Morokuma. "Why Do Molecules Interact? The Origin of Electron Donor-Acceptor Complexes, Hydrogen Bonding and Proton Affinity". In: *Acc. Chem. Res.* 10.8 (1977), pp. 294–300.

[87] Jane S Murray and Peter Politzer. "Hydrogen bonding: A Coulombic σ-hole interaction". In: *J. Indian Inst. Sci.* 100.1 (2020), pp. 21–30.

[88] T. Nagata et al. "Basis set superposition error free self-consistent field method for molecular interaction in multi-component systems: Projection operator formalism". In: *J. Chem. Phys.* 115 (2001), p. 3553. DOI: 10.1063/1.1388039.

[89] F. Neese et al. "Assessment of orbital-optimized, spin-component scaled second-order many-body perturbation theory for thermochemistry and kinetics". In: *J. Chem. Theory Comput.* 5 (2009), pp. 3060–3073.

[90] Brian D Nguyen et al. "Divergence of many-body perturbation theory for noncovalent interactions of large molecules". In: *J. Chem. Theory Comput.* 16.4 (2020), pp. 2258–2273.

[91] Subrata Pal, Prabal K. Maiti, and Biman Bagchi. "Exploring DNA groove water dynamics through hydrogen bond lifetime and orientational relaxation". In: *J. Chem. Phys.* 125.23 (2006). DOI: 10.1063/1.2403872.

[92] K. Patkowski. "Recent developments in symmetry-adapted perturbation theory". In: *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* (2020), e1452. DOI: 10.1002/wcms.1452.

[93] M. J.S. Phipps et al. "Energy decomposition analysis based on absolutely localized molecular orbitals for large-scale density functional theory calculations in drug design". In: *Journal of Chemical Theory and Computation* 12.7 (2016), pp. 3135–3148. ISSN: 15499626. DOI: 10.1021/acs.jctc.6b00272.

[94] Avijit Rakshit et al. "Atlas of putative minima and low-lying energy networks of water clusters n= 3–25". In: *J. Chem. Phys.* 151.21 (2019), p. 214307.

[95] Adam Rettig et al. "Revisiting the Orbital Energy-Dependent Regularization of Orbital-Optimized Second-Order Møller–Plesset Theory". In: *J. Chem. Theory Comput.* 18.9 (2022), pp. 5382–5392.

[96] Jan Rezac and Pavel Hobza. "Benchmark calculations of interaction energies in noncovalent complexes and their applications". In: *Chemical reviews* 116.9 (2016), pp. 5038–5071.

[97] Craig C Robertson et al. "A solvent-resistant halogen bond". In: *Chemical Science* 5.11 (2014), pp. 4179–4183.

[98] Dipak Kumar Sahoo et al. "Nature and Strength of M-H···S and M-H···Se (M = Mn, Fe, & Co) Hydrogen Bond". In: *J. Phys. Chem. A* 123.11 (2019), pp. 2227–2236. ISSN: 15205215. DOI: 10.1021/acs.jpca.8b12003.

[99] Andreas Savin et al. "ELF: The electron localization function". In: *Angewandte Chemie International Edition in English* 36.17 (1997), pp. 1808–1832.

[100] Steve Scheiner. "Origins and properties of the tetrel bond". In: *Phys. Chem. Chem. Phys.* 23.10 (2021), pp. 5702–5717. ISSN: 14639076. DOI: 10.1039/d1cp00242b.

[101] Wolfgang B Schneider et al. "Decomposition of intermolecular interaction energies within the local pair natural orbital coupled cluster framework". In: *J. Chem. Theory Comput.* 12.10 (2016), pp. 4778–4792.

[102] Martin Schütz, Guntram Rauhut, and Hans-Joachim Werner. "Local treatment of electron correlation in molecular clusters: Structures and stabilities of (H2O) n, n= 2- 4". In: *J. Phys. Chem. A* 102.29 (1998), pp. 5997–6003.

[103] Robert Sedlak et al. "Accuracy of quantum chemical methods for large noncovalent complexes". In: *J. Chem. Theory Comput.* 9.8 (2013), pp. 3364–3374.

[104] James Shee et al. "Regularized second-order Møller–Plesset theory: A more accurate alternative to conventional MP2 for noncovalent interactions and transition metal thermochemistry for the same computational cost". In: *J. Phys. Chem. Lett.* 12 (2021), pp. 12084–12097.

[105] John P Simons et al. "Sugars in the gas phase. Spectroscopy, conformation, hydration, co-operativity and selectivity". In: *Int. Rev. Phys. Chem.* 24.3-4 (2005), pp. 489–531.

[106] J. C. Slater. In: *Phys. Rev.* 34 (1929), p. 1293. DOI: 10.1103/PhysRev.34.1293.

[107] W Andrzej Sokalski and Szczepan M Roszak. "Efficient techniques for the decomposition of intermolecular interaction energy at SCF level and beyond". In: *J. Mol. Struct. (Theochem)* 234 (1991), pp. 387–400.

[108] H. Stoll, G. Wagenblast, and H. Preuss. "On the Use of Local Basis Sets for Localized Molecular Orbitals". In: *Theor. Chem. Acc.* 57 (1980), p. 169. DOI: 10.1007/BF00574903.

[109] A. J. Stone and A. J. Misquitta. "Charge-transfer in symmetry-adapted perturbation theory". In: *Chem. Phys. Lett.* 473 (2009), pp. 201–205. DOI: 10.1016/j.cplett.2009.03.073.

[110] Anthony Stone. *The theory of intermolecular forces*. Oxford University Press, 2013.

[111] Anthony J Stone. "Natural bond orbitals and the nature of the hydrogen bond". In: *J. Phys. Chem. A* 121.7 (2017), pp. 1531–1534.

[112] Peifeng Su and Hui Li. "Energy decomposition analysis of covalent bonds and intermolecular interactions". In: *J. Chem. Phys.* 131.1 (2009), p. 014102. URL: http://www.ncbi.nlm.nih.gov/pubmed/19586091.

[113] Peifeng Su, Zhen Tang, and Wei Wu. "Generalized Kohn-Sham energy decomposition analysis and its applications". In: *WIREs: Comput. Mol. Sci.* (2020), e1460.

[114] A. Szabo and N. S. Ostlund. *Modern Quantum Chemistry*. Dover, 1996.

[115] K. Szalewicz. "Symmetry-adapted perturbation theory of intermolecular forces". In: *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* 2 (2012), pp. 254–272. DOI: 10.1002/wcms.86.

[116] J. Thirman and M. Head-Gordon. "An energy decomposition analysis for second-order Møller–Plesset perturbation theory based on absolutely localized molecular orbitals". In: *J. Chem. Phys.* 143 (2015), p. 084124. DOI: 10.1063/1.4929479.

[117] J. Thirman and M. Head-Gordon. "Efficient implementation of energy decomposition analysis for second-order Møller–Plesset perturbation theory and application to anion–π interactions". In: *J. Phys. Chem. A* 121 (2017), pp. 717–728. DOI: 10.1021/acs.jpca.6b11516.

[118] J. Thirman and M. Head-Gordon. "Electrostatic domination of the effect of electron correlation in intermolecular interactions". In: *J. Phys. Chem. Lett.* 5 (2014), pp. 1380–1385. DOI: 10.1021/jz500165u.

[119] Jonathan Thirman et al. "Characterizing the interplay of Pauli repulsion, electrostatics, dispersion and charge transfer in halogen bonding with energy decomposition analysis". In: *Phys. Chem. Chem. Phys.* 20.2 (2018), pp. 905–915.

[120] Ehud Tsivion et al. "A computational study of CH4storage in porous framework materials with metalated linkers: Connecting the atomistic character of CH4binding sites to usable capacity". In: *Chemical Science* 7.7 (2016), pp. 4503–4518. ISSN: 20416539. DOI: 10.1039/c6sc00529b.

[121] L Urban, TH Thompson, and C Ochsenfeld. "A scaled explicitly correlated F12 correction to second-order Møller–Plesset perturbation theory". In: *J. Chem. Phys* 154.4 (2021), p. 044101.

[122] S. P. Veccham et al. "A non-perturbative pairwise-additive analysis of charge transfer contributions to intermolecular interaction energies". In: *Phys. Chem. Chem. Phys.* 23 (2021), pp. 928–943. DOI: 10.1039/D0CP05852A.

[123] F. Weinhold. In: *Computational Methods in Photochemistry*. Ed. by A. G. Kutateladze. Vol. 13. Molecular and Supramolecular Photochemistry. Taylor & Francis, 2005, p. 393.

[124] F. Weinhold, C. R. Landis, and E. D. Glendening. "What is NBO analysis and how is it useful?" In: *Int. Rev. Phys. Chem.* 35 (2016), pp. 399–440. DOI: 10.1080/0144235X.2016.1192262.

[125] Frank Weinhold and Roger A Klein. "What is a hydrogen bond? Resonance covalency in the supramolecular domain". In: *Chem. Educ. Res. Pract.* 15.3 (2014), pp. 276–285.

[126] Qiaozhuo Wu et al. "Enhancement of tetrel bond involving tetrazole-TtR 3 (Tt= C, Si; R= H, F). Promotion of SiR 3 transfer by a triel bond". In: *Physical Chemistry Chemical Physics* 24.42 (2022), pp. 25895–25903.

[127] Soohaeng Yoo et al. "High-level ab initio electronic structure calculations of water clusters (H2O) 16 and (H2O) 17: A new global minimum for (H2O) 16". In: *J. Phys. Chem. Lett.* 1.20 (2010), pp. 3122–3127.

[128] Stefan Zahn, Douglas R MacFarlane, and Ekaterina I Izgorodina. "Assessment of Kohn–Sham density functional theory and Møller–Plesset perturbation theory for ionic liquids". In: *Phys. Chem. Chem. Phys.* 15.32 (2013), pp. 13664–13675.

[129] Tom Ziegler and Arvi Rauk. "A Theoretical Study of the Ethylene-Metal Bond in Complexes between Cu^+, Ag^+, Pt^0, or Pt^{2+} and Ethylene, Based on Hartree-Fock-Slater Transition-State Method". In: *Inorg. Chem.* 18.6 (1979), pp. 1558–1565.

[130] Tom Ziegler and Arvi Rauk. "On the Calculation of Bonding Energies by the Hartree Fock Slater Method". In: *Theor. Chem. Acc.* 46.1 (1977), pp. 1–10.