## Title

Exploring the Influence of Orbitals and Regularization on Molecular Properties Calculated Via Many-body Perturbation Theory



## Author

Wong, Jonathan LC

## Publication Date

2023

Peer reviewed|Thesis/dissertation

# Exploring the Influence of Orbitals and Regularization on Molecular Properties Calculated Via Many-body Perturbation Theory

by

Jonathan LC Wong

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Chemistry

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Martin Head-Gordon, Chair
Professor Eran Rabani
Associate Professor Kranthi K. Mandadapu

Fall 2022

Exploring the Influence of Orbitals and Regularization on Molecular Properties Calculated Via Many-body Perturbation Theory

Abstract

Exploring the Influence of Orbitals and Regularization on Molecular Properties Calculated Via Many-body Perturbation Theory

by

Jonathan LC Wong

Doctor of Philosophy in Chemistry

University of California, Berkeley

Professor Martin Head-Gordon, Chair

Molecular properties act as the interface between electronic structure theory and experimental chemistry. Several recently developed electronic structure methods have proved to produce accurate electronic energies in an efficient manner, but their applications to molecular properties calculations remain under-developed. This thesis explores the application of accurate and efficient electronic structure methods for energy calculations to the calculation of molecular properties. In Chapter 2, a method independent, fully numerical finite difference approach to calculating molecular magnetic properties is presented. Two recently proposed methods, regularized second-order Møller-Plesset perturbation theory ($\kappa$-MP2) and variants of the third-order Møller-Plesset perturbation theory (MP3) are used to calculate nuclear magnetic resonance chemical shifts. The scaled MP3 method, MP2.X, is found to be the best performing among the methods inspected: it produces lower errors for heavy nuclei than coupled cluster with singles and doubles (CCSD), a method with the same cost scaling and higher prefactor. In particular, MP2.5 produced an impressive two-fold improvement over CCSD for $^{13}C$ shieldings, while MP2.6 produced an extraordinary six-fold improvement over CCSD for $^{15}N$ shieldings. The simpler and cheaper $\kappa$-MP2, with weak regularization, had smaller errors for $^{13}$C shieldings than CCSD as well. In Chapter 3, analytic nuclear gradients for the regularized orbital optimized second-order Møller-Plesset perturbation theory ($\kappa$-OOMP2) is developed. It was then applied to several interesting chemical problems, including radical geometry and vibrational frequencies, chalcogen bonding, and singlet-triplet gaps. $\kappa$-OOMP2 was able to resolve most of the artifacts stemming from HF orbitals. Compared to the regularization strength suitable for thermochemistry, weaker regularization is more suitable for the problems examined in this work. Finally, Chapter 4 presents a mechanistic study of solvent oxygen incorporation into the products of CO reduction on Cu using a recently developed density functional $\omega$B97M-V. This work found that isotopic scrambling of carbonyl-containing intermediate reaction products that are reduced further to oxygenated products provides a simple explanation to the observed solvent oxygen incorporation.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

The field of electronic structure theory is considered a sub-field of theoretical chemistry, but it really lies at the intersection of chemistry, physics, computer science, and mathematics. Its goal is to model the quantum behaviour of electrons in atoms, molecules, and solid-state materials, and to better understand how their interactions give rise to experimentally observable phenomena such as chemical reactivity, spectral measurements, and various physical properties. Despite the fact that the exact laws governing these interactions have been known for over a century, applying these laws to interesting chemical systems is not tractable in practice. With practicality in mind, electronic structure theorists develop physically motivated approximations to these exact laws, exploit techniques from mathematics to develop novel algorithms, and harness the latest computer architecture to bring these algorithms to life, always trying to balance computational cost and accuracy. The result of the last few decades of effort is a slew of methods and algorithms that, if used wisely, can provide valuable insights into the underlying mechanisms of chemical reactions and other chemical processes. These insights aid experimentalists in designing new experiments and interpreting the results of existing experiments in a more nuanced and accurate way. As such, theorists and experimentalists enjoy a complementary relationship, where theorists provide *ab initio* insight, and experimentalists validate theoretical models with measurements.

In this thesis, we will first introduce some common electronic structure methods and discuss a general framework for calculating molecular properties. Then, we will discuss new developments in methods for calculating molecular properties. Lastly, we will discuss our work on elucidating the reactivity and mechanisms of different chemical systems using electronic structure theory.

## 1.1 Quantum Chemistry Methods

### Schrödinger Equation

Proposed by Erwin Schrödinger in 1926[1], the time dependent Schrödinger equation:

$$i\hbar\frac{\partial}{\partial t}|\Psi\rangle = \hat{H}|\Psi\rangle \tag{1.1}$$

describes the time evolution of the wavefunction $|\Psi\rangle$ governed by the Hamiltonian $\hat{H}$. The wavefunction completely defines the quantum mechanical state of the system being studied. The time independent Schrödinger equation:

$$\hat{H}|\Psi\rangle = E|\Psi\rangle \tag{1.2}$$

describes the energy $E$ and structure of the wavefunction $|\Psi\rangle$ under the influence of the Hamiltonian $\hat{H}$.

Given eigenstates $|\Psi\rangle$ of $\hat{H}$, $E$ are the corresponding eigenvalues, which are the energies of eigenstate $|\Psi\rangle$.

## Hamiltonian

The Hamiltonian for $N$ nuclei and $n$ electrons in atomic units is given by

$$\begin{aligned}\hat{H} &= \hat{T}_e + \hat{T}_n + \hat{V}_{ne} + \hat{V}_{ee} + \hat{V}_{nn} \\ &= -\sum_i^n \frac{1}{2}\boldsymbol{\nabla}_i^2 - \sum_K^N \frac{1}{2m_K}\boldsymbol{\nabla}_K^2 - \sum_i^n \sum_K^N \frac{Z_K}{|\mathbf{r}_i - \mathbf{R}_K|} + \sum_{i<j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{K<L} \frac{Z_K Z_L}{|\mathbf{R}_K - \mathbf{R}_L|}\end{aligned} \tag{1.3}$$

where $\hat{T}_e, \hat{T}_n, \hat{V}_{ne}, \hat{V}_{ee}, \hat{V}_{nn}$ are the electronic kinetic, nuclear kinetic, nuclear-electron potential, electron-electron potential, and nuclear-nuclear potential operators respectively. $Z_K$ is the charge of the $K^{\text{th}}$ nuclei. The corresponding wavefunction $\Psi(\mathbf{r}, \mathbf{R})$ then is a function of all electronic and nuclear coordinates, where $\boldsymbol{r} = \{\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3, ...\}$ and $\boldsymbol{R} = \{\boldsymbol{R}_1, \boldsymbol{R}_2, \boldsymbol{R}_3, ...\}$ are the sets of coordinates for the $n$ electrons and $N$ nuclei. The $3^{\text{rd}}$ postulate of quantum mechanics says that the squared modulus of the wavefunction $|\Psi(\mathbf{r}, \mathbf{R})|^2$ is the probability density of finding electrons at coordinates $\mathbf{r}$ and nuclei at $\mathbf{R}$.

The time independent Schrödinger equation is thus a partial differential equation (PDE) of $3n + 3N$ variables. Solving such a complex PDE is a non-trivial problem. Quantum chemistry circumvents this problem with a few approximations, which we will introduce below.

### Born-Oppenheimer

The Born-Oppenheimer (BO) approximation is based on the idea that the much heavier nuclei move much slower than electrons, and thus we can separate their degrees of freedom. The full wavefunction is factorized into two parts:

$$\Psi(\mathbf{r}, \mathbf{R}) = \Psi_e(\mathbf{r}; \mathbf{R})\Phi_N(\mathbf{R}) \tag{1.4}$$

Here $\Psi_e(\mathbf{r}; \mathbf{R})$ is the electronic wavefunction at a particular configuration of nuclear coordinate $\mathbf{R}$; $\Phi_N(\mathbf{R})$ is the nuclear wavefunction. The electronic Hamiltonian is

$$\hat{H}_e = \hat{T}_e + \hat{V}_{ne} + \hat{V}_{ee} \tag{1.5}$$

And the corresponding electronic Schrödinger equation is

$$(\hat{T}_e + \hat{V}_{ne} + \hat{V}_{ee})|\Psi_e(\mathbf{r}; \mathbf{R})\rangle = E_e(\mathbf{R})|\Psi_e(\mathbf{r}; \mathbf{R})\rangle \tag{1.6}$$

which is now a $3n$ dimensional PDE. Solving this equation accurately and efficiently is the central concern of the field of electronic structure theory. Armed with the electronic energy $E_e$ for desired nuclear configurations, the BO potential energy surface (PES) on which nuclei move is given by

$$V(\mathbf{R}) = E_e(\mathbf{R}) + V_{NN}(\mathbf{R}) \tag{1.7}$$

where the BO nuclear Schrödinger equation is then

$$[T_N(\mathbf{R}) + V(\mathbf{R})]\Phi_N(\mathbf{R}) = E_N\Phi_N(\mathbf{R}) \tag{1.8}$$

The nuclear Schrödinger equation is often not solved directly, but approximated in electronic structure theory by modeling the nuclei as a set of coupled quantum harmonic oscillators. In other words, we can approximate the PES near local extrema $\boldsymbol{R}_0$ by

$$V_{HO}(\boldsymbol{R}) = E_N|_{\boldsymbol{R_0}} + \frac{1}{2}(\boldsymbol{R} - \boldsymbol{R_0})^T\mathbf{K}(\boldsymbol{R} - \boldsymbol{R_0}) \tag{1.9}$$

where $\mathbf{K}$ is the force constants matrix. The nuclear Hessian at PES extrema is calculated, followed by normal mode analysis. From the normal modes, zero-point vibrational energies (ZPE) are computed and used to approximate nuclear effects at zero-temperature.

## Hartree Fock

The variational principal states that for Hermitian $\hat{H}$, minimizing the expectation value of the Hamiltonian with a trial wavefunction $\Psi$ parameterized by some variational parameter $C$ gives an upper bound to the exact ground state energy $E_0$:

$$\min_C \frac{\langle\tilde{\Psi}(C)|\hat{H}|\tilde{\Psi}(C)\rangle}{\langle\tilde{\Psi}(C)|\tilde{\Psi}(C)\rangle} \geq E_0 \tag{1.10}$$

This gives an approximate solution to the electronic Hamiltonian.

The mathematical form of the electronic wavefunction can be deduced from the fermionic nature of electrons. Electrons are spin $\frac{1}{2}$ fermions, so the electronic wavefunction should be antisymmetric with respect to two particle permutations. Since electrons are indistinguishable, the two particle permutation operator $\hat{P}_{ij}$ commutes with the electronic Hamiltonian $\hat{H}$. Therefore an eigenfunction of $\hat{P}_{ij}$ will also be an eigenfunction of $\hat{H}$. The determinant of a square matrix has the convenient property of being antisymmetric with respect to exchanging two columns. The Slater Determinant, with rows indexing one-electron spin orbitals $\chi_i$ and columns indexing particle coordinates $\mathbf{x}_j$, is therefore a good trial wavefunction:

$$\Psi_S(\mathbf{x}_1, \mathbf{x}_2, ...\mathbf{x}_n) = \frac{1}{\sqrt{n}}\begin{vmatrix} \chi_1(\mathbf{x}_1) & \chi_1(\mathbf{x}_2) & \cdots & \chi_1(\mathbf{x}_n) \\ \chi_2(\mathbf{x}_1) & \chi_2(\mathbf{x}_2) & \cdots & \chi_2(\mathbf{x}_n) \\ \vdots & \vdots & \vdots & \ddots \\ \chi_n(\mathbf{x}_1) & \chi_n(\mathbf{x}_2) & \cdots & \chi_n(\mathbf{x}_n) \end{vmatrix} \tag{1.11}$$

The spin orbitals are products of spatial and spin parts:

$$\chi_i(\mathbf{x}_j) = \psi_i(\mathbf{r}_j)\sigma_i(\omega_j) \tag{1.12}$$

where $\psi_j$ is the spatial orbital with spatial coordinate $r_j$; $\sigma_j$ is the spin component with spin coordinate $\omega_j$. $\sigma$ is usually chosen to be $\alpha$ and $\beta$, the two eigenvectors of the spin operator along the z axis, $\hat{s}_z$.

In practice, the spatial orbital $\psi$ is expanded as a linear combination of atomic orbitals (LCAO):

$$\psi_i(\mathbf{r}) = \sum_{\mu}^{M} C_{\mu i}\phi_\mu(\mathbf{r}) \tag{1.13}$$

where $\phi$ is an atomic orbital, $\mu$ is the AO index, $M$ is the number of AO's, and $\mathbf{C}$ is the molecular orbital (MO) coefficient matrix. Thus we have specified the variational space by the set of AO's, and $\mathbf{C}$ represents a matrix of variational parameters. Considering both $\alpha$ and $\beta$ spin, this gives $n$ occupied orbitals and $2M - n$ virtual (unoccupied) orbitals. The Aufbau principle says that the $n$ lowest orbitals will be occupied in a ground state wavefunction. We note that the AO's are in general not orthonormal. This will have to be accounted for with the overlap matrix $\boldsymbol{S}$ given by

$$S_{\mu\nu} = \langle\phi_\mu|\phi_\nu\rangle \tag{1.14}$$

Using the Slater determinant as trial wavefunction and variationally optimizing coefficients $\mathbf{C}$ gives us the Hartree-Fock (HF) approximation, given by

$$E_{\text{HF}} = \min_{\mathbf{C}}\langle\Psi_S|\hat{H}|\Psi_S\rangle \tag{1.15}$$

For convenience, we will rewrite the electronic Hamiltonian Eq 1.6 into two parts: a one electron part $\hat{O}_1$ and a two electron part $\hat{O}_2$:

$$\hat{O}_1 = \sum_{i}^{n}\hat{h}(\mathbf{r}_i) = -\sum_{i}^{n}\frac{1}{2}\boldsymbol{\nabla}_i^2 - \sum_{i}^{n}\sum_{K}^{N}\frac{Z_K}{|\mathbf{r}_i - \mathbf{R}_K|} \tag{1.16}$$

$$\hat{O}_2 = \sum_{i<j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{1.17}$$

$\hat{h}$ is often referred to as the core Hamiltonian. The Hartree-Fock ground state energy written in Dirac bra-ket notation is

$$E_{HF} = \sum_{i}^{n}\langle i|\hat{h}|i\rangle + \frac{1}{2}\sum_{i}^{n}\sum_{j}^{n}\langle ij||ij\rangle \tag{1.18}$$

where

$$\langle i|\hat{h}|i\rangle = \int \chi_i^*(\mathbf{x}_1)h(\mathbf{r}_1)\chi_i(\mathbf{x}_1)d(\mathbf{x}_1) \tag{1.19}$$

$$\langle ij||ij\rangle = \langle ij|ij\rangle - \langle ij|ji\rangle \tag{1.20}$$

$$\langle pq|rs\rangle = \int\int \chi_p^*(\mathbf{x}_1)\chi_q^*(\mathbf{x}_2)\frac{1}{r_{12}}\chi_r(\mathbf{x}_1)\chi_s(\mathbf{x}_2)d\mathbf{x}_1 d\mathbf{x}_2 \tag{1.21}$$

The first part of Eq 1.20 is the classical Coulomb interaction between two charge densities that arise from $\chi_i$ and $\chi_j$; the second part is the purely quantum exchange interaction, due to the antisymmetric property of the wavefunction. We minimize Hartree-Fock energy Eq 1.18 with respect to spin orbitals $\chi$'s by the Lagrange multiplier method. The Lagrangian is

$$\mathcal{L}[\chi_i] = E_{HF}[\chi_i] - \sum_{ij}\epsilon_{ij}(\langle\chi_i|\chi_j\rangle - \delta_{ij}) \tag{1.22}$$

where the constraint is spin orbitals $\chi$ being orthogonal.

Setting the variance of $\mathcal{L}$ to 0 gives the equation:

$$\left(\hat{h}(\mathbf{r}_1) + \sum_j^n \int \chi_j^*(\mathbf{x}_2)\frac{1}{r_{12}}(1-\hat{P}_{12})\chi_j(\mathbf{x}_2)\right)|\chi_i\rangle = \sum_j^n \epsilon_{ij}|\chi_i\rangle \tag{1.23}$$

We define the fock operator $\hat{f}$ as

$$\hat{f} = \hat{h}(\mathbf{r}_1) + \sum_j^n \int \chi_j^*(\mathbf{x}_2)\frac{1}{r_{12}}(1-\hat{P}_{12})\chi_j(\mathbf{x}_2) \tag{1.24}$$

The Hartree-Fock equation is then formulated as an eigenvalue problem:

$$\hat{f}|\chi_i\rangle = \epsilon_i|\chi_i\rangle \quad i = 1, 2, ..., n \tag{1.25}$$

Fully expanding with the LCAO approximation (Eq 1.13) gives the Roothaan equations[2]:

$$\mathbf{FC} = \mathbf{SC}\epsilon \tag{1.26}$$

$$\mathbf{F}_{\mu\nu} = \langle\phi_\mu|\hat{f}|\phi_\nu\rangle \tag{1.27}$$

Since **F** depends on **C**, solving the Roothaan equations is a non-linear optimization problem, and will be handled by the self-consistent field (SCF) method, which involves iterative diagonalization and updates until some convergence criteria is satisfied.

## Correlated Methods

Revisiting the fock operator (Eq 1.24), we see that electron $\mathbf{x}_1$ experiences all other electrons through an averaged potential. The electrons are not correlated aside from the exchange interaction, which

arises from the antisymmetric wavefunction. Hartree-Fock is therefore a minimally-correlated method. We define the correlation energy as the difference between the HF energy and the exact ground state energy:

$$E_{\text{corr}} = E_0 - E_{\text{HF}} \tag{1.28}$$

Since HF is variational, $E_{\text{corr}}$ is negative semi-definite. Correlation is often separated into static correlation and dynamic correlation. They are delineated as follows: when the wavefunction has a large overlap with the Hartree-Fock determinant, correlation energy will mostly be dynamic, described by instantaneous electron excitations from occupied orbitals to virtual orbitals; when the ground state wavefunction is not dominated by a single determinant, rather, more than one determinant become nearly degenerate, then static correlation becomes important. Static correlation dominated systems are referred to as multi-reference (MR).

Before discussing correlated methods, we will briefly introduce second quantization notation. Starting with a vacuum state with no electrons $|0\rangle$, the creation operator $\hat{a}_p^\dagger$ creates electron $p$, and the annihilation operator $\hat{a}_q$ destroys electron $q$.

$$\hat{a}_p^\dagger|0\rangle = |p\rangle \tag{1.29}$$

$$\hat{a}_q|q\rangle = |0\rangle \tag{1.30}$$

A pair of creation and annihilation operators acting on some state creates an excitation:

$$\hat{a}_p^\dagger \hat{a}_q|\phi\rangle = |\phi_q^p\rangle \tag{1.31}$$

where an electron is excited from orbital $q$ to $p$. Now we can define single a excitation operator $\hat{T}_1$:

$$\hat{T}_1 = \sum_{i,a} \hat{a}_a^\dagger \hat{a}_i t_i^a \tag{1.32}$$

and a double excitation operator $\hat{T}_2$:

$$\hat{T}_2 = \sum_{i<j,a<b} \hat{a}_b^\dagger \hat{a}_a^\dagger \hat{a}_i \hat{a}_j t_{ij}^{ab} \tag{1.33}$$

where **t** are the t-amplitudes, $i, j$ are occupied orbital indices, and $a, b$ are virtual orbital indices. $\hat{T}_1$ acting on some reference state gives all singly excited states, $\hat{T}_2$ acting on some reference state gives all doubly excited states etc.

The action of all $n$ excitation operators acting on ground state Slater determinant $|\Phi\rangle$ spans the full Hilbert space associated with orbital expansion.

## Configuration Interaction

Armed with the set of determinants spanning the Hilbert space, a straightforward way to calculate correlation is to use a trial wavefunction that spans the entire space, in other words, a linear

combination of all determinants. In this context, the Slater determinants are sometimes referred to configurations, thus the name full configuration interaction (FCI). We define the FCI wavefunction:

$$|\Psi_{\text{FCI}}\rangle = (\hat{T}_0 + \hat{T}_1 + \hat{T}_2 + ... + \hat{T}_n)|\Phi_0\rangle \tag{1.34}$$

$$= t_0|\Phi_0\rangle + \sum_{i,a} t_i^a|\Phi_i^a\rangle + \sum_{i<j,a<b} t_{ij}^{ab}|\Phi_{ij}^{ab}\rangle + ... \tag{1.35}$$

Because the wavefunction spans the entire Fock space, FCI energy is exact within the basis set used. In addition, since all configurations are considered, FCI is invariant with respect to the reference wavefunction $|\Phi_0\rangle$. The total number of configurations is $\binom{2M}{n}$, which scales factorially with system size. Factorial scaling renders FCI intractable for any system with more than a few atoms.

One approach to make CI tractable is by truncating excitations to a certain order. For example, summing up to $\hat{T}_2$ gives CISD. This approach has a few drawbacks. First, since the wavefunction no longer spans the entire Hilbert space, CI energies depend on the reference wavefunction. This is not a huge problem when the HF determinant has a large overlap with the exact ground state. Second, truncated CI energies are size inconsistent. Size consistency is defined as

$$E_{\text{A,B}(R_{AB}\to\infty)} = E_{\text{A}} + E_{\text{B}} \tag{1.36}$$

The non-interacting supersystem has the same energy as the sum of the energies of the subsystems. This can be understood with a CISD example. The individual subsystem CISD energies includes up to double excitations, so the sum of their energies would include some effects from quadruple excitations. This would be missing in the supersystem energy since only double excitations are included. Without size consistency, truncated CI is not suitable for treating problems such as chemical reaction energies.

## Coupled Cluster Theory

Another approach to calculate correlation is the family of coupled cluster (CC) methods. First we introduce the CC exponential ansatz:

$$|\Psi_{CC}\rangle = e^{\hat{T}_1+\hat{T}_2+\hat{T}_3+\cdots}|\Phi_0\rangle \tag{1.37}$$

Since this wavefunction is equivalent to the FCI wavefunction when all excitation operators are in the exponent, solving for the energy variationally costs combinatorial time as well.

To calculate truncated CC energies in polynomial time, we first consider the similarity transformed Hamiltonian:

$$\hat{H}_{CC} = e^{-\hat{T}}\hat{H}e^{\hat{T}} \tag{1.38}$$

We left project $\hat{H}_{CC}|\Phi_0\rangle$ with all ground and excited wavefunctions:

$$\langle\Phi_0|\hat{H}_{CC}|\Phi_0\rangle = E_{CC} \tag{1.39}$$
$$\langle\Phi_i^a|\hat{H}_{CC}|\Phi_0\rangle = 0 \tag{1.40}$$
$$\langle\Phi_{ij}^{ab}|\hat{H}_{CC}|\Phi_0\rangle = 0 \tag{1.41}$$
$$...$$

We then iteratively solve for all t-amplitudes and CC energy. CCSD[3, 4] includes only single and double excitations, and it costs $O(N^6)$ per iteration; CCSDT[5], which includes triple excitations as well, costs $O(N^8)$ per iteration. CCSD(T)[6] attempts to reduce this high cost by approximating triples contribution with a perturbative correction that costs $O(N^7)$ for the non-iterative step. With high accuracy for its cost, many consider this method the gold standard of electronic structure calculations.

## Møller-Plesset Perturbation Theory

Another way to calculate correlation energy is by perturbation theory. Møller-Plesset (MP) Perturbation Theory[7] can be considered an extension of Rayleigh-Schrödinger (RS) perturbation theory. Starting from RS formalism, we split the Hamiltonian $\hat{H}$ into an unperturbed part $\hat{H}_0$ and the perturbation $\hat{V}$,

$$\hat{H} = \hat{H}_0 + \lambda\hat{V} \tag{1.42}$$

where $\lambda$ is the order of perturbation. The eigenfunctions and eigenvalues of $\hat{H}_0$ are known: $\hat{H}_0|\psi_n^{(0)}\rangle = E_n^{(0)}|\psi_n^{(0)}\rangle$, $n = i, \{k\}$. Subscript $i$ denotes ground state and $k$ denotes excited states. We then expand the wavefunction and energy as power series in $\lambda$:

$$|\psi_i\rangle = |\psi_i^{(0)}\rangle + \lambda|\psi_i^{(1)}\rangle + \lambda^2|\psi_i^{(2)}\rangle + \lambda^3|\psi_i^{(3)}\rangle + ... \tag{1.43}$$
$$E_i = E_i^{(0)} + \lambda E_i^{(1)} + \lambda^2 E_i^{(2)} + \lambda^3 E_i^{(3)} + ... \tag{1.44}$$

Substituting the expanded wavefunction and energy into the Schrödinger equation and collecting by powers of $\lambda$,

$$\hat{H}_0|\psi_i^{(0)}\rangle = E_i^{(0)}|\psi_i^{(0)}\rangle \tag{1.45}$$
$$\hat{H}_0|\psi_i^{(1)}\rangle + \hat{V}|\psi_i^{(0)}\rangle = E_i^{(0)}|\psi_i^{(1)}\rangle + E_i^{(1)}|\psi_i^{(0)}\rangle \tag{1.46}$$
$$\hat{H}_0|\psi_i^{(2)}\rangle + \hat{V}|\psi_i^{(1)}\rangle = E_i^{(0)}|\psi_i^{(2)}\rangle + E_i^{(1)}|\psi_i^{(1)}\rangle + E_i^{(2)}|\psi_i^{(0)}\rangle \tag{1.47}$$
$$\hat{H}_0|\psi_i^{(3)}\rangle + \hat{V}|\psi_i^{(2)}\rangle = E_i^{(0)}|\psi_i^{(3)}\rangle + E_i^{(1)}|\psi_i^{(2)}\rangle + E_i^{(2)}|\psi_i^{(1)}\rangle + E_i^{(3)}|\psi_i^{(0)}\rangle \tag{1.48}$$
$$...$$

Using the convention of intermediate normalization:

$$\langle \psi_i^{(0)} | \psi_i \rangle = 1 \tag{1.49}$$

we left project Eq 1.45 and 1.46 with $\langle \psi_i^{(0)}|$ to obtain

$$\begin{aligned} E_i^{(0)} &= \langle \psi_i^{(0)} | \hat{H}_0 | \psi_i^{(0)} \rangle \\ E_i^{(1)} &= \langle \psi_i^{(0)} | \hat{V} | \psi_i^{(0)} \rangle \end{aligned} \tag{1.50}$$

The first order wavefunction $\psi_i^{(1)}$ is given by a linear combination of excited states of $\psi^{(0)}$:

$$|\psi_i^{(1)}\rangle = -\sum_{k \neq i} \frac{\langle \psi_k^{(0)} | \hat{V} | \psi_i^{(0)} \rangle}{E_k^{(0)} - E_i^{(0)}} |\psi_k^{(0)}\rangle \tag{1.51}$$

Then we left project Eq 1.47 with $\langle \psi_i^{(0)}|$ to obtain

$$\begin{aligned} E_i^{(2)} &= \langle \psi_i^{(0)} | \hat{V} | \psi_i^{(1)} \rangle \\ &= -\sum_{k \neq i} \frac{\langle \psi_i^{(0)} | \hat{V} | \psi_k^{(0)} \rangle \langle \psi_k^{(0)} | \hat{V} | \psi_i^{(0)} \rangle}{E_k^{(0)} - E_i^{(0)}} \end{aligned} \tag{1.52}$$

Similarly, the third order correction is

$$E_i^{(3)} = \langle \psi_i^{(0)} | \hat{V} | \psi_i^{(2)} \rangle \tag{1.53}$$

The above form would require the second order wavefunction. Wigner's $(2n+1)$ rule [8] provides a work around. It states that the $2n^{th}$ and $(2n+1)^{th}$ order energy can be evaluated with the $n^{th}$ order and lower wavefunctions:

$$E_i^{(2n+1)} = \langle \psi_i^{(n)} | \hat{V} | \psi_i^{(n)} \rangle - \sum_{s,p=1}^{n} E_i^{(2n+1-s-p)} \langle \psi_i^{(s)} | \psi_i^{(p)} \rangle \tag{1.54}$$

As such, we have

$$E_i^{(3)} = \langle \psi_i^{(1)} | \hat{V} | \psi_i^{(1)} \rangle - E_i^{(1)} \langle \psi_i^{(1)} | \psi_i^{(1)} \rangle \tag{1.55}$$

$$E_i^{(3)} = \sum_{l \neq i} \sum_{m \neq i} \frac{\langle \psi_i^{(0)} | \hat{V} | \psi_l^{(0)} \rangle \langle \psi_l^{(0)} | \hat{V} | \psi_m^{(0)} \rangle \langle \psi_m^{(0)} | \hat{V} | \psi_i^{(0)} \rangle}{(E_l^{(0)} - E_i^{(0)})(E_m^{(0)} - E_i^{(0)})} - E_i^{(1)} \sum_{l \neq i} \frac{|\langle \psi_i^{(0)} | \hat{V} | \psi_l^{(0)} \rangle|^2}{(E_l^{(0)} - E_i^{(0)})^2} \tag{1.56}$$

Moving on to Møller-Plesset perturbation theory, we use the Fock operator $\hat{F}$ as $\hat{H}_0$; $\hat{V}$ then equals $\hat{H} - \hat{F}$.

Using Hartree-Fock canonical orbitals, the occupied-occupied (oo) and virtual-vitual (vv) blocks of **F** are diagonal, and the occupied-virtual (ov) blocks are zero,

$$\begin{aligned} E_{\text{MP0}} &= \langle \phi_0 | \hat{F} | \phi_0 \rangle \\ &= \sum_i \epsilon_i \end{aligned} \tag{1.57}$$

where $\epsilon_i$ is the orbital energy of orbital $i$

$$\begin{aligned} E_{\text{MP1}} &= \langle \phi_0 | \hat{H} - \hat{F} | \phi_0 \rangle \\ &= E_{HF} - \sum_i \epsilon_i \end{aligned} \tag{1.58}$$

For the first order wavefunction, we observe that in Eq 1.51, denominator $\langle \psi_k^{(0)} | \hat{V} | \psi_i^{(0)} \rangle$ couples the $0^{th}$ order ground state with $0^{th}$ order excited states through $\hat{V}$. Since $\hat{V}$ only contains up to 2 electron operators, and the ov blocks of **F** are zero for canonical orbitals, only double excitations survive.

$$|\psi^{(1)}\rangle = \sum_{i<j,a<b} t_{ij}^{ab} |\phi_{ij}^{ab}\rangle \tag{1.59}$$

$$t_{ij}^{ab} = -\frac{\langle ab||ij\rangle}{\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j} = -\frac{\langle ab||ij\rangle}{\Delta_{ij}^{ab}} \tag{1.60}$$

where we have first order t-amplitudes $t_{ij}^{ab}$, and we used $\Delta_{ij}^{ab}$ as shorthand for $\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j$. The second order energy is then

$$\begin{aligned} E_{\text{MP2}} &= -\sum_{i<j,a<b} \frac{\langle ij||ab\rangle \langle ab||ij\rangle}{\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j} \\ &= -\frac{1}{4} \sum_{ijab} \frac{|\langle ij||ab\rangle|^2}{\Delta_{ij}^{ab}} \end{aligned} \tag{1.61}$$

where the factor of $\frac{1}{4}$ arises from unrestricting the sum indices. The MP2 energy exhibits divergence when $\Delta_{ij}^{ab}$ approaches 0, which leads to a catastrophic over estimation of correlation energy. One remedy for this problem is by introducing an energy based regularizer. Several have been proposed [9–11], one example being the $\kappa$ regularizer. The $\kappa$-regularized t-amplitudes $\bar{\mathbf{t}}$ are

$$\bar{t}_{ij}^{ab} = -\frac{\langle ab||ij\rangle}{\Delta_{ij}^{ab}} \left(1 - e^{-\kappa \Delta_{ij}^{ab}}\right) \tag{1.62}$$

The $\kappa$-regularized MP2 energy is given by

$$E_{\kappa\text{-MP2}} = -\frac{1}{4} \sum_{ijab} \frac{|\langle ij||ab\rangle|^2}{\Delta_{ij}^{ab}} \left(1 - e^{-\kappa \Delta_{ij}^{ab}}\right)^2 \tag{1.63}$$

As $\Delta_{ij}^{ab} \to 0$, $\bar{t}_{ij}^{ab} \to \kappa\langle ij||ab\rangle$, and the correlation contribution goes to zero.

For the MP3 energy, we start from Eq 1.55 and note that the two excited states can differ by 2 indices: 2 occupied, 2 virtual, 1 occupied 1 virtual. This gives

$$E_{\text{MP3}} = \frac{1}{8}\sum_{ijabcd}(t_{ij}^{ab})^*\langle ab||cd\rangle t_{ij}^{cd} + \frac{1}{8}\sum_{ijklab}(t_{ij}^{ab})^*\langle kl||ij\rangle t_{kl}^{ab} - \sum_{ijkabc}(t_{ij}^{ab})^*\langle kb||ic\rangle t_{kj}^{ac} \tag{1.64}$$

For open-shell systems, spin symmetry broken HF wavefunction may not be a good reference for perturbation theory. This leads to subpar MP2 results[12]. We can obtain a better reference wavefunction by optimizing the orbitals. For this purpose, we first introduce the Hylleras functional[13]:

$$J_H = \langle\psi_0|\hat{V}|\psi_1\rangle + \langle\psi_1|\hat{V}|\psi_0\rangle + \langle\psi_1|\hat{H}_0 - E_{\text{MP0}}|\psi_1\rangle \tag{1.65}$$

The MP2 energy can be thought of as the minimum of the Hylleras functional with respect to t-amplitudes:

$$E_{\text{MP2}} = \arg\min_{\mathbf{t}}\{J_H\} \tag{1.66}$$

By minimizing with respect to orbital rotation parameter $\mathbf{\Theta}$ as well, we obtain the OOMP2 method. The corresponding Lagrangian is

$$\mathscr{L}\{\mathbf{t}, \mathbf{\Theta}\} = E_{HF}[\mathbf{\Theta}] + J_H[\boldsymbol{t}, \mathbf{\Theta}] \tag{1.67}$$

After orbital rotation, Brillouin's theorem no longer holds. The first order wave function should now contain single excitation contributions. However, It has been shown that inclusion of singles does not have a big effect on the accuracy of OOMP2[14]. In addition, orbital rotation should account for some of the effects of the single contributions. Therefore we will discard explicit singles for OOMP2.

OOMP2 inherits the divergence problem of MP2: small $\Delta_{ij}^{ab}$ can lead to catastrophic error during optimization. Regularized OOMP2 was introduced to tackle this issue. $\kappa$-OOMP2[9] in particular had shown great success in spin symmetry restoration and thermochemistry prediction. This method will be used extensively in the following chapters.

## Density Functional Theory

The above approaches, usually named wavefunction methods, all center around the wavefunction. Correlation effects are treated by explicitly considering electronic excitations upon the reference HF determinant, either perturbatively or incorporated into the ansatz. This improvement comes at a cost: the least costly method, MP2, still scales $O(N^5)$, compared to HF's $O(N^3)$ scaling; the gold standard CCSD(T) scales $O(N^7)$. An alternative approach that scales similar to HF is density functional theory (DFT). Given its amiable cost and relative accuracy, DFT has become the workhorse of quantum chemistry, and is arguably the most successful achievement of the

field. Below we will introduce the basics of DFT and some successful approaches to functional approximations.

The basis of DFT is two theorems proposed by Hohenberg and Kohn (HK) in 1964[15]. The first HK theorem states that the external potential $v_{\text{ext}}$ is determined by the electron density $\rho$, up to an additive constant. Since $\rho$ also determines the number of electrons $N$, $\rho$ then also determines the ground state wavefunction $\Psi$ and thus all other ground state properties. In other words, there is a one to one mapping between electron density $\rho(\boldsymbol{r})$ and $v_{\text{ext}}(\boldsymbol{r})$. Since the number of electrons $N$ and the external potential $v_{\text{ext}}$ completely determines the Hamiltonian and thus the wavefunction $\Psi$, the electron density $\rho$ also determines the wavefunction $\Psi$ and all related properties. For isolated molecular systems, $V_{\text{ext}}$ is simply $V_{en}$.

From this, we can write the ground state energy as a functional of $\rho(\boldsymbol{r})$:

$$\begin{aligned} E[\rho(\boldsymbol{r})] &= \langle \Psi[\rho(\boldsymbol{r})] | \hat{T} + \hat{V}_{ee} + \hat{V}_{ne} | \Psi[\rho(\boldsymbol{r})] \rangle \\ &= T[\rho(\boldsymbol{r})] + V_{ee}[\rho(\boldsymbol{r})] + \int v_{ne}(\boldsymbol{r})\rho(\boldsymbol{r}) d\boldsymbol{r} \\ &= T[\rho(\boldsymbol{r})] + J[\rho(\boldsymbol{r})] + Q[\rho(\boldsymbol{r})] + \int v_{ne}(\boldsymbol{r})\rho(\boldsymbol{r}) d\boldsymbol{r} \\ &= F_{\text{HK}}[\rho(\boldsymbol{r})] + \int v_{ne}(\boldsymbol{r})\rho(\boldsymbol{r}) d\boldsymbol{r} \end{aligned} \tag{1.68}$$

$V_{ee}$ was split into the classical Coulomb $J$ given exactly by

$$J[\rho(\boldsymbol{r})] = \frac{1}{2} \int \frac{\rho(\boldsymbol{r})\rho(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} d\boldsymbol{r} \tag{1.69}$$

and a non-classical term $Q$, which describes the purely quantum exchange and correlation energies. $F_{\text{HK}}$ is the universal functional, so called because it is independent of external potential $v_{\text{ext}}$: it is the same for any system.

The second HK theorem gives the variational principle for the optimum electron density. For some trial density $\tilde{\rho}(\boldsymbol{r})$ that satisfies $\tilde{\rho}(\boldsymbol{r}) \geq 0$ and $\int \tilde{\rho}(\boldsymbol{r}) = N$

$$E[\tilde{\rho}(\boldsymbol{r})] \geq E_0 \tag{1.70}$$

Analogous to Eq 1.10, this ensures that a trial density gives an upper bound to the exact ground state energy, and that we can optimize $\tilde{\rho}$ and thus $E$.

The two HK theorems provides an attractive framework for calculating the exact ground state energy of a system, often refered to as HK-DFT or orbital free DFT. We have transformed our $3n$ dimensional $\Psi$-dependent problem into one only dependent on the 3 dimensional $\rho$. However, there are two significant limitations that have hampered its practical use. First, the second HK theorem requires that the trial density $\tilde{\rho}$ be $v$-representable. This means that $\tilde{\rho}$ must correspond to some ground state $\Psi$ that is associated with $v_{\text{ext}}$. However the conditions that ensure a $v$-representable $\rho$ are unknown. The second limitation is that in the universal functional $F_{HK}$, the exact forms of both the kinetic energy functional and the non-classical $V_{ee}$ term are unknown. Since kinetic energy

often has the same magnitude as the total energy, its accurate description is essential to the success of HK-DFT.

One of the first attempts at approximating the kinetic energy functional, the Thomas-Fermi (TF) model[16, 17], derives the functional form from the uniform electron gas (UEG).

$$T_{TF}[\rho] = \frac{3}{10}(3\pi^2)^{\frac{2}{3}} \int \rho^{\frac{5}{3}}(\mathbf{r}) d\mathbf{r} \tag{1.71}$$

While conceptually elegant, the TF model is not applicable to systems larger than an atom: it does not predict molecular bonds. Even for atoms, its energy prediction is not very accurate. Later attempts to seek more accurate approximations have mostly been limited to systems similar to the UEG[18, 19].

The key approach that transformed DFT from a theoretical intrigue to today's quantum chemistry workhorse came in 1965 by Kohn and Sham(KS). Trading simplicity for accuracy, the KS method introduced orbitals into the problem. It introduces a fictitious system of $n$ non-interacting electrons ($V_{ee} = 0$) represented by $n$ spin orbitals, for which the exact solution to the Schrödinger equation is simply the Slater determinant of the $n$ spin orbitals. The exact kinetic energy of this fictitious system $T_s$ is then

$$T_s = \sum_i^n \langle \psi_i | -\frac{1}{2}\nabla_i^2 | \psi_i \rangle \tag{1.72}$$

The universal functional $F$ is then written as

$$F[\rho(\boldsymbol{r})] = T_s[\rho(\boldsymbol{r})] + J[\rho(\boldsymbol{r})] + E_{xc}[\rho(\boldsymbol{r})] \tag{1.73}$$

where

$$E_{xc}[\rho(\boldsymbol{r})] = T[\rho(\boldsymbol{r})] - T_s[\rho(\boldsymbol{r})] + V_{ee}[\rho(\boldsymbol{r})] - J[\rho(\boldsymbol{r})] \tag{1.74}$$

The exchange-correlation(XC) functional now contains the error in the kinetic energy due to the lack of electron-electron interaction, which crucially seems to be small, and the non-classical part of the electron-electron interaction.

Akin to Eq 1.25, the KS equations are

$$\hat{f}^{KS} | \psi_i \rangle = \epsilon_i | \psi_i \rangle \tag{1.75}$$

where the KS Fock operator $\hat{f}^{KS}$ is defined as

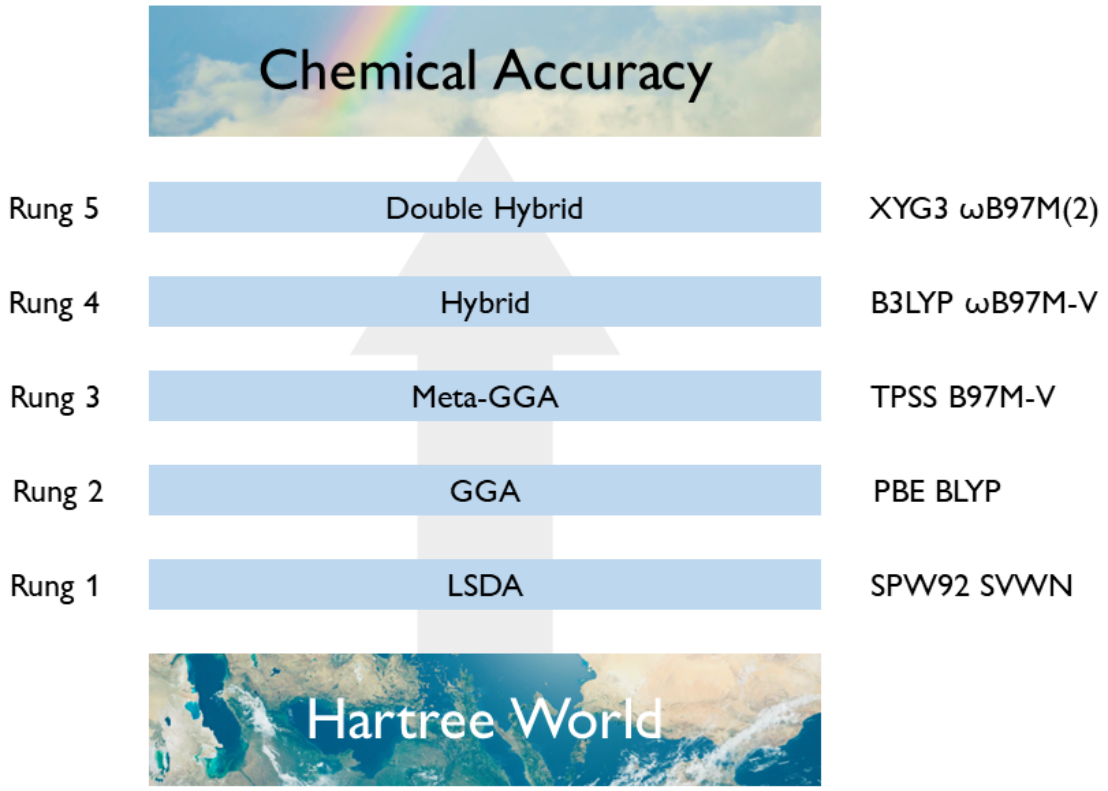
$$\hat{f}^{KS} = \hat{h}(\mathbf{r}) + \int \frac{\rho(\mathbf{r})}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \frac{\delta E_{xc}}{\delta \rho(\mathbf{r})} \tag{1.76}$$

The last term, the functional derivative of the XC functional with respect to electron density, is the XC potential. Eq 1.76 can again be expanded in an AO basis, giving

$$\mathbf{F}^{KS}\mathbf{C} = \mathbf{S}\mathbf{C}\epsilon \tag{1.77}$$

This again can be solved with the SCF procedure.

Figure 1.1: A graphical rendition of Perdew's Jacob's ladder. The name of each rung is shown in the middle, with two example functionals shown on the right




KS-DFT is formally exact. However the exchange-correlation functional is still unknown. The last few decades of DFT development has focused on approximating this functional, and to great success indeed. Since the kinetic energy of the reference non-interacting density is usually similar in magnitude to the exact kinetic energy, the XC functional is smaller in magnitude than the exact kinetic energy functional, and thus easier to approximate. These density functional approximations (DFA) will be summarized briefly in the following.

We must note that DFA's are not systematically improvable: there is no clear path to converge to the exact functional. Nonetheless, one way to navigate the density functional "zoo"[20] is by Perdew's Jacob's ladder[21].

Fig 1.1 depicts a graphical rendition of this ladder. At level 0 is the "Hartree world", where the mean-field potential only includes the classical Coulomb repulsion. At the top of the ladder is the "chemical accuracy heaven". Every step up the ladder increases the complexity of the DFA. A well designed DFA should perform better than the ones below, stepping up the ladder should in theory get us closer to chemical accuracy. Rung 1 are the functionals using the local spin density approximation(LSDA). $E_{xc}$ is a function of $\rho_\alpha$ and $\rho_\beta$ only. Rung 2 functionals are a function of the

gradient of the density, $\nabla\rho$, as well, and are called generalized-gradient approximations(GGAs). Rung 3 functionals in addition use the kinetic energy density of the non-interacting reference system, $\tau = \sum_i |\nabla\psi_i|^2$, and are referred to as meta-GGA functionals. The functionals in the first 3 rungs are collectively called local or semi-local functionals. They suffer from the infamous self interaction error, where the coulomb $\hat{J}$ and XC potentials $v_{xc}$ fail to cancel out for a single orbital. This results in an artificial self repulsion, leading to overly delocalized electron densities. At rung 4, Global hybrid (GH) functionals partly alleviate this error by partially adding Hartree-Fock exchange to the XC functional globally. These GH GGA and GH meta-GGA functionals are in the 4th rung. Also in rung 4 are the range separated hydrid (RSH) functionals, which use a more sophisticated way of adding in HF exchange. The electron-electron coulomb operator is separated into short range and long range parts:

$$\frac{1}{r_{12}} = \frac{\text{erfc}(\omega r_{12})}{r_{12}} + \frac{\text{erf}(\omega r_{12})}{r_{12}} \tag{1.78}$$

The error function and complementary error function smoothly interpolates the strength of the added HF exchange.

The functionals up to rung 4 do not explicitly handle dispersion/van der Waals(VDW) interactions, which are crucial in non-covalently bound systems. One approach is adding empirical dispersion $C_n$ potentials, with the general form:

$$E_{\text{disp}} = -\sum_{K<L}^{N} \left( \frac{C_{6,KL}}{R_{KL}^6} + \frac{C_{8,KL}}{R_{KL}^8} + \frac{C_{10,KL}}{R_{KL}^{10}} + ... \right) f_{\text{damp}}(R_{KL}) \tag{1.79}$$

where $R_{KL}$ are the interatomic distances, and $f_{\text{damp}}$ is a damping function that smoothly switched off the correction at small $R_{KL}$'s. The most widely used example is Grimme's DFT-D family of methods[22–24]. A more ab initio approach is the non-local correlation (NLC) functional, with the general form:

$$E_{NLC} = \int\int \rho(\boldsymbol{r})\Phi(\boldsymbol{r},\boldsymbol{r}')\rho(\boldsymbol{r}')d\boldsymbol{r}d\boldsymbol{r}' \tag{1.80}$$

where $\Phi(\boldsymbol{r},\boldsymbol{r}')$ is the NLC kernel. One notable example is the VV10 Vydrov-Van Voorhis kernel[25], which is employed in the rung 3 meta-GGA B97M-V[26], rung 4 RSH GGA $\omega$B97X-V[27] and RSH meta-GGA $\omega$B97M-V[28] functionals. Large scale benchmarks have shown that these functionals are the top performers in their respective categories[29].

Rung 5 functionals continue the idea of incorporating wavefunction theory components as in rung 4. They are referred to as double hybrid functionals. Using virtual KS orbitals, MP2 or RPA like correlation corrections are incorporated. A few recently developed examples include B2PLYP[30] and $\omega$B97M(2)[31].

## 1.2 Molecular Properties

The preceding section introduced a number of electronic structure methods and how these methods allow the calculation of the energy of isolated molecular systems. In contrast, often times chemically

relevant systems are not isolated. They are interacting with external fields e.g. light-matter interactions or in a condensed phase medium. In the following section, we will discuss how molecules interact with external fields.

## Response to external fields

The energy of a molecule subject to some relatively weak perturbation $\boldsymbol{x}$ changes in response:

$$E(\boldsymbol{x}) = E^{(0)} + E^{(1)}\boldsymbol{x} + \frac{1}{2}\boldsymbol{x}^T E^{(2)}\boldsymbol{x} + ... \tag{1.81}$$

where the new energy is expanded as a Taylor series. The coefficients $E^{(n)}$ of the $n^{th}$ order response in this expansion are characteristic to the molecular system, and thus are called molecular properties. These correspond to measured quantities in an experiment. For time-independent perturbations, the properties are given by differentiation of the perturbed energy:

$$\boldsymbol{E}^{(n)} = \frac{\mathrm{d}^n E}{\mathrm{d}\boldsymbol{x}...\mathrm{d}\boldsymbol{x}}\bigg|_{\boldsymbol{x}=0} \tag{1.82}$$

The electric field is probably the simplest perturbation. The energy of a molecule subject to the uniform electric field $\vec{\mathcal{E}} = \mathcal{E}_x\hat{x} + \mathcal{E}_y\hat{y} + \mathcal{E}_z\hat{z}$, written up to second order is

$$E(\vec{\mathcal{E}}) = E(0) - \vec{\mu}\cdot\vec{\mathcal{E}} - \frac{1}{2}\vec{\mathcal{E}}^{\dagger}\boldsymbol{\alpha}\vec{\mathcal{E}} + ... \tag{1.83}$$

where vector $\vec{\mu}$ is the electric dipole moment and rank 2 tensor $\alpha$ is the electric polarizability.

The energy of a closed-shell molecule subject to a uniform magnetic field $\vec{B}$, and in the presence of non-zero spin nuclei with magnetic moments $\vec{m}_k$, written up to second order is

$$E(\vec{B}, \vec{m}_K) = E(0) - \frac{1}{2}\vec{B}^{\dagger}\boldsymbol{\xi}\vec{B} + \sum_K \vec{B}^{\dagger}\boldsymbol{\sigma}_K\vec{m}_K + \frac{1}{2}\sum_{K\neq L}\vec{m}_K^{\dagger}\boldsymbol{J}_{KL}\vec{m}_L + ... \tag{1.84}$$

where the rank 2 tensor $\boldsymbol{\xi}$ is the magnetizability, the rank 2 tensor $\boldsymbol{\sigma}$ is the nuclear shielding, and rank 2 tensor $\boldsymbol{J}$ is the spin-spin coupling. Odd order terms in $\vec{B}$ vanish for closed shell molecules.

These molecular properties are then the corresponding ordered derivatives of the energy with respect to the corresponding field strengths evaluated at 0 field strength:

$$\vec{\mu} = \frac{\mathrm{d}E}{\mathrm{d}\vec{\mathcal{E}}}\Big|_{\vec{\mathcal{E}}=0} \tag{1.85}$$

$$\boldsymbol{\alpha} = \frac{\mathrm{d}^2E}{\mathrm{d}\vec{\mathcal{E}}^2}\Big|_{\vec{\mathcal{E}}=0} \tag{1.86}$$

$$\boldsymbol{\xi} = \frac{\mathrm{d}^2E}{\mathrm{d}\vec{B}^2}\Big|_{\vec{B}=0} \tag{1.87}$$

$$\boldsymbol{\sigma}_K = \frac{\mathrm{d}^2E}{\mathrm{d}\vec{B}\mathrm{d}\vec{m_K}}\Big|_{\vec{B},\vec{m_K}=0} \tag{1.88}$$

$$\boldsymbol{J}_{KL} = \frac{\mathrm{d}^2E}{\mathrm{d}\vec{m_K}\mathrm{d}\vec{m_L}}\Big|_{\vec{m_K},\vec{m_L}=0} \tag{1.89}$$

We note in passing that higher order static properties such as hyperpolarizability, hypermagnetizability, magnetic circular dichroism, can be expressed in the same way.

## Hamiltonian

A classical charged particle with charge $z$ moving in an electric field $\vec{E}(r,t)$ and magnetic field $\vec{B}(r,t)$ with velocity $\vec{v}$ experiences the Lorentz Force $\vec{F}$, given by

$$\vec{F} = z(\vec{E} + \vec{v} \times \vec{B}) \tag{1.90}$$

The particle's motion is given be Newton's equation of motion:

$$\vec{F} = m\vec{a} \tag{1.91}$$

The electric field and magnetic field satisfy the Maxwell's equations, written in SI based atomic units:

$$\nabla \cdot \vec{E} = 4\pi\rho \tag{1.92}$$

$$\nabla \times \vec{B} - c^{-2}\frac{\partial E}{\partial t} = 4\pi c^{-2}\vec{J} \tag{1.93}$$

$$\nabla \cdot \vec{B} = 0 \tag{1.94}$$

$$\nabla \times \vec{E} + \frac{\partial \vec{B}}{\partial t} = 0 \tag{1.95}$$

where c is the speed of light, $\rho$ is the electric charge density, and $\vec{J}$ is the electric current density. The two homogeneous equations, Eq 1.94 and 1.95, exactly define the relationships between $\vec{E}$ and $\vec{B}$. A solution of these two equations can be written in terms of a scalar potential $\phi$ and the vector potential $\vec{A}$

$$\vec{E}(r,t) = -\nabla\phi(r,t) - \frac{\partial \vec{A}(r,t)}{\partial t} \tag{1.96}$$

$$\vec{B}(r,t) = \nabla \times \vec{A}(r,t) \tag{1.97}$$

It is important to note that $\phi$ and $\vec{A}$ are not unique with respect to gauge transformations that take the form of

$$\phi'(r,t) = \phi(r,t) - \frac{\partial f(r,t)}{\partial t} \tag{1.98}$$

$$\vec{A}'(r,t) = \vec{A}(r,t) + \nabla f(r,t) \tag{1.99}$$

where $f$ is the gauge function. The transformed potentials $\phi$' and $\vec{A}$' correspond to the same physical $\vec{E}$ and $\vec{B}$ fields. There are a few common ways to specify $f$. We will follow the Coulomb gauge, where the chosen gauge function satisfies

$$\nabla \cdot \vec{A} = 0 \tag{1.100}$$

The classical Hamiltonian for the above charged particle system is

$$H(\vec{r},\vec{p}) = \frac{\pi^2}{2m} + z\phi \tag{1.101}$$

where $\vec{\pi} = \vec{p} - z\vec{A}$ is a generalized momentum. To convert to the quantum 1 electron Hamiltonian, we perform substitution: $p \to -i\nabla, z \to -1$, and add the coupling between electron spin and external field to get:

$$\begin{aligned} \hat{H}_{1e} &= \frac{1}{2}(-i\nabla + A)^2 + B \cdot s - \phi \\ &= \frac{1}{2}\nabla^2 - iA \cdot \nabla + \frac{1}{2}A^2 + B \cdot s - \phi \end{aligned} \tag{1.102}$$

Moving to $n$ electrons, we recover the Pauli Hamiltonian:

$$\hat{H} = \sum_i^n \left( \frac{1}{2}\nabla_i^2 - iA_i \cdot \nabla_i + \frac{1}{2}A_i^2 + B \cdot s_i - \phi_i \right) + \sum_{i \neq j} \frac{1}{r_{ij}} \tag{1.103}$$

With NMR properties in mind, for the following, we will assume the absence of an applied external electric field. The scalar potential then is due to nuclei only: $\phi = \phi_N$. Including only the first order in the multipole expansion:

$$\phi_N = \sum_K^N \frac{Z_K}{r_{iK}} \tag{1.104}$$

The vector potential has contributions from both the external magnetic field and the nuclear magnetic moments $A = A_B + A_{m_K}$

$$\boldsymbol{A}_B(\boldsymbol{r}_i) = \frac{1}{2}\boldsymbol{B} \times (\boldsymbol{r}_i - \boldsymbol{R}_O) \tag{1.105}$$

$$\boldsymbol{A}_{m_K}(\boldsymbol{r}_i) = \alpha^2 \sum_K \frac{\boldsymbol{m}_K \times \boldsymbol{r}_{iK}}{r_{iK}^3} \tag{1.106}$$

$\alpha$ is the fine structure constant, which in atomic units is $\alpha \approx \frac{1}{137}$. The corresponding magnetic field due to the nuclear magnetic moments is

$$\boldsymbol{B}_{m_K}(\boldsymbol{r}_i) = \alpha^2 \sum_K \frac{3(\boldsymbol{m}_K \cdot \boldsymbol{r}_{iK})\boldsymbol{r}_{iK} - r_{iK}^2 \boldsymbol{m}_K}{r_{iK}^5} + \frac{8\pi}{3}\alpha^2 \sum_K \delta(r_{iK})\boldsymbol{m}_K \tag{1.107}$$

The final form of the Hamiltonian is given by

$$\begin{aligned}
\hat{H} = &-\frac{1}{2}\sum_i^n \nabla_i^2 + \frac{1}{2}\sum_i^n \boldsymbol{B}\cdot\boldsymbol{L_i} + \frac{1}{8}\sum_i^n [B^2 r_i^2 - (\boldsymbol{B}\cdot\boldsymbol{r_i})^2] \\
&+ \alpha^2 \sum_i^n \sum_K^N \frac{\boldsymbol{m}_K \cdot \boldsymbol{L}_{iK}}{r_{iK}^3} \\
&+ \frac{\alpha^2}{2}\sum_i^n \sum_K^N \frac{(\boldsymbol{B}\cdot\boldsymbol{m}_K)(\boldsymbol{r}_{iN}\cdot\boldsymbol{r}_{iK}) - (\boldsymbol{B}\cdot\boldsymbol{r}_{iK})(\boldsymbol{r}_{iN}\cdot\boldsymbol{m}_K)}{r_{iK}^3} \\
&+ \frac{\alpha^4}{2}\sum_{KL} \frac{(\boldsymbol{m}_K\cdot\boldsymbol{m}_L)(\boldsymbol{r}_{iK}\cdot\boldsymbol{r}_{iL}) - (\boldsymbol{m}_K\cdot\boldsymbol{r}_{iL})(\boldsymbol{r}_{iC}\cdot\boldsymbol{m}_D)}{r_{iC}^3 r_{iD}^3} \\
&+ \alpha^2 \sum_i^n \sum_K^N \frac{3(\boldsymbol{m}_K\cdot\boldsymbol{r}_{iK})\boldsymbol{r}_{iK}\cdot\boldsymbol{s}_i - r_{iK}^2 \boldsymbol{m}_K\cdot\boldsymbol{s}_i}{r_{iK}^5} + \frac{8\pi}{3}\alpha^2 \sum_i^n \sum_K^N \delta(r_{iK})\boldsymbol{m}_K\cdot\boldsymbol{s}_i \\
&- \sum_i^n \sum_K^N \frac{Z_K}{|\mathbf{r}_{iK}|} + \sum_{i\neq j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}
\end{aligned} \tag{1.108}$$

## Derivative evaluation

For exact wavefunctions, we can apply perturbation theory to obtain properties. A first order molecular property is given by

$$\frac{\mathrm{d}E(x)}{\mathrm{d}x_i} = \langle 0|\frac{\mathrm{d}H}{\mathrm{d}x_i}|0\rangle \tag{1.109}$$

A second order molecular property is given by

$$\frac{\mathrm{d}^2E(x)}{\mathrm{d}x_i\mathrm{d}x_j} = \langle 0|\frac{\mathrm{d}^2H}{\mathrm{d}x_i\mathrm{d}x_j}|0\rangle - 2\sum_{nne0} \frac{\langle 0|\frac{\mathrm{d}H}{\mathrm{d}x_i}|n\rangle\langle n|\frac{\mathrm{d}H}{\mathrm{d}x_j}|0\rangle}{E_n - E_0} \tag{1.110}$$

$|0\rangle$ is the exact ground state and $|n\rangle$ is the $n^{th}$ exact excited state. These results are consequences of the Hellmann-Feynman theorem, and $H^{(0)}|n\rangle = E_n|0\rangle$.

The sum over states expression requires all excited states explicitly. Accurate excited state calculations for even a few states is not a trivial task. In the following, we will discuss general approaches to calculating properties with the electronic structure methods introduced in the previous section.

Energy derivatives with approximate wavefunctions can be evaluated analytically or numerically. First we will briefly describe the general formalism of the analytic approach. For this purpose we write some variationally optimized energy E (for example the HF energy) as a function of general perturbations $\boldsymbol{A}$ and $\boldsymbol{B}$, and a set of wavefunction variational parameters $\boldsymbol{\lambda}$: $E(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{\lambda})$. Denoting $\boldsymbol{\lambda}^*$ as the optimized parameter, the optimized electronic energy is then given by $E(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{\lambda}^*)$. The variational condition can be expressed as

$$\frac{\partial E(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}}\Big|_{\boldsymbol{\lambda}=\boldsymbol{\lambda}^*} = 0 \tag{1.111}$$

It holds for all values of perturbations $\boldsymbol{A}$ and $\boldsymbol{B}$.

The first derivative is given by applying the chain rule:

$$\frac{\mathrm{d}E(\boldsymbol{A}, \boldsymbol{B})}{\mathrm{d}\boldsymbol{A}} = \left(\frac{\partial E(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{\lambda})}{\partial \boldsymbol{A}} + \frac{\partial E(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} \frac{\partial \boldsymbol{\lambda}}{\partial \boldsymbol{A}}\right)\Big|_{\boldsymbol{\lambda}=\boldsymbol{\lambda}^*} \tag{1.112}$$

The second portion vanishes due to the variational condition, Eq 1.111. The gradient of electronic energy with respect to perturbation $\boldsymbol{A}$ is then

$$\frac{\mathrm{d}E(\boldsymbol{A}, \boldsymbol{B})}{\mathrm{d}\boldsymbol{A}} = \frac{\partial E(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{\lambda}^*)}{\partial \boldsymbol{A}} \tag{1.113}$$

This is one form of the Hellmann-Feynman theorem.

We proceed similarly for the second derivative by applying the chain rule again:

$$\frac{\mathrm{d}^2E(\boldsymbol{A}, \boldsymbol{B})}{\mathrm{d}\boldsymbol{B}\mathrm{d}\boldsymbol{A}} = \left(\frac{\partial^2 E(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{\lambda})}{\partial \boldsymbol{B}\partial \boldsymbol{A}} + \frac{\partial^2 E(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}\partial \boldsymbol{A}} \frac{\partial \boldsymbol{\lambda}}{\partial \boldsymbol{B}}\right)\Big|_{\boldsymbol{\lambda}=\boldsymbol{\lambda}^*} \tag{1.114}$$

From this result, we see that to calculate second derivatives of the energy, only the wavefunction's first order response to perturbation $\boldsymbol{B}$, represented by $\frac{\partial \boldsymbol{\lambda}}{\partial \boldsymbol{B}}$, is required. To obtain a form for this response, we differentiate Eq 1.111 on both sides with respect to $\boldsymbol{B}$:

$$\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{B}} \frac{\partial E(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} = 0 \tag{1.115}$$

$$\frac{\partial^2 E(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{\lambda})}{\partial \boldsymbol{B}\partial \boldsymbol{\lambda}} + \frac{\partial^2 E(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}^2} \frac{\partial \boldsymbol{\lambda}}{\partial \boldsymbol{B}} = 0 \tag{1.116}$$

$$\frac{\partial^2 E(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}^2} \frac{\partial \boldsymbol{\lambda}}{\partial \boldsymbol{B}} = -\frac{\partial^2 E(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{\lambda})}{\partial \boldsymbol{B}\partial \boldsymbol{\lambda}} \tag{1.117}$$

Eq 1.117 is a set of linear equations known as response equations. The first term on the left hand side $\frac{\partial^2 E}{\partial \lambda^2}$ is the electronic Hessian. The wavefunction response can be calculated by inverting the Hessian and multiplying it to the right hand side. This is not feasible when there are a large number of wavefunction parameters. In practice, $\frac{\partial \lambda}{\partial \boldsymbol{B}}$ is solved iteratively.

The above results, with some modifications, are applicable to non-variational methods as well. We first define a fictitious variational energy functional that, at its minima, gives the same energy as the reference non-variational energy.

$$E'(\boldsymbol{A}, \boldsymbol{B}) = E(\boldsymbol{A}, \boldsymbol{B}, \lambda^*) \tag{1.118}$$

where the nonvariational parameters $\lambda^*$, are determined from a set of equations:

$$\boldsymbol{e}(\boldsymbol{A}, \boldsymbol{B}, \lambda^*) = 0 \tag{1.119}$$

Then, we define a Lagrangian for this variational optimization:

$$L(\boldsymbol{A}, \boldsymbol{B}, \lambda, \lambda') = E(\boldsymbol{A}, \boldsymbol{B}, \lambda) + (\lambda')^T \boldsymbol{e}(\boldsymbol{A}, \boldsymbol{B}, \lambda) \tag{1.120}$$

where $\boldsymbol{e} = 0$ is the constraint, and $\lambda'$ are the set of associated multipliers. Substituting $L$ as $E$ into Eq 1.114, we get

$$\begin{aligned}\frac{\mathrm{d}^2 E(\boldsymbol{A}, \boldsymbol{B})}{\mathrm{d}\boldsymbol{B}\mathrm{d}\boldsymbol{A}} =& \Bigg[\frac{\partial^2 E(\boldsymbol{A}, \boldsymbol{B}, \lambda)}{\partial \boldsymbol{B} \partial \boldsymbol{A}} + (\lambda')^T \frac{\partial^2 \boldsymbol{e}(\boldsymbol{A}, \boldsymbol{B}, \lambda)}{\partial \boldsymbol{B} \partial \boldsymbol{A}} + \\ & \left(\frac{\partial^2 E(\boldsymbol{A}, \boldsymbol{B}, \lambda)}{\partial \lambda \partial \boldsymbol{A}} + \frac{\partial^2 \boldsymbol{e}(\boldsymbol{A}, \boldsymbol{B}, \lambda)}{\partial \lambda \partial \boldsymbol{A}}\right) \frac{\partial \lambda}{\partial \boldsymbol{B}} + \\ & \frac{\partial \boldsymbol{e}(\boldsymbol{A}, \boldsymbol{B}, \lambda)^T}{\partial \boldsymbol{A}} \frac{\partial \lambda'}{\partial \boldsymbol{B}}\Bigg]\Bigg|_{\lambda=\lambda^*}\end{aligned} \tag{1.121}$$

where we observe that in addition to the primary parameter $\lambda$'s response, the response of the multiplier $\lambda'$ is needed as well. These are determined by solving the following sets of equations:

$$\frac{\partial \boldsymbol{e}(\boldsymbol{A}, \boldsymbol{B}, \lambda)}{\partial \lambda} \frac{\partial \lambda}{\partial \boldsymbol{B}} = -\frac{\partial \boldsymbol{e}(\boldsymbol{A}, \boldsymbol{B}, \lambda)}{\partial \boldsymbol{B}} \tag{1.122}$$

$$\begin{aligned}\frac{\partial \lambda'}{\partial \boldsymbol{B}}^T \frac{\partial \boldsymbol{e}(\boldsymbol{A}, \boldsymbol{B}, \lambda)}{\partial \lambda} =& -\left(\frac{\partial^2 E(\boldsymbol{A}, \boldsymbol{B}, \lambda)}{\partial \boldsymbol{B} \partial \lambda} + (\lambda')^T \frac{\partial^2 \boldsymbol{e}(\boldsymbol{A}, \boldsymbol{B}, \lambda)}{\partial \boldsymbol{B} \partial \lambda}\right) \\ & -\left(\frac{\partial^2 E(\boldsymbol{A}, \boldsymbol{B}, \lambda)}{\partial \lambda^2} + (\lambda')^T \frac{\partial^2 \boldsymbol{e}(\boldsymbol{A}, \boldsymbol{B}, \lambda)}{\partial \lambda^2}\right) \frac{\partial \lambda}{\partial \boldsymbol{B}}\end{aligned} \tag{1.123}$$

The first set of equations gives the response of the primary parameters to the perturbation $\boldsymbol{B}$ as $\frac{\partial \lambda}{\partial \boldsymbol{B}}$. The second gives the response of the multipliers to perturbation $\frac{\partial \lambda'}{\partial \boldsymbol{B}}$.

In contrast, the numerical approach to calculating derivatives is simple on paper. They are given by finite differences of energies evaluated at finite perturbations. For example, second order derivatives with central finite differences to second order are given by

$$\frac{\mathrm{d}^2 E(\boldsymbol{A}, \boldsymbol{B})}{\mathrm{d}A_i \mathrm{d}B_j} = -\frac{E(\Delta A_i, \Delta B_j) - E(\Delta A_i, -\Delta B_j) - E(-\Delta A_i, -\Delta B_j) + E(-\Delta A_i, -\Delta B_j)}{4 \Delta A_i \Delta B_j} \tag{1.124}$$

where 4 energy evaluations are required for a single element of the property tensor. Similarly, one can take a mixed derivative approach, where we take finite differences to second order of gradients

$$\frac{\mathrm{d}^2 E(\boldsymbol{A}, \boldsymbol{B})}{\mathrm{d}A_i \mathrm{d}B_j} = -\frac{F_i(\Delta B_j) - F_i(-\Delta B_j)}{2 \Delta B_j} \tag{1.125}$$

where we have written the gradient of energy with respect to perturbation $\frac{\mathrm{d}E(\boldsymbol{A}, \boldsymbol{B})}{\mathrm{d}\boldsymbol{A}}$ as $\boldsymbol{F}$.

Comparing the analytical and numerical approaches, the main drawback of the analytical approach is that for each electronic structure method, its response equations need to be derived and implemented. The numerical approach only needs the implementation of the perturbed energy. However, it suffers from numerical stability issues. While it is in general less efficient than the analytic approach, it is naively parallelizable, no communication between processes are required. We note in passing that the mixed derivative approach is a good middle ground, especially for variational methods.

## Gauge Origin problem

The vector potential given by Eq 1.105 satisfies

$$\boldsymbol{B} = \nabla \times \boldsymbol{A}_B(\boldsymbol{r}) \tag{1.126}$$

$\boldsymbol{A}_B$ is not uniquely defined as any gauge origin $\boldsymbol{R}_O$ satisfies Eq 1.126. As a result, the field Hamiltonian is also not uniquely defined. This stands in contrast with the energy and properties of the system, which should not be dependent on the choice of gauge origin.

Consider a gauge transformed vector potential:

$$\boldsymbol{A}'_B(\boldsymbol{r}) = \boldsymbol{A}_B(\boldsymbol{r}) + \nabla f(\boldsymbol{r}) \tag{1.127}$$

where $f$ is some scalar function. We note that this corresponds to the same $\boldsymbol{B}$, since $\nabla \times \nabla f = 0$. The gauge transformation represents a unitary transformation of the Hamiltonian:

$$\hat{H}' = e^{-if(r)} \hat{H} e^{if(r)} \tag{1.128}$$

The exact wavefunction therefore has to transform as

$$\psi'(\boldsymbol{r}) = e^{-if(\boldsymbol{r})} \psi(\boldsymbol{r}) \tag{1.129}$$

For this wavefunction, Hamiltonian expectation value is unchanged after gauge transformation:

$$\langle \psi' | H' | \psi' \rangle = \langle \psi | H | \psi \rangle \tag{1.130}$$

Going back to the effect of the gauge origin, the scalar function corresponding to a gauge origin shift is

$$f(\boldsymbol{r}) = \frac{1}{2} \boldsymbol{B} \times (\boldsymbol{O}' - \boldsymbol{O}) \cdot \boldsymbol{r} \tag{1.131}$$

And the transformed wavefunction is

$$\psi_{O'}(\boldsymbol{r}) = \exp\left[ -\frac{i}{2} \boldsymbol{B} \times (\boldsymbol{O}' - \boldsymbol{O}) \cdot \boldsymbol{r} \right] \psi_O(\boldsymbol{r}) \tag{1.132}$$

However Eq 1.132 is only true for exact wavefunctions. Approximate wavefunctions in general do not transform correctly, causing unphysical gauge origin dependence for calculated quantities.

This leads to slow basis set size convergence for magnetic properties. This limited early attempts at NMR calculations to small systems, since large basis sets are needed for accurate results. There have been numerous attempts to remedy this problem. The prevalent approach nowadays is the use of gauge-including atomic orbitals (GIAO), sometimes called London atomic orbitals in the literature. This approach was introduced by London in 1937 for a study on ring currents of aromatic molecules. A field dependent complex phase factor is attached to each atomic orbital. A GIAO $\omega_\mu$ is defined as

$$\omega_\mu(\mathbf{r}) = \exp\left(-\frac{i}{2}\mathbf{B} \times (\mathbf{R}_\mu - \mathbf{R}_O) \cdot \mathbf{r}\right) \cdot \chi_\mu(\mathbf{r}) \tag{1.133}$$

where $\chi_\mu(\mathbf{r})$ is a regular Gaussian-type orbital(GTO) and $\mathbf{R}_O$ is the vector for gauge origin $O$, which is arbitrarily chosen as the Cartesian origin for simplicity.

In a GIAO calculation, all GTO's are replaced by GIAO's. Introduction of this complex factor leads to complex AO matrix elements for all operators in the Hamiltonian, Eq 1.108, and a complex SCF procedure with complex MO coefficients $\boldsymbol{C}$ is needed. This carries the extra compute cost of complex arithmetic. In addition, we encounter loss of symmetry in the 2-electron MO integrals given by Eq 1.21. For real orbitals, we have eightfold permutation symmetry where

$$\langle pq|rs\rangle = \langle qp|sr\rangle = \langle rs|pq\rangle = \langle sr|qp\rangle = \langle rq|ps\rangle = \langle sp|qr\rangle = \langle ps|rq\rangle = \langle qr|sp\rangle \tag{1.134}$$

For complex orbitals, we are left with fourfold permutation symmetry where

$$\langle pq|rs\rangle = \langle qp|sr\rangle = \langle rs|pq\rangle^* = \langle sr|qp\rangle^* \tag{1.135}$$

## 1.3 Outline

### Nuclear Shieldings Computed Via Finite Fields: Assessing Alternatives to MP2 and MP3

In Chapter 2, we present the development and implementation of a method independent, fully numerical finite difference approach to calculating NMR shieldings, using gauge-including atomic orbitals. The resulting capability can be used to explore non-standard methods given only the energy as a function of finite applied magnetic fields and nuclear spins. For example, standard second order Møller-Plesset theory (MP2) has well-known efficacy for $^{1}$H and $^{13}$C shieldings, and known limitations for other nuclei such as $^{15}$N and $^{17}$O. It is therefore interesting to seek methods that offer good accuracy for $^{15}$N and $^{17}$O shieldings without greatly increased compute costs, as well as exploring whether such methods can further improve $^{1}$H and $^{13}$C shieldings. Using a small molecule test set of 28 species, we assessed the accuracy for shieldings of two MP2 alternatives: $\kappa$ regularized MP2 ($\kappa$-MP2) which provides energy-dependent damping of large amplitudes and MP2.X, which includes a variable fraction, X, of third order correlation (MP3). The aug-cc-pVTZ basis was used, and CCSD(T) results were taken as reference values. Our $\kappa$-MP2 results reveal that significant improvements are possible over MP2 for $^{13}$C andand $^{15}$N, with the optimal $\kappa$ value

being strongly element-specific. $\kappa$-MP2 with $\kappa = 2$ offers roughly 30% RMS error reduction over MP2 and outperforms CCSD for $^{13}$C shieldings. For $^{15}$N, $\kappa$-MP2 with $\kappa = 1.1$ provides over 90% error reduction vs MP2, and around 60% error reduction vs CCSD. On the other hand, the scaled MP3 method MP2.X with scaling factor 0.6 impressively outperformed CCSD (as well as far outperforming MP3 itself) for all heavy nuclei. These results can be understood as providing renormalization of doubles amplitudes to partially account for neglected triple and higher substitutions, and offer promising opportunities for future applications.

## Analytic First-Order Nuclear Derivatives for Regularized Orbital-Optimized MP2

In Chapter 3, we present the derivation and implementation of analytical nuclear gradients of the $\kappa$ regularized orbital optimized MP2 ($\kappa$-OOMP2) method. One of the most important first order molecular property in electronic structure theory is the force on each atom, which enables walking on potential energy surfaces via geometry optimization or trajectory calculations. MP2, being the simplest and computationally cheapest correction to HF theory for electron correlation effects, breaks down in the presence of spin-contamination and small HOMO-LUMO gaps. Recent work have shown that $\kappa$-OOMP2 resolves these issues effectively, producing accurate thermochemistry results. To test the efficacy of $\kappa$-OOMP2 nuclear gradients, we applied it to a few systems with varying complexities, including the methanol radical cation, transanular chalcogen-chalcogen interactions, vibrational frequencies of radical species, and single-triplet gap of the $C_{60}$ fullerene. We found that $\kappa$-OOMP2 was able to resolve most of the artifacts of UMP2 that arised from the HF reference. It restored the energy landscape of radical species, and produced high quality vibrational frequencies. In addition, compared to the optimal $\kappa$ value of 1.45 fitted for thermochemistry, weaker regularization produced smaller errors for the systems we investigated.

## Mechanism of Solvent Water Oxygen Incorporation into the Oxygenated Products of CO Reduction over Cu

In Chapter 4, we explore the mechanism of a reaction involved in the process of $CO_2$ reduction reaction (CO2RR). The electrochemical reduction of CO and $CO_2$ over Cu produces a variety of multicarbon products. Interestingly, recent isotope experiments have suggested that the oxygen atoms contained in the multicarbon alcohols produced over Cu are derived from solvent water. This observation has brought into question many of the proposed reaction mechanisms by which these multicarbon alcohols are produced over Cu. However, these surprising experimental observations are likely the result of isotopic scrambling between transiently produced carbonyl-containing intermediate reaction products, such as acetaldehyde, with solvent water and not another mechanism. The existence of such carbonyl-containing intermediate reaction products is supported by both experimental and theoretical studies. Furthermore, theoretical calculations support the notion that the reversible hydration of these carbonyl-containing species is facile in the vicinity of the Cu surface.

# Chapter 2

# Nuclear Shieldings Computed Via Finite Fields: Assessing Alternatives to MP2 and MP3

## 2.1 Introduction

Nuclear magnetic resonance (NMR) spectroscopy is a widely used technique in organic chemistry and biochemistry for determining the identity and structure of molecules[32–35]. It works by applying a strong magnetic field to a sample and measuring the response of the nuclear spins in the sample to the applied field. The resulting spectrum contains information about the chemical shifts and coupling constants of the nuclei in the molecular sample, which can be used to determine the structure (and the identity) of the molecule. For large systems, structure determination from the NMR spectrum is not a straightforward task. As a result, NMR spectra prediction with quantum chemistry methods has become quite widely used in aiding the assignment of experimental spectra[36–39].

To this end, a plethora of methods has been developed over the years. Lower cost methods such as Hartree-Fock (HF) theory [40] and density functional theory (DFT) provide lower accuracy approximations to chemical shifts. The level of accuracy attainable for magnetic properties using pure or hybrid density functionals is generally not considered sufficient for predictive, quantitative applications,[41, 42] in contrast to the steady continued developments in the accuracy of DFT methods for molecular structures and relative energies[20, 29]. The basic reason for the poor performance of DFT for chemical shifts is that the standard existence theorems upon which modern functional developments are based are formally inapplicable to molecules in magnetic fields[43, 44]. One extension (CDFT) requires that the functionals depend not just upon the electron density, but also the current (or paramagnetic current (C)) density[45, 46], which is typically not done at present. Another (BDFT) requires that the XC functional depends explicitly on both the density and the magnetic field (B)[47], which is currently not widely used. Routine DFT-based calculations of NMR shieldings and chemical shifts use neither of these extensions, and nonetheless offer significantly improved accuracy over HF theory, although accuracy of course varies significantly from functional to functional.[42, 48–52]. Additionally, linear and sublinear scaling implementations have allowed application of DFT to large systems with more than 1000 heavy atoms[53–55].

There is ample precedent showing that wave function-based methods based on coupled cluster (CC) theory[56, 57] are the best solution to the problem of inadequate accuracy in DFT for NMR chemical shifts. For example, RMS errors for absolute $^{13}$C chemical shieldings of small molecules in the gas phase using coupled cluster with singles and doubles and perturbative triples (CCSD(T))[6] could yield a standard deviation of less than 0.8 ppm versus experiment[41]: these errors are reduced by a further factor of 5-10 or so when relative shieldings in related environments are evaluated.[42, 51, 58, 59] However there are very strong size-limitations on the applicability of CCSD(T): CCSD(T) scales as $\mathcal{O}(A^7)$ with number of atoms, $A$, which limits production calculations to small molecules and small numbers of calculations. Therefore CCSD(T) has been used to provide most of the benchmark data for assessing and developing computational methods for NMR shielding that have lower compute costs.

Alternatively, second order Møller-Plesset perturbation theory (MP2) is the lowest cost post-HF method that accounts for correlation, which as already implied is essential for accurate chemical shift predictions. [41, 60, 61] MP2 has the relatively low computational scaling of $\mathcal{O}(A^5)$ and offers much higher quality prediction than HF or low level DFT's particularly for $^1$H shieldings (and

to a lesser degree, $^{13}C$ shieldings)[48, 62]. As a result, MP2 has seen considerable development for calculating chemical shifts. These include the early z-vector approach[63], local-correlation approximation[64–66], AO-MP2[67–71], and Cholesky decomposition (CD)[72, 73]. There were efforts to improve accuracy as well, for example, spin component scaled MP2 (SCS-MP2)[74], and spin opposite scaled MP2 (SOS-MP2)[75]. The closely related double-hybrid (DH) DFT has also received attention in this context[76, 77].

Yet MP2 is known to fail dramatically when the energy gap between the energy of highest occupied and lowest unoccupied molecular orbitals (HOMO and LUMO) become small, because the correlation amplitudes may diverge. Various novel approaches for relatively low cost stabilization (or regularization) of MP2 have been developed recently, including the $\delta$[78], $\sigma$, and $\kappa$[79] regularizers. In particular $\kappa$-MP2,[80] reviewed in detail below, has been shown to not just avoid divergences, but also significantly improve the accuracy of MP2 for non-bonded interactions where it exhibits large errors.[81] In a related vein, MP3, which scales as $\mathcal{O}(A^6)$, yields such disappointing results for molecular relative energies that one typically does not employ it in practice, resorting instead to the infinite order CCSD method with iterative $\mathcal{O}(A^6)$ compute cost. This problem, associated with poor convergence of the MP series, can be greatly ameliorated for non-bonded interactions by damping the third order contribution by 0.5 in the MP2.5 method, or by some variable fraction X in the MP2.X method. Such scaled MP3 approaches[82–84] have been recently shown to work even better for relative energies using non-canonical reference orbitals[85–87].

We are interested in applying novel methods of intermediate compute cost such as those reviewed above to the calculation of NMR properties. However the shielding is a 2nd order properties[37, 88], whose analytical implementation (i.e. as second order responses) is enormously challenging in terms of human labor. Selecting a new theory for painstaking implementation is a high-risk project that has historically been addressed one wavefunction method at a time. Instead, building on our recent work on magnetic susceptibilities[89], we have chosen a different approach. Here we report our implementation of an in-silico laboratory for magnetic shielding by applying finite nuclear spins at a nucleus, $K$, ($\mathbf{m}_K$) and magnetic fields ($\mathbf{B}$) directly to the molecular Hamiltonian:

$$\hat{H}(\mathbf{m}_K, \mathbf{B}) = \hat{H}^{(0)} + \hat{H}^{(1)}(\mathbf{m}_K, \mathbf{B}) + \hat{H}^{(2)}(\mathbf{m}_K, \mathbf{B}) \tag{2.1}$$

Here the superscript (1) indicates linear terms, and the superscript (2) indicates quadratic terms. Solving for the electronic structure energy, $E(\mathbf{m}_K, \mathbf{B})$ for different finite values of $\mathbf{B}$ and $\mathbf{m}_K$ in principle permits the evaluation of the magnetic shielding as second order finite differences, approximating the true derivative. Thus the usual laborious implementation of the analytical derivative is completely avoided. This approach has been successfully used before by other groups[90–92] for diverse purposes, including properties of molecules in strong fields, and magnetic properties.

The organization of the remainder of the paper is as follows. In Section 2.2, we discuss the theoretical background of the calculation of NMR shielding tensors with $\kappa$-MP2 and MP2.X, describe details of our implementation, introduce the test set from Lutnæs and co-workers,[93] and evaluate the accuracy of the numerical procedure. In Section 2.3, we assess the performance of $\kappa$-MP2 and MP2.X compared to CCSD(T) reference data, using the Lutnæs test set. Finally, conclusions are given in Section 2.4.

## 2.2 Theoretical Background

For in-depth treatment of the quantum chemical calculation of magnetic properties, we refer interested readers to these reviews on the subject.[37, 88, 94] We will present our approach by first introducing the general design idea of our code, then discuss the matrix elements that are required for the calculation of nuclear shieldings. We then briefly introduce the theoretical basis of $\kappa$-MP2 and MP3 for calculating the nuclear shielding tensor. Lastly, we will discuss the implementation and verification of our code, as well as the numerical performance compared to analytical results.

### General Approach

The nuclear shielding $\boldsymbol{\sigma}^K$ for nuclei *K* is a 3x3 tensor and can be calculated as the derivative of the energy *E* with respect to the magnetic field **B** and the induced magnetic moment $\mathbf{m}_K$,[37] with elements

$$\sigma_{ij}^K = \left.\frac{\partial^2 E\left(\{\mathbf{R}_N\}, \mathbf{B}, \mathbf{m}_K\right)}{\partial \mathrm{m}_{K,i} \partial \mathrm{B}_j}\right|_{\mathbf{B}=\mathbf{0},\mathbf{m}_K=\mathbf{0}} . \tag{2.2}$$

Here, $\{\mathbf{R}_N\}$ is the nuclear configuration. The isotropic shielding is $\sigma_{\text{iso}} = \frac{1}{3}\text{tr}(\boldsymbol{\sigma})$. The nuclear shielding tensor is related to the experimentally observed chemical shift

$$\delta = \frac{\sigma_{\text{iso,ref}} - \sigma_{\text{iso}}}{1 - \sigma_{\text{iso,ref}}} \tag{2.3}$$

where $\sigma_{ref}$ is some reference system.

We use Gauge-Including Atomic Orbitals (GIAOs),[95] also known as London orbitals, in the calculation of NMR properties to address the gauge-origin problem.[37] A GIAO $\omega_\mu$, centered on nucleus $N$, is given as

$$\omega_\mu(\mathbf{r}; \mathbf{A}_N) = \exp\left(-i\mathbf{A}_N \cdot \mathbf{r}\right) \cdot \chi(\mathbf{r}), \tag{2.4}$$

where **r** is the electron coordinate and $\chi(\mathbf{r})$ is a regular Gaussian-type orbital and the vector potential $\mathbf{A}_N$ is

$$\mathbf{A}_N = \frac{1}{2}\mathbf{B} \times \mathbf{R}_{NO}. \tag{2.5}$$

$\mathbf{R}_{NO}$ is the vector from the nucleus $N$ to an arbitrarily chosen gauge-origin $O$, which we choose to be the zero vector for convenience.

We calculate the second derivative of the energy with respect to the magnetic field and induced magnetic moment to second order *via*[96]

$$\sigma_{ij}^K = -\frac{E(\Delta B_i, \Delta m_{K,j}) - E(\Delta B_i, -\Delta m_{K,j}) - E(-\Delta B_i, -\Delta m_{K,j}) + E(-\Delta B_i, -\Delta m_{K,j})}{4\Delta B_i \Delta m_{K,j}} \tag{2.6}$$

$\Delta B_i$ and $\Delta m_{K,i}$ are steps in the Cartesian component $i$ of the magnetic field and nuclear magnetic moment respectively. The magnitude of the two steps determines the accuracy of the numerical results via roundoff error and/or contamination from higher derivatives. An assessment of $\Delta B_i$,

$\Delta m_{K,i}$ and their relation to the numerical error is given in Section 2.2. The calculation of the nuclear shieldings *via* eqs. 2.6 necessitates the calculation of several energies at different values of the magnetic field and nuclear magnetic moment. Due to the use of GIAOs, we use a fully complex-valued code throughout our calculations.

### Required Matrix Elements

The Hamiltonian in an external magnetic field is discussed in detail in ref. [97] and an excellent overview on the calculation of matrix elements involving GIAOs can be found in ref. [49]. For the sake of brevity, we will discuss only the terms in the field Hamiltonian that enter the nuclear magnetic shielding expression. The Hamiltonian with explicit dependence on the external magnetic field and nuclear magnetic moment is given by

$$\begin{aligned}\hat{H} = &-\frac{1}{2}\sum_{i}^{n}\nabla_i^2 - \sum_{i}^{n}\sum_{K}^{N}\frac{Z_K}{|\mathbf{r}_{iK}|} + \frac{1}{2}\sum_{i}^{n}\boldsymbol{B}\cdot\boldsymbol{L_i} + \alpha^2\sum_{i}^{n}\sum_{K}^{N}\frac{\boldsymbol{m}_K\cdot\boldsymbol{L}_{iK}}{r_{iK}^3} \\ &+ \frac{\alpha^2}{2}\sum_{i}^{n}\sum_{K}^{N}\frac{(\boldsymbol{B}\cdot\boldsymbol{m}_K)(\boldsymbol{r}_{iN}\cdot\boldsymbol{r}_{iK}) - (\boldsymbol{B}\cdot\boldsymbol{r}_{iK})(\boldsymbol{r}_{iN}\cdot\boldsymbol{m}_K)}{r_{iK}^3} + \sum_{i\neq j}\frac{1}{|\mathbf{r}_i-\mathbf{r}_j|}\end{aligned} \tag{2.7}$$

The terms in the Hamiltonian are, in order, kinetic, nuclear attraction, paramagnetic shielding, paramagnetic spin-orbit, diamagnetic shielding, and electron-electron repulsion. For details on the implementation of matrix elements of GIAO matrix elements of kinetic, nuclear attraction, paramagnetic shielding, and electron-electron repulsion, we refer readers to our previous work[89].

The diamagnetic shielding (DS) operator is[97]

$$\hat{h}_i^{\mathrm{DS}} = \frac{\alpha^2}{2}\sum_{K}\left[\frac{(\mathbf{B}\cdot\mathbf{m}_K)(\mathbf{r}_N\cdot\mathbf{r}_K) - (\mathbf{B}\cdot\mathbf{r}_K)(\mathbf{r}_N\cdot\mathbf{m}_K)}{|\mathbf{r}_K|^3}\right]. \tag{2.8}$$

The paramagnetic spin-orbit coupling (PSO) operator is [97]

$$\hat{h}_i^{\mathrm{PSO}} = -\frac{i\alpha^2}{2}\sum_{K}\left(\frac{\mathbf{m}_K\cdot\mathbf{r}_K\times\nabla}{|\mathbf{r}_K|^3}\right). \tag{2.9}$$

After some straightforward yet cumbersome algebra, the DS and PSO operators lead to energy expressions that involve only field integrals with increased or decreased angular momentum on GIAO $\omega_\nu$. The field integrals can also be further decomposed into nuclear attraction integrals with increased or decreased angular momentum on GIAO $\omega_\nu$. (see Appendix) The first field integral based approach is much less cumbersome than using modified nuclear attraction integrals, and is used for all the results in this paper.

### Nuclear Shieldings with $\kappa$-MP2

The following is a brief description of regularization theory and the $\kappa$ regularized MP2 approach ($\kappa$-MP2), recently proposed by our group [79, 80, 98, 99] We begin from the textbook canonical

MP2 correlation energy expression, which is given by

$$E_{\text{MP2}} = -\frac{1}{4}\sum_{ijab}\frac{|\langle ab||ij\rangle|^2}{\Delta_{ij}^{ab}} \tag{2.10}$$

Here $\Delta_{ij}^{ab} = \epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j$, which is composed of sums and differences of orbital energies, $\epsilon$, of virtual levels, $a, b$ and occupied levels, $i, j$.

The above expression, Eq. 2.10, for $E_{\text{MP2}}$, diverges when $\Delta_{ij}^{ab}$ is small. This will be the case for metals, and is almost the case for small gap systems such as semiconductor clusters, or, in particular, some unstable reactive species, such as molecules with stretched bonds or high energy transition structures. Addressing this issue defines the regularization problem.

The simplest possible approach is to use a level shift, defining $\delta$ regularization. We thus shift the denominator $\Delta_{ij}^{ab}$ by a constant factor $\delta$: $\Delta_{ij}^{ab} \longleftarrow \Delta_{ij}^{ab} + \delta$. This achieves the dual purpose of preventing divergence and damping MP2's over correlation. However, it was found that for the purpose of Coulson-Fischer point restoration for double and triple bond dissociations, $\delta$ had to be really large.[100] Such a large level shift deteriorates the performance of $\delta$-MP2 for thermochemistry and other properties because it damps all correlation contributions, regardless of how close $\Delta_{ij}^{ab}$ is to zero. This is clearly not a satisfactory approach.

An alternative approach is the $\sigma$-regularizer, with the idea of adjusting the damping strength of the correlation contribution as a function of $\Delta_{ij}^{ab}$. The $\sigma$ regularized MP2 ($\sigma$-MP2) expression is given by

$$E_{\sigma\text{-MP2}} = -\frac{1}{4}\sum_{ijab}\frac{|\langle ab||ij\rangle|^2}{\Delta_{ij}^{ab}}\left(1-\exp\left(-\sigma\Delta_{ij}^{ab}\right)\right), \tag{2.11}$$

A closed related scheme is $\kappa$ regularization. In $\kappa$ regularized MP2 ($\kappa$-MP2), the two-electron integrals are damped via $\langle ab||ij\rangle \longleftarrow \langle ab||ij\rangle\left(1-\exp\left(-\kappa\Delta_{ij}^{ab}\right)\right)$ such that the MP2 amplitudes $t$ are modified as

$$t_{ij}^{ab} = \frac{\langle ab||ij\rangle}{\Delta_{ij}^{ab}}\left(1-\exp\left(-\kappa\Delta_{ij}^{ab}\right)\right), \tag{2.12}$$

These regularized amplitudes lead to the $\kappa$-MP2 energy expression

$$E_{\kappa\text{-MP2}} = -\frac{1}{4}\sum_{ijab}\frac{|\langle ab||ij\rangle|^2}{\Delta_{ij}^{ab}}\left(1-\exp\left(-\kappa\Delta_{ij}^{ab}\right)\right)^2, \tag{2.13}$$

In the limit of $\Delta_{ij}^{ab} \to 0$, $t_{ij}^{ab} \to 0$ instead of $\infty$, and the corresponding correlation energy contribution is accordingly also zero. At the same time, when $\Delta_{ij}^{ab}$ is large, $t_{ij}^{ab}$ is undamped, and the correlation contribution approaches the MP2 value. This energy-dependent regularization (small gap amplitudes strongly damped, large gap amplitudes unaffected) has been demonstrated to significantly improve chemical results for intermolecular interactions, ligand binding energies, and chemical bond and atomization energies.[80]

**Nuclear Shieldings with MP3**

Adding the next order correction in the MP series defines the MP3 method. The MP3 correlation energy is given by

$$E_{\text{MP3}} = \frac{1}{8}\sum_{ijabcd}(t_{ij}^{ab})^*\,\langle ab||cd\rangle\,t_{ij}^{cd} + \frac{1}{8}\sum_{ijklab}(t_{ij}^{ab})^*\,\langle kl||ij\rangle\,t_{kl}^{ab} - \sum_{ijkabc}(t_{ij}^{ab})^*\,\langle kb||ic\rangle\,t_{kj}^{ac} \tag{2.14}$$

MP3 formally scales as $\mathcal{O}(A^6)$ with respect to molecule size. At the expense of some loss of accuracy, applying tensor hypercontraction (THC) techniques[101] allows a lower scaling of $\mathcal{O}(A^4)$.

From an accuracy standpoint, it is well known that MP3 often produces results that are deteriorated relative to the simpler and less computationally costly MP2 method.[102, 103] This is due to the oscillating behaviour of the MP series, in which even orders of perturbation increases the number of electrons correlated simultaneously, while odd orders of perturbation theory recouple the existing correlation amplitudes. In effect, the odd orders temper the over-correlation which is present at even orders. However, MP3 correlation often overcorrects, resulting in an overall under-correlation. This observation has inspired several attempts to damp the third order overcorrection. Grimme proposed the SCS-MP3 method [74], where MP3 correlation contribution was scaled by 0.25 in addition to the SCS-MP2 scaling parameters. Improvements in atomization, reaction, and ionization energies were reported, in some cases approaching the quality of the more costly QCISD method. Alternatively, Hobza and coworkers proposed MP2.5 and MP2.X methods [82–84], where the MP3 correlation contribution is scaled with a single empirical parameter, $c$, usually less than 1. The MP2.X energy is given by

$$E_{\text{MP2.X}}(c) = E_{\text{HF}} + E_{\text{MP2}} + cE_{\text{MP3}} \tag{2.15}$$

For MP2.5, c = 0.5. Hobza et al. reported significant improvements in the description of non-covalent interactions. Like $\kappa$-MP2, this method has the attractive quality of having only one empirical parameter.

**Implementation and Code Verification**

We have implemented all one- and two-electron integrals given in Section 2.2 and $\kappa$-MP2 and MP3 using GIAOs in the integral library `libqints` in a developer's version of Q-Chem 6.0 [104] Our code is able to run complex-restricted[98, 105] and complex-general[99, 106] SCF procedures as implemented in `libgscf` and `libgmbpt`, the latter of which is important for the calculation of indirect spin-spin coupling constants (to be presented in a future publication).

In a first step, the new matrix elements were tested in order to verify our code. Since the nuclear attraction integrals were already tested, we tested the field integral based implementation against the nuclear attraction integral based implementation.

In the second verification step, nuclear shieldings at the Hartree-Fock level and at the RI-MP2 level with very large auxiliary basis sets were tested against analytical shieldings evaluated using the CFOUR[107] code. We also employed the CFOUR package to generate CCSD and CCSD(T) results against which the various MP methods can be assessed.

**Lutnæs data set**

In the following sections, we will use the data set curated by Lutnæs and coworkers [93] with the purpose of benchmarking magnetic properties. The data set contains the following 28 small molecules: AlF, $C_2H_4$, $C_3H_4$, $CH_2O$, $CH_3F$, $CH_4$, CO, FCCH, FCN, $H_2C_2O$, $H_2O$, $H_2S$, $H_4C_2O$, HCN, HCP, HF, HFCO, HOF, LiF, LiH, $N_2$, $N_2O$, $NH_3$, $O_3$, OCS, $OF_2$, PN, $SO_2$. Geometries were optimized at the CCSD(T)/cc-pVTZ level of theory. Of the 28 molecules, 4 are considered "difficult" cases for NMR shieldings. They are $O_3$, $OF_2$, PN, and $SO_2$. $O_3$ exhibits significant multireference character, while $SO_2$ exhibits moderate multireference character. PN does not display distinct multireference character, but has a large paramagnetic contribution to the nuclear shielding tensor. For $OF_2$, inclusion of triple excitations is important for accurate nuclear shieldings. The rest of the 24 molecules are considered "easy" in the sense that electron correlation effects are not strong. Even so, electron correlation effects are essential to obtain results of useful accuracy. Given the large absolute value of the $^{17}$O shieldings of $O_3$, we will exclude it from all statistical assessment in order to avoid skew. This procedure is in line with previous benchmark studies.[93, 108, 109] The data set then consists of unique nuclear shieldings for 18 $^{1}$H nuclei, 17 $^{13}$C nuclei, 7 $^{15}$N nuclei, 11 $^{17}$O nuclei, and 9 $^{17}$F nuclei.

**Numerical Accuracy of Finite Differences Nuclear Shieldings**

To evaluate the numerical accuracy of the finite difference scheme, we compared the finite difference HF isotropic nuclear shieldings of the Lutnæs data set against analytical HF isotropic nuclear shieldings obtained via CFOUR. The cc-pVDZ basis was used for this purpose. The SCF convergence criteria was $10^{-11}$ a.u., and the cutoff threshold for two electron integrals was $10^{-17}$ a.u.

In Table 2.1 we present the mean error (ME) and root mean square error (RMSE) for HF nuclear shieldings for the Lutnæs data set. With both the second order and fourth order finite difference scheme, the optimal step size of $10^{-2.3}$ a.u. in both external magnetic field $B$ and nuclear magnetic moment $m_K$ leads to mean errors on the order of $10^{-2}$ ppm and $10^{-6}$ ppm respectively. Second order differences with the optimal step size of $10^{-2.3}$ a.u. for both perturbations are then used to evaluate all results reported here.

## 2.3 Results and Discussion

We assess the performance of various MP2 and MP3 alternatives with the Lutnæs data set. To contextualize the errors in the following sections, we introduce a 3 level target error scale. The high accuracy target is 0.1 ppm for $^{1}$H nuclei, 1 ppm for $^{13}$C nuclei, 3 ppm for $^{15}$N nuclei, and 4 ppm for $^{17}$O nuclei. The medium accuracy target is 0.2 ppm for $^{1}$H nuclei, 2 ppm for $^{13}$C nuclei, 6 ppm for $^{15}$N nuclei, and 8 ppm for $^{17}$O nuclei. Finally the low accuracy target is 0.4 ppm for $^{1}$H nuclei, 4 ppm for $^{13}$C nuclei, 12 ppm for $^{15}$N nuclei, and 16 ppm for $^{17}$O nuclei. This can be loosely applied to statistical error measures of either RMSE or MAE. Various benchmark studies[48, 51, 52, 110] have investigated the nuclear shielding performance of HF, MP2, and many DFT functionals with various basis sets. Notably, for $^{1}$H, HF and DFT at various rungs with a large basis set achieve

Table 2.1: Numerical performance of the $2^{nd}$ order finite difference scheme for the HF/cc-pVDZ nuclear shieldings of 28 molecules with varying step sizes. Values presented are the ME and RMSE (where the latter is given in parentheses), in units of ppm. Reference values were calculated analytically at the same level of theory.

| $2^{nd}$ order | B step size ($log_{10}$ a.u.) | |
|---|---|---|
| $m_k$ step size ($log_{10}$ a.u.) | -2.3 | -2 |
| -2.3 | 0.027 (0.096) | 0.113 (0.384) |
| -2 | 0.029 (0.096) | 0.112 (0.382) |
| $4^{th}$ order | B step size ($log_{10}$ a.u.) | |
| $m_k$ step size ($log_{10}$ a.u.) | -2.3 | -2 |
| -2.3 | $-1.5\times10^{-6}$ (0.008) | 0.002 (0.007) |
| -2 | $7\times10^{-4}$ (0.005) | 0.002 (0.006) |

the medium accuracy target of 0.2 ppm, while MP2 comes close to the high accuracy target of 0.1 ppm. With a small basis, most DFT functionals are in the range of the low accuracy target. $^{13}$C proved to be slightly more difficult, with MP2 at large basis achieving the medium accuracy target of 2 ppm, and rung 5 DFT functionals achieving the low accuracy target of 4 ppm. $^{15}$N and $^{17}$O shieldings were shown to be difficult problems. MP2 performed poorly, beyond the range of the low accuracy target. Only a few rung 5 DFT functionals were in the low accuracy target range. It is therefore a worthwhile endeavor to explore alternatives in the MPn family of methods, seeking further improvements over MP2's already impressive performance for $^{1}$H and $^{13}$C, and seeking suitable methods for the more difficult $^{15}$N and $^{17}$O shieldings.

## Accuracy of $\kappa$-MP2 nuclear shieldings for the Lutnæs set

First, we consider the $\kappa$-MP2 method. We present $\kappa$-MP2 shielding results in Table 2.3, using the medium-sized aug-cc-pVTZ basis set. Since $\kappa$-MP2 energies depend non-linearly on $\kappa$, each chosen value requires its own calculations. Therefore we compare MP2 (or $\kappa \to \infty$) against HF (or $\kappa = 0$) and three judiciously chosen finite non-zero $\kappa$ values. The strongest regularization used is $\kappa = 1.1$, which is a value that was demonstrated to work well for ligand binding energies to transition metal complexes as well as strong intermolecular dispersion interactions.[80] The intermediate value is $\kappa = 1.45$, which was originally recommended for the orbital optimized version, based on thermochemistry.[9] To provide weak regularization, we also assess $\kappa = 2$.

Turning to the results in Table 2.3, we first observe that compared to MP2, overall error (i.e. across *all* nuclei) is reduced at all three regularization strengths. The largest gain was for the intermediate regularization parameter of $\kappa$ = 1.45, with a one quarter reduction in RMSE relative to MP2, and more than two thirds error reduction relative to HF. However, it is still double the error of

CCSD. Furthermore we notice that the overall RMSE is very similar for all three $\kappa$ values. At first glance this does not appear too encouraging. However, we recall the conclusion from other recent work[80, 111] assessing relative energies using $\kappa$-MP2 and $\kappa$-MP2 stating that no single $\kappa$ value is optimal for all classes of relative energies. It may well be that the optimal $\kappa$ value varies from nucleus to nucleus. Therefore we will next inspect the results for each individual nucleus type.

Table 2.2: $\kappa$-MP2 aug-cc-pVTZ Isotropic NMR shielding statistics in ppm

| | | | $\kappa$-MP2 | | | |
|---|---|---|---|---|---|---|
| All | HF | $\kappa$=1.1 | 1.45 | 2 | MP2 | CCSD |
| ME | -19.01 | 2.35 | 4.84 | 7.24 | 8.90 | -1.44 |
| RMSE | 49.03 | 15.93 | 14.18 | 15.82 | 19.92 | 7.83 |
| $^{1}$H | | | | | | |
| ME | 0.01 | -0.0004 | -0.015 | -0.035 | -0.08 | 0.06 |
| RMSE | 0.39 | 0.22 | 0.17 | 0.14 | 0.13 | 0.09 |
| $^{13}$C | | | | | | |
| ME | -11.31 | -2.97 | -1.24 | 0.40 | 1.65 | -1.56 |
| RMSE | 15.07 | 4.34 | 2.59 | 1.95 | 3.20 | 2.43 |
| $^{15}$N | | | | | | |
| ME | -58.12 | 1.18 | 10.77 | 19.02 | 26.07 | -7.41 |
| RMSE | 76.67 | 3.21 | 15.41 | 26.71 | 37.98 | 9.52 |
| $^{17}$O | | | | | | |
| ME | -38.17 | 16.91 | 19.33 | 21.40 | 19.00 | -0.06 |
| RMSE | 50.54 | 24.50 | 27.27 | 29.26 | 27.36 | 8.93 |
| $^{19}$F | | | | | | |
| ME | 19.76 | 16.69 | 12.16 | 8.18 | 3.97 | 5.87 |
| RMSE | 34.73 | 20.80 | 14.68 | 9.67 | 5.59 | 7.50 |

For $^{1}$H, no improvement over MP2 was observed: with the strongest regularization ($\kappa = 1.1$), RMSE almost doubled to 0.22 ppm from the unregularized MP2 value of 0.13. The weakest regularization ($\kappa = 2$) has similar RMSE as MP2. In fact that RMS changes monotonically across all 4 regularization values: $\kappa = 0, 1.1, 1.45, 2$. This is not unexpected, given the relatively high accuracy of unregularized MP2. We also note the trade-off between RMSE and ME. To say it another way, unregularized MP2 evidently does not systematically overestimate correlation effects

in $^{1}$H magnetic shieldings across the set of molecules studied here.

For $^{13}$C, the situation is qualitatively different. While MP2 performs reasonably well, a subset of regularization strengths lead to quite noticeable improvements over MP2 in the RMSE over our dataset. In particular, intermediate regularization $\kappa$ = 1.45 modestly decreases the RMSE from 3.2 ppm for MP2 to 2.7 ppm, while the ME changes sign from 1.7 ppm for MP2 to -1.3 ppm. The RMSE increases further (along with the ME) as we move to strong regularization. We observe optimal results With weak regularization ($\kappa$ = 2), where RMSE is reduced by a third to 2.0 ppm, and the ME is only 0.4 ppm. In fact, the RMSE and ME for weak regularization was lower than for CCSD, and reaches the medium level of accuracy ($\leq 2$ ppm). From the fact that the ME changes sign between HF and MP2, whilst it does not change sign for CCSD, we infer that unregularized MP2 slightly overcorrelates $^{13}$C shieldings, whilst CCSD slightly undercorrelates them. Therefore $\kappa$-MP2 with weak regularization ($\kappa$ = 2) can be a useful improvement over MP2 for $^{13}$C shieldings.

For $^{15}$N, the statistics summarized in Table 2.3 reveal a far more dramatic role for regularization than was the case for either $^{1}$H or $^{13}$C. Comparing HF and MP2, we see only a 2-fold reduction in RMSE for MP2, which is accompanied by only a 2-fold reduction in ME, together with a change of sign. This indicates significant overestimation of the correlation effects on $^{15}$N shielding by MP2. As a consequence, we observe error reduction at all $\kappa$'s, with strong regularization ($\kappa = 1.1$) being optimal. With strong regularization, RMSE is reduced dramatically from 38 ppm for MP2 to only 3.2 ppm, which an extraordinary 12-fold reduction in error. This result is nearly three times smaller error than CCSD, and approaches the high accuracy target we identified for $^{15}$N. Evidently a large part of the systematic overestimation of correlation effects by MP2 is removed by the strong regularization. This is confirmed by the very large 22-fold reduction in ME from 26 ppm to 1.2 ppm.

Why is strong regularization preferred for $^{15}$N whilst weak regularization is optimal for $^{13}$C? We cannot be entirely sure, although we note that the subset of molecules containing nitrogen nuclei includes many nitriles where N is triple bonded to its neighbor. It is reasonable to speculate that triply bonded molecules exhibit stronger correlation effects than singly bonded species, and this could be a plausible explanation for why stronger regularization leads to error reduction for $^{15}$N. However, each nitrile also has a $^{13}$C nucleus, and inspection of individual results (see Supporting Information) shows that the $^{13}$C shieldings in nitriles are generally better predicted via $\kappa$-MP2 with weak regularization than strong regularization. Furthermore chemical shieldings at sp$^3$ $^{15}$N are also better predicted with strong regularization. We provisionally conclude that the character of correlation effects on shielding at $^{15}$N are systematically different to $^{13}$C, and each is systematically different to $^{1}$H. Accordingly optimal regularization is different for each case.

For $^{17}$O, MP2 also performs rather poorly with slightly less than a 2-fold reduction in RMSE relative to HF (from 50.5 ppm to 27.4 ppm). The ME for MP2 changes sign relative to HF and is reduced by only a little more than a factor of two, to a value of 19 ppm. At first glance this appears similar to the case of $^{15}$N, and suggests the possibility of significant improvements due to regularization. However, no significant changes over MP2 are observed! Strong regularization ($\kappa = 1.1$) leads to slight improvements, from the RMSE of 27.4 ppm for MP2 to 24.5 ppm. The weaker regularizers lead to results that are essentially unchanged from MP2. Even strong regularization leads to only modest reductions in ME vs MP2, and the sign of the ME is the

same as MP2, indicating that correlation effects on $^{17}$O shieldings are *still* overestimated. It is clearly interesting to assess still stronger regularization for $^{17}$O shieldings in the future, as values of $\kappa \leq 1$ should apparently yield systematic improvements. From a regularized MP2 perspective, $^{17}$O shieldings appear to require extremely strong regularization, perhaps comparable to the values for transition metal energetics (e.g. $\kappa = 0.8$ was optimal for the MOR39 data set[80]).

Finally we consider the case of $^{19}$F magnetic shieldings. In this case, unregularized MP2 outperforms CCSD, and also exhibits a lower ME than CCSD. The improvement over HF is dramatic at the MP2 level: a five-fold reduction in ME, and a most impressive six-fold reduction in RMSE. There is also no change of sign in the ME from HF to MP2, suggesting that on average, correlation effects on $^{19}$F nuclei are not overestimated by MP2. This situation is unpromising for $\kappa$-MP2 because all regularization strengths provide some damping of correlation effects. Not surprisingly we indeed observe that all 3 finite regularization strengths in $\kappa$-MP2 lead to higher error, with $\kappa$ = 2 having the lowest RMSE of 9.7 ppm, versus 5.6 ppm for MP2 and 7.5 ppm for CCSD.

## Accuracy of MP3 nuclear shieldings for the Lutnæs Set

Following the detailed discussion of $\kappa$-MP2 discussed above, where regularization was used to damp the MP2 contribution to magnetic shieldings, we now turn to the inclusion of third order correlation as an alternative strategy to improve upon MP2. We will assess 3 models. First, as a reference calculation, standard MP3. Second, in light of its very good performance for non-covalent interactions, MP2.5, which damps the third order contribution by a factor of two. And finally, because MP2.X depends linearly on the third order contribution, we can optimize the fraction, either nucleus by nucleus, or for the overall dataset. Bearing in mind the significant increase in computational cost relative to MP2 or $\kappa$-MP2, we must also achieve significant improvements in accuracy in order to justify the additional computing expense.

Table 2.3 contains the shielding results obtained with the MP3, MP2.5 and MP2.X models. Considering all atoms, adding third order correction slightly improves upon MP2, by reducing the MP2 RMSE of 20.0 ppm to MP3 RMSE of 15.1 ppm. We also note that the ME changes sign from the MP2 value of +8.9 ppm to the MP3 value of -5.46 ppm, showing that the third order correction statistically overcorrects MP2 to yield an ME with the same sign as HF. Scaling with the optimal parameter of $\mathrm{X} = 0.6$ leads to remarkable improvements, with RMSE reduced to 3.3 ppm, and ME reduced to 0.72 ppm. This represents half the RMSE of CCSD. The traditional scaling of 0.5 was slightly less effective, giving an RMSE of 4.7 ppm and ME of 2.3 ppm. We can contrast these scaled MP3 results with the $\kappa$-MP2 data discussed above. In the $\kappa$-MP2 case, there was only modest improvement over MP2 by using either weak, medium or strong regularization, indicating that regularization strength was element-specific rather transferable. By contrast, these large improvements in scaled MP3 over MP3 across the entire dataset imply rather good transferability from one element to another. This is very encouraging.

Nevertheless, it may be possible to do even better by optimizing the scaling factor for each nucleus individually. So next we consider the individual nuclei, starting with $^{1}$H shieldings, where MP2 performed quite remarkably, and where regularization did not lead to any improvements.

Table 2.3: RIMP3 and scaled RIMP2.X aug-cc-pVTZ Isotropic NMR shielding statistics in ppm

| All | RIMP3 | RIMP2.5 | RIMP2.X | Scaling factor X |
|---|---|---|---|---|
| ME | -5.46 | 2.27 | 0.72 | 0.6 |
| RMSE | 15.1 | 4.67 | 3.29 | |
| $^{1}$H | | | | |
| ME | 0.07 | 0.002 | -0.013 | 0.4 |
| RMSE | 0.18 | 0.113 | 0.108 | |
| $^{13}$C | | | | |
| ME | -3.13 | -0.41 | -0.41 | 0.5 |
| RMSE | 4.35 | 1.15 | 1.15 | |
| $^{15}$N | | | | |
| ME | -17.58 | 4.90 | 0.41 | 0.6 |
| RMSE | 25.31 | 7.28 | 1.45 | |
| $^{17}$O | | | | |
| ME | -11.0 | 4.51 | 2.18 | 0.575 |
| RMSE | 21.29 | 6.22 | 4.98 | |
| $^{19}$F | | | | |
| ME | 2.78 | 3.63 | 3.29 | 0.7 |
| RMSE | 4.39 | 4.16 | 4.01 | |

Standard MP3 performs slightly worse than MP2, with RMSE of 0.19 ppm and ME of 0.09 ppm, compared to the RMSE of 0.13 ppm for MP2. This is consistent with the poor reputation of MP3. By contrast, with the optimal scaling factor of 0.4, the MP2.4 RMSE is reduced to 0.11 ppm, which is a slight improvement over MP2. In fact, MP2.4 approaches the high accuracy target of 0.1 ppm. Nonetheless, it is perhaps debatable as to whether the 0.02 ppm reduction in RMSE is justifiable for the higher cost of the third order correction.

For $^{13}$C shieldings, MP2's already very good performance was improved by weak regularization, reducing RMSE from 3.2 ppm to 2.0 ppm. By contrast, MP3 exhibits an RMSE of 4.4 ppm and ME of -3.13 ppm, which is slightly worse than MP2. This situation is significantly improved With the optimal scaling factor of 0.5, whereupon we observe RMSE reduction to 1.2 ppm and ME to -0.4 ppm. This is a quite impressive result, which is not only better than both MP2 and CCSD, but also approaches the high accuracy target of 1 ppm at far less compute cost than CCSD(T). Accordingly, we suggest that MP2.5 may be a potentially valuable first principles approach to $^{13}$C magnetic

shieldings.

For $^{15}$N shieldings, MP3 has quite large errors, with an RMSE of 25.3 ppm and a ME of -17.6 ppm. This however is still a useful improvement over MP2 (RMSE 38.0 ppm). With the optimal scaling factor of 0.6, RMSE is dramatically reduced to only 1.5 ppm, (ME reduced to 0.4 ppm). This is more than a factor of two better than the $\kappa$-MP2 with strong regularization (RMSE of 3.2 ppm), and well within the range of the high accuracy target of 3 ppm for $^{15}N$. It is an extraordinary six-fold improvement over CCSD itself, indicating that the optimal scaling is effectively accounting for a very significant fraction of the triples correlation effect by a simple and apparently very effective renormalization of the doubles. It appears that MP2.6 may be a very promising method for $^{15}$N magnetic shieldings.

Similarly for $^{17}$O shieldings, MP3 again has quite large errors, with RMSE of 21.3 ppm and ME of -11 ppm, which can be compared against the MP2 value of 27.4 ppm. With the optimal scaling factor of 0.575, RMSE reduced by a striking factor of four to 5.0 ppm, while the ME is reduced to 2.2. This represents close to a factor of two improvement over CCSD's 8.9 ppm RMSE, as well as approaching the 4 ppm high accuracy target. While not quite as strikingly high quality as the $^{15}$N shieldings, this improvement suggests that approximately half of the triples contribution to $^{17}$O shieldings is captured by optimally scaled MP3.

Lastly for $^{19}$F, where MP2 with RMSE 5.7 ppm outperformed CCSD's 7 ppm RMSE, and where regularization led to no improvements, unscaled MP3 has RMSE of 4.4 ppm and ME of 2.8 ppm. This is a particularly interesting case, because it is a rare outcome where standard MP3 already statistically performs better than CCSD. With optimal scaling factor of 0.7, we observe a slightly improved RMSE of 4.0 ppm, while the ME is slightly increased to 3.3 ppm. Relative to the CCSD RMSE of 7.5 ppm, we infer that almost half of the triples contribution to correlation (i.e. the CCSD error) is captured by the renormalization associated with scaled MP2.7. This is a quite similar outcome to the $^{17}$O and $^{13}$C shielding results.

Overall, considering $^{1}$H, MP2.4 gives a slight but meaningful improvement over MP2, albeit at a higher computational cost. Considering only heavy atoms, 0.6 seems to be a suitable scaling factor overall, giving greatly improved shieldings for $^{13}$C, $^{15}$N, and $^{17}$O, in all cases producing smaller RMSE than CCSD. The value of 0.6 is quite close to the *ad hoc* scaling factor of 0.5, and in all cases gives similar statistical performance. In addition, with optimal scaling factors, we observe very promising performance, where RMSE's have either neared or reached the respective high accuracy targets.

## 2.4 Conclusions

We have implemented and tested a method independent finite difference based software using gauge including atomic orbitals to calculate NMR shieldings. This finite field approach uses second order finite differences of the energy, as a function of an applied magnetic field and an applied nuclear spin to evaluate the magnetic shielding tensor. This required development of the appropriate complex matrix elements for magnetically perturbed Hamiltonian, as well as generalization of the corresponding self-consistent field and correlation energies to work with complex orbitals and

amplitudes. Once this infrastructure is in place, assessment of magnetic properties of non-standard methods may then be accomplished with much less effort than if it were necessary to implement analytic second derivatives. The finite difference approach is also naively parallelizable.

The MP2 method is to known to have excellent performance for $^{1}$H shieldings and good performance for $^{13}$C, but poor performance for $^{15}$N and $^{17}$O on normal closed shell molecules. This motivated us to assess the possibility of using regularized MP2 ($\kappa$-MP2)[80] to improve upon MP2. We tested on a set of 28 small molecules[93] that was used previously to benchmark approximate electronic structure methods for magnetic properties. The results showed that $\kappa$-MP2 could perform very well, provided element-specific regularization was used. For $^{1}$H zero regularization is optimal; for $^{13}$C, weak regularization reduces MP2 errors by 30%; for $^{15}$N, strong regularization reduces MP2 errors by a remarkable 90%, and for $^{19}$F zero regularization is optimal. For $^{13}$C and $^{15}$N, $\kappa$-MP2 with element-specific regularization yields errors that are smaller than CCSD.

On the other hand, testing on the same data set, we found that the scaled third order perturbative method MP2.X is the best performing among the methods we have inspected. Our results showed that with element-specific scaling factor $c$, MP2.X performs very well. For $^{1}$H $c = 0.3$ is optimal, reducing MP2 errors by almost 20%; for $^{13}$C, $c = 0.5$ reduces MP2 errors by more than 60%; for $^{15}$N, $c = 0.6$ reduces MP2 errors by an impressive 95%; for $^{17}$O, $c = 0.575$ reduces MP2 errors by a remarkable 80%, and for $^{19}$F, $c = 0.7$ reduces MP2 errors by close to 30%. For all those elements but $^{1}$H, MP2.X with element-specific scaling factor yields errors that are smaller than CCSD, making these approaches promising for chemical applications with accuracy significantly higher than is attainable with DFT.[51, 52]

The $\kappa$-MP2 and MP2.X results can be understood as regularization or partial inclusion of perturbative third order correction renormalizes the doubles amplitudes, partially accounting for the neglected triple and higher correlation effects.

In the near future, it would be interesting to apply the same framework to the regularized orbital optimized MP2 ($\kappa$-OOMP2) method, as well as MP3 evaluated with non-canonical orbitals such as $\kappa$-OOMP2 or Kohn-Sham orbitals. We intend to apply our framework to evaluate scalar couplings, and to complete our mixed analytical and numerical derivative implementation.

# Chapter 3

# Analytic First-Order Nuclear Derivatives for Regularized Orbital-Optimized MP2: Implementation and Applications

## 3.1 Introduction

The most important first order molecular property in quantum chemistry is the force on each atom, $-\mathrm{d}E/\mathrm{d}\mathbf{R}$, which enables walking on potential energy surfaces[112] via geometry optimization or trajectory calculations. The quality of structures and structure-related properties such as harmonic vibrational frequencies can then be assessed relative to other quantum chemistry models. There is a very large literature[113] on the results for widely used models which in turn informs the selection of the most appropriate methods to solve a target problem or set of problems.

Therefore any electronic structure method that can be viewed as potentially promising for large-scale applications requires an implementation of the analytic nuclear gradient. Today there are two alternatives. Automatic differentiation[114, 115] provides an implicit implementation of the gradient given sufficient information about how the code for the energy depends on parameters with relatively little human effort. AD has now been applied to nuclear derivatives in computational quantum chemistry.[116, 117] However, relative to the alternative of explicitly formulating and implementing the gradient within dedicated software for the gradient, AD typically involves a significant loss of efficiency.[115]

Within wavefunction-based theory, second order Møller-Plesset theory[118] (MP2) is very well-established as the simplest and computationally cheapest correction to mean field Hartree-Fock (HF) theory for electron correlation effects. With direct methods,[119, 120] and use of auxiliary basis sets,[121, 122] MP2 is computationally efficient for medium-sized molecules, with $O(N^5)$ scaling vs $O(N^6)$ and higher for coupled cluster theory.[56] On-going developments in low-scaling local MP2 methods make very large calculations feasible as well.[123, 124] From the perspective of chemistry, MP2 has been strikingly successful for some classes of problems, such as structures and frequencies of stable closed shell organic molecules,[125] and binding energies

and intermolecular interactions such as water clusters,[126–128], interaction of simple cations and anions with water.[129, 130], and ionic liquids[131, 132].

However MP2 is also well-known to break down in several domains. First, with unrestricted HF (UHF) orbitals, UMP2 is sensitive to spin-contamination.[12, 102] This means that UMP2 does not always perform well for radical species and bond separation processes, and even geometries and vibrational frequencies of radicals[133]. One way to overcome this problem is to optimize orbitals in the presence of MP2 correlation.[14, 134] Orbital-optimized MP2 (OOMP2) greatly reduces spin-contamination, and has been shown to improve geometries and frequencies of radicals.[134] Considerable work has been reported on efficient implementation and assessment of OOMP2.[135–138] Building on MP2 analytical gradient theory,[139–142] the analytical gradient for its opposite spin approximation,[143] and full OOMP2[144] have been available for some time.

A second failure mode of MP2 is that for systems with small HOMO-LUMO gaps, the MP2 correlation energy is not variationally bounded from below, and becomes too negative, diverging in the zero gap limit. This problem is exacerbated in OOMP2, where a variety of artifacts arise from decreasing the HF gap via orbital optimization.[10] Perhaps the most striking is the absence of a Coulson-Fischer point for spin-polarization;[10] OOMP2 suffers from false symmetry restoration in contrast to the the artificial symmetry breaking of HF theory.[145] To resolve this issue requires regularization of some type. Regularization via a level shift[10, 11] resolves the issue, but at the loss of significant accuracy. The most promising approach to date is "energy-dependent" regularization, as exemplified by the $\kappa$-OOMP2 approach[9, 111] (and of course the related $\kappa$-MP2 method[80, 146]). Other alternatives also exist.[80, 147, 148]

In $\kappa$-OOMP2, stability against non-variational collapse can be accompanied by significant *improvements* in accuracy,[9] although no single value of $\kappa$ is optimal for all classes of molecules[111]. As examples of significant improvement, large MP2 errors for dispersion-dominanted intermolecular interactions,[81] as well as transition metal binding energies[80] can be reduced by factors of 3-5 in the regularized $\kappa$-OOMP2 method at negligible additional compute cost relative to OOMP2 itself. $\kappa$-OOMP2 has also been successfully applied to treat biradicaloids, when combined with spin projection.[149]

In light of the exciting data reported for $\kappa$-OOMP2 relative energies, it is clearly desirable to have available the analytic nuclear gradient to enable broader assessment and usage of the method. That is the first purpose of the present paper, where the nuclear gradient formulation, and pertinent details of the implementation are summarized in the section on Theory. A variety of applications are then reported to highlight areas where usage of the $\kappa$-OOMP2 forces may be advantageous relative to either MP2 or OOMP2.

## 3.2 Theory

The derivation of the RI-MP2 nuclear gradient is well established.[139, 141, 142] Furthermore, the relatively small modifications to the RI-MP2 nuclear gradient expression for performing OOMP2 were well documented by Lochan *et al*.[134] Therefore, we briefly recapitulate the RI-MP2 and OOMP2 nuclear gradient expressions here and highlight some additional modifications necessary to

incorporate the $\kappa$-regularizer.[9] We use $\{i, j, k, l, \cdots\}$ for occupied orbitals, $\{a, b, c, d, \cdots\}$ for virtual (unoccupied) orbitals, $\{P, Q, R, S, \cdots\}$ for auxiliary (RI) basis functions, and $\{\mu, \nu, \lambda, \sigma, \cdots\}$ for atomic orbitals,

## RI-MP2 and OOMP2 Nuclear Gradient

The MP2 correlation energy is written in the real-valued spin-orbital basis as

$$E_c^{\text{RI-MP2}} = -\frac{1}{4}\sum_{ijab} t_{ij}^{ab}\langle ij||ab\rangle_{\text{RI}} = -\frac{1}{4}\sum_{ijab}\frac{|\langle ij||ab\rangle_{\text{RI}}|^2}{\Delta_{ij}^{ab}} \tag{3.1}$$

where $\Delta_{ij}^{ab} = \epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j$. Following standard practice[150], we write the nuclear gradient of the RI-MP2 correlation energy as

$$\frac{\mathrm{d}E_c^{\text{RI-MP2}}}{\mathrm{d}\mathbf{R}} = \sum_{pq} P_{pq}^{(2)} F_{pq}^{(\mathbf{R})} + \sum_{pq} W_{pq}^{(2)} S_{pq}^{(\mathbf{R})} + 2\sum_{\mu\nu Q}(\mu\nu|Q)^{(\mathbf{R})}\Gamma_{\mu\nu}^{Q} - \sum_{PQ}(P|Q)^{(\mathbf{R})}\Gamma_{PQ} \tag{3.2}$$

where we used the superscript $(\mathbf{R})$ to denote the derivative of the corresponding quantity with respect to $\mathbf{R}$. $\mathbf{F}$ and $\mathbf{S}$ denote the Fock matrix and the atomic-orbital overlap matrix, respectively. It is worth defining individual terms further so that additional terms arising due to the $\kappa$-regularizer can be clearly identified later. We first define the second-order correction to the one-particle density matrix (1-PDM):

$$P_{ik}^{(2)} = -\frac{1}{2}\sum_{jab} t_{ij}^{ab} t_{kj}^{ab} \tag{3.3}$$

$$P_{ac}^{(2)} = \frac{1}{2}\sum_{ijb} t_{ij}^{ab} t_{ij}^{cb} \tag{3.4}$$

$$P_{ai}^{(2)} = Z_{ai} \tag{3.5}$$

where $Z_{ai}$ is the z-vector[63] which can be obtained by solving the following linear equation

$$\sum_{jb} A_{ai,bj} Z_{bj} = \frac{\partial E_{\text{RI-MP2}}}{\partial \theta_{ai}} \tag{3.6}$$

with $\mathbf{A}$ being the HF orbital Hessian. Here, we used $\theta_{ai}$ to denote the orbital rotation parameter between a virtual orbital $a$ and an occupied orbital $i$. The second-order correction to the energy-weighted 1-PDM is

$$W_{ik}^{(2)} = -\frac{1}{2}\sum_{jab}\langle kj||ab\rangle_{\text{RI}} t_{ij}^{ab} - \frac{1}{2}P_{ik}^{(2)}(\epsilon_i + \epsilon_k) - \frac{1}{2}\sum_{pq} P_{pq}^{(2)} A_{pq,ki} \tag{3.7}$$

$$W_{ac}^{(2)} = -\frac{1}{2}\sum_{ijb}\langle ij||cb\rangle_{\text{RI}} t_{ij}^{ab} - \frac{1}{2}P_{ac}^{(2)}(\epsilon_a + \epsilon_c) \tag{3.8}$$

$$W_{aj}^{(2)} = -\frac{1}{2}\sum_{ikb}\langle ik||jb\rangle_{\text{RI}} t_{ik}^{ab} - P_{ai}^{(2)}\epsilon_i - \frac{1}{2}\sum_{pq} P_{pq}^{(2)} A_{pq,aj} \tag{3.9}$$

Now, the two-particle density matrix (2-PDM) contributions are defined by

$$\Gamma^Q_{ia} = \sum_{jb} t^{ab}_{ij} C^Q_{jb} \tag{3.10}$$

$$\Gamma^Q_{\mu\nu} = \sum_{ia} C_{\mu i} C_{\nu a} \Gamma^Q_{ia} \tag{3.11}$$

$$\Gamma_{PQ} = \sum_{ia} C^P_{ia} \Gamma^Q_{ia} \tag{3.12}$$

where

$$C^P_{ia} = \sum_Q (ia|Q)(Q|P)^{-1} \tag{3.13}$$

Orbitals are no longer canonical in OOMP2, which leads to an extra term in the energy-weighted 1PDM. This extra term only modifies the virtual-occupied block of $\mathbf{W}^{(2)}$,

$$\widetilde{W}^{(2)}_{aj} = W^{(2)}_{aj} - \sum_i F_{ai} P^{(2)}_{ij} - \sum_b F_{jb} P^{(2)}_{ba} \tag{3.14}$$

Moreover, the RI-MP2 Lagrangian (i.e., the right hand side of Eq. (3.6)) is zero at the SCF convergence. Therefore, there is no need for solving the Z-vector equation and $P^{(2)}_{ai} = 0$. These changes complete the implementation of the OOMP2 nuclear gradient.

## $\kappa$-OOMP2 Nuclear Gradient

In $\kappa$-MP2, the MP2 correlation energy is modified into the following form:

$$E_c^{\kappa\text{-RI-MP2}} = -\frac{1}{4} \sum_{ijab} \bar{t}^{ab}_{ij} \langle ij||ab\rangle_{\text{RI}} \left(1 - e^{-\kappa \Delta^{ab}_{ij}}\right) = -\frac{1}{4} \sum_{ijab} \frac{|\langle ij||ab\rangle_{\text{RI}}|^2}{\Delta^{ab}_{ij}} \left(1 - e^{-\kappa \Delta^{ab}_{ij}}\right)^2 \tag{3.15}$$

where we defined the regularized amplitudes to be

$$\bar{t}^{ab}_{ij} = \frac{\langle ij||ab\rangle_{\text{RI}}}{\Delta^{ab}_{ij}} \left(1 - e^{-\kappa \Delta^{ab}_{ij}}\right) \tag{3.16}$$

We then define the regularized second-order correction to the 1-PDM:

$$\overline{P}^{(2)}_{ik} = -\frac{1}{2} \sum_{jab} \bar{t}^{ab}_{ij} \bar{t}^{ab}_{kj} \tag{3.17}$$

$$\overline{P}^{(2)}_{ac} = \frac{1}{2} \sum_{ijb} \bar{t}^{ab}_{ij} \bar{t}^{cb}_{ij} \tag{3.18}$$

$$\overline{P}^{(2)}_{ai} = \overline{Z}_{ai} \tag{3.19}$$

where $\overline{Z}_{ai}$ is now the z-vector which is be obtained by solving the modified linear equation

$$\sum_{jb} A_{ai,bj} \overline{Z}_{bj} = \frac{\partial E_{\kappa\text{-RI-MP2}}}{\partial \theta_{ai}} \tag{3.20}$$

The regularized second-order correction to the energy-weighted 1-PDM is

$$\overline{W}^{(2)}_{ik} = -\frac{1}{2} \sum_{jab} \langle kj||ab\rangle_{\text{RI}} \left(1 - e^{-\kappa \Delta^{ab}_{kj}}\right) \bar{t}^{ab}_{ij} - \frac{1}{2} \overline{P}^{(2)}_{ik} (\epsilon_i + \epsilon_k) - \frac{1}{2} \sum_{pq} \overline{P}^{(2)}_{pq} A_{pq,ki} \tag{3.21}$$

$$\overline{W}^{(2)}_{ac} = -\frac{1}{2} \sum_{ijb} \langle ij||cb\rangle_{\text{RI}} \left(1 - e^{-\kappa \Delta^{cb}_{ij}}\right) \bar{t}^{ab}_{ij} - \frac{1}{2} \overline{P}^{(2)}_{ac} (\epsilon_a + \epsilon_c) \tag{3.22}$$

$$\overline{W}^{(2)}_{aj} = -\frac{1}{2} \sum_{ikb} \langle ik||jb\rangle_{\text{RI}} \left(1 - e^{-\kappa \Delta^{jb}_{ik}}\right) \bar{t}^{ab}_{ik} - \overline{P}^{(2)}_{ai} \epsilon_i - \frac{1}{2} \sum_{pq} \overline{P}^{(2)}_{pq} A_{pq,aj} \tag{3.23}$$

For $\kappa$-OOMP2, the extra term in the energy-weighted 1PDM of the virtual-occupied block of $W^{(2)}$ is

$$\widetilde{\overline{W}}^{(2)}_{aj} = \overline{W}^{(2)}_{aj} - \sum_{i} F_{ai} \overline{P}^{(2)}_{ij} - \sum_{b} F_{jb} \overline{P}^{(2)}_{ba} \tag{3.24}$$

## 3.3 Computational Details

The theory described above has been implemented in the Q-Chem program package[104]. This implementation is fully in-core, and avoids storage of the quartic four-center two-electron integrals and amplitudes. Instead, storage is dominated by 3-center quantities, including $(ia|P)$, $C^P_{ia}$, $\Gamma^P_{ia}$, $\Gamma^P_{\mu\nu}$, etc. All computationally significant steps are OpenMP parallelized. The overall computational cost of the $\kappa$-OOMP2 analytical gradient is dominated by $\mathcal{O}(O^2V^3)$ computational steps ($O$ is the number of occupied levels, $V$ is the number of virtual orbitals) in the large-molecule limit (just like the MP2 gradient).

As already mentioned in the Introduction, the value of $\kappa$ used for regularization in OOMP2 is a parameter that must be selected. In the applications reported below, we will use $\kappa = 1.45$, as recommended originally[9] based on results for the W4-11 benchmark set [151]. We will also consider $\kappa = 1.1$, corresponding to significantly stronger regularization (i.e. correlation is more strongly damped), as has proven useful in treating strong dispersion interactions and transition-metal complexes.[111] And we will also use $\kappa = 2.0$, corresponding to significantly weaker regularization (i.e. correlation is less strongly damped) in order to probe the effects of regularization strength.

## 3.4 Results and Discussion

We assess the numerical performance of $\kappa$-OOMP2 with four applications of increasing complexity: methanol cation radical torsional surface, weak transannular chalcogen···chalcogen interactions, small molecule vibrational frequencies, and adiabatic singlet-triplet gap C60 fullerene.

## Methanol cation torsional and bond stretch surfaces

Molecular ions are ubiquitous as reactive intermediates, for instance created by photoionization in the interstellar medium, or even in mass spectral analysis of stable molecules. Standard quantum chemical methods, such as MP2 theory, have enjoyed long-standing success in predicting the structures of molecular ions, even though such structures can be far more diverse and less predictable than those for stable molecules. However, there are certainly cases where MP2 theory fails to perform at its usual level of accuracy, for instance because the potential energy surface is relatively flat, or because a correlation effect is exaggerated by the simple MP2 approach. The methanol cation, $CH_3OH^{+\cdot}$, is one such case, where Gauld et al previously found that MP2 predicts an artificially short C-O bond length[152]. They reported that CCSD(T)/6–311G(df,p) predicted a C$-$O bond length of 1.370 Å at the eclipsed conformation, while MP2 predicted a C$-$O bond length of less than 1.3 Å. This error is roughly 5-10 times the usual quite good accuracy of computed MP2 bond lengths in large basis sets. We note in passing that the global minimum on the $[C,H_4,O]^+$ PES is actually the $CH_2-OH_2^{+\cdot}$ structure.[153]

In an effort to reproduce the above results, we found 3 local $CH_3OH^{+\cdot}$ minima with MP2: the global minimum structure is the eclipsed short C-O bond $C_s$ structure; a local minimum staggered short C-O bond $C_s$ structure; and a staggered normal C-O bond $C_s$ structure, which is a saddle point with 1 torsional imaginary frequency. We didn't locate any $C_1$ structures that were local minima with MP2 in the cc-pVTZ basis set. This corroborates the findings of Gauld et al, the $C_1$ eclipsed long C-O bond structure seems to be an artifact of small basis sets.

To generate a reference potential energy surface, we computed H-O-C-H torsional scan and C-O bond stretch scan with CCSD/cc-pVTZ, followed by CCSD(T)/cc-pVTZ energy single points. We conducted the same procedure with HF, OOMP2, $\kappa$-OOMP2 with $\kappa$=1.1, 1.45 at the same basis set. These results are presented in Figure 3.1.

First, we observe that HF predicts too long C-O bond lengths, around 1.47 Å, compared to around 1.375 Å with CCSD. The H-O-C-H torsional surface displays opposite peaks and troughs: energy minima is at the staggered geometry and saddle points are at the eclipsed geometry. This is in contrast to CCSD(T), which predicts minima at the eclipsed geometry, and saddle points at staggered geometry. This suggests that HF might not be a good reference for perturbative treatment.

Second, we observe the dramatic shortened C-O bond in MP2, predicting around 1.29 Å at the eclipsed geometry. This corresponds to the eclipsed short C-O bond $C_s$ structure global minimum referenced above. It also displays two electronic surfaces, shown by discontinuities in both torsional energy surface and bond length. In addition, the torsional barrier is dramatically overestimated: while CCSD(T) predicts around 0.5 kcal/mol, MP2 predicts around 2.5 kcal/mol, representing a 5 fold error.

Third, we observe that orbital optimization restores the torsional energy surface: the minima bond length with $\kappa$-OOMP2 is predicted to be 1.34 Å, and non-regularized OOMP2 predicted 1.35 Å, representing much closer values to CCSD prediction of 1.375 Å than MP2's 1.29 Å. The torsional barrier is also decreased: $\kappa$-OOMP2 predicts around 0.8 kcal/mol, while OOMP2 predicts around 0.75 kcal/mol. This represents an almost 8 fold decrease from MP2's error. We also note that for this system, regularization strength minimally affects the qualitative results.

We note in passing an interesting feature in the CCSD surface: a flat energetic region around the eclipsed minima. This was not observed for the unregularized and regularized variants of OOMP2 nor CCSD(T).

The dual energy surface in MP2 is an interesting artifact. We present an analysis of this artifact in Figure 3.2, which shows the C-O stretch profile computed at MP2/cc-pVTZ, $\kappa$-OOMP2/cc-pVTZ, and CCSD/cc-pVTZ; and corresponding C-H bond lengths, which is cis to OH in eclipsed configuration and trans to OH in staggered configuration. In the MP2 panel, for both staggered and eclipsed configuration respectively, there exist two distinct energy surfaces after a Coulson-Fischer-like point, corresponding to smooth C-H bond lengthening or abrupt shortening. For the staggered surface, the short C-H long C-O is the torsional saddle point, meanwhile there is no other local minima on the eclipsed surface. This is in contrast to the CCSD and $\kappa$-OOMP2 surfaces, which display no such divergences: smooth shortening of the C-H bond coinciding with the C-O bond lengthening is observed instead. We speculate that the long C-H surface in MP2 lowered in energy, resulting in overlap with the proper minimum surface with the short C-H bond.

## Intramolecular chalcogen interactions

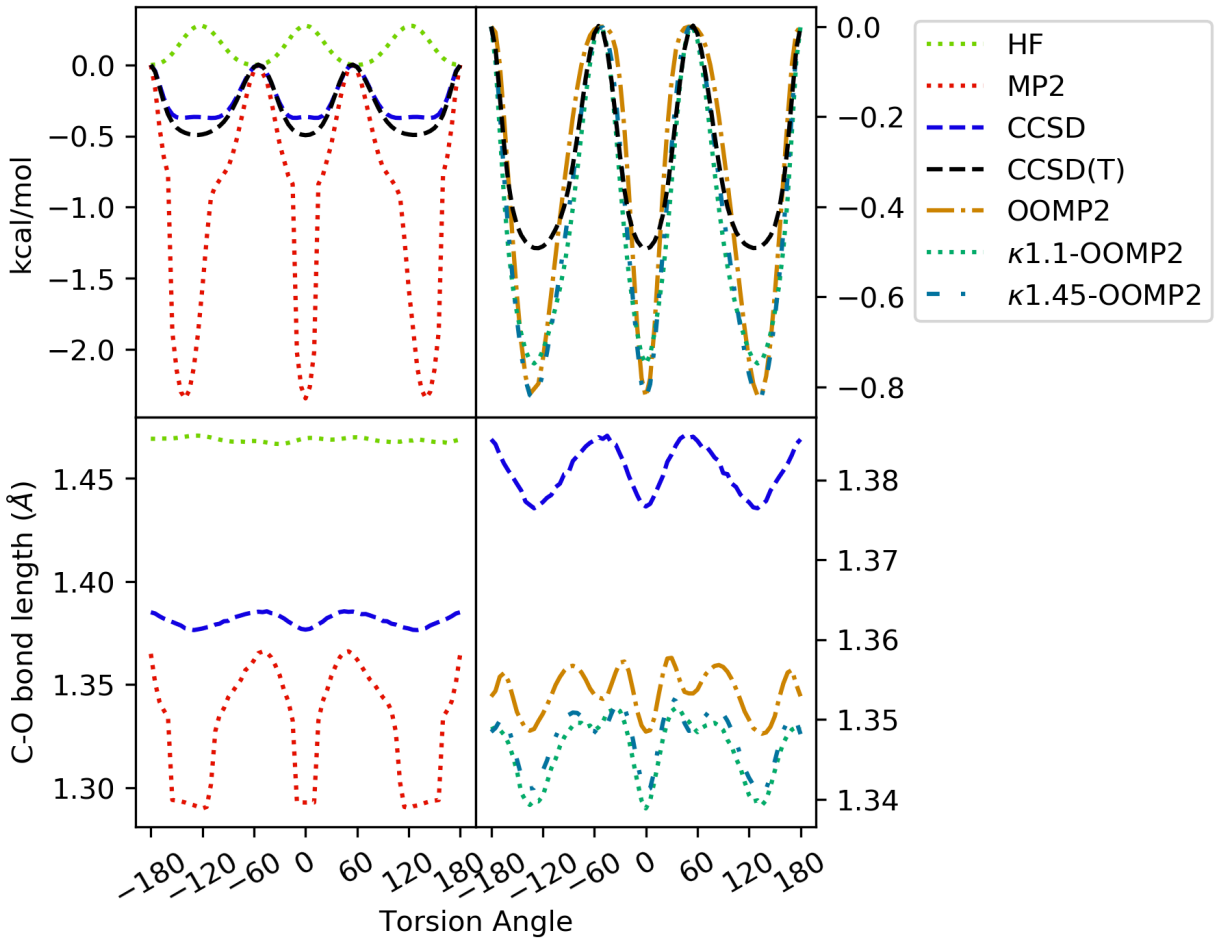
The nature of transannular chalcogen interactions in heterocyclic molecules has been in debate since the 70's. The presense of the weak chalcogen-chalcogen interaction can not be explained by simple molecular orbital theory. This class of molecules has recently received renewed interest. Moilanen and coworkers reported the importance of correlation description in the intramolecular cross ring chalcogen bond in tetrachalcogen tetranitrides, $E_4N_4$, in the cage form, and octachalcogen dications $E_8^{2+}$ where E = S,Se[154]. Fig. 3.3 shows a graphical depiction of these two classes molecules. In particular correlation weakens and lengthens the cross-ring E•••E bond. Hartree-Fock underestimates the E•••E distances in E4N4 and particularly in $E_8^{2+}$, whereas MP2 predicts too large E•••E distances, predicting no interaction between the two chalcogen centers. This was a result of MP2 overestimating correlation contribution. On the other hand, DFT results were highly dependent on the chosen functional and the amount of exact exchange employed. Therefore it is interesting to see whether regularized OOMP2 provides a good description of this class of molecules.

We present bond lengths, angles, and E•••E distances results in Table 3.1. First, comparing to CCSD(T), we observe the too large E•••E distances predicted by MP2 and the too short E•••E distances predicted by HF. With $\kappa$-OOMP2 and OOMP2, we observe that OOMP2 reduced E•••E distance RMSE from 0.89 Å in MP2 to 0.49 Å. Applying regularization with $\kappa$=2 results in further reduction to 0.36 Å. We also observe a larger error for the two $E_8^{2+}$ molecules. This might be related to the multi-reference nature of these two molecules, as suggested by Moilanen and coworkers. There also seems to be a trade-off between accuracy in bond lengths and angles, where stronger regularization gave more accurate angles but less accurate bond lengths.

## Harmonic vibrational frequencies of small molecules

As described in the introduction, MP2 for closed shell systems in general performs pretty well for vibrational frequencies[155]. However, for open shell systems, the performance of UMP2 is quite

Figure 3.1: Methanol cation radical $CH_3OH^{+\cdot}$ H-O-C-H Torsional energy surface (top left and top right) and corresponding C-O bond length profile (bottom left and bottom right). Results were split into left and right panels for clarity.




poor[133], which can be attibuted to the symmetry broken UHF reference, which is not longer an eigenfunction of the $\hat{S}^2$ operator. $\kappa$-OOMP2 offers an attractive way to resolve this problem. We assess the performance of $\kappa$-OOMP2 with two molecular vibrational frequency data sets. The first data set, compiled by Lochan and coworkers[156], contains 15 closed shell molecules and 18 open shell molecules, with experimental data as reference. The 15 closed shell molecules are: $N_2$, CO, HF, $Li_2$, LiF, LiH, BH, $CH_4$, $NH_3$, $H_2O$, $C_2H_2$, HCN, HNC, $H_2CO$, and $CO_2$, with 38 vibrational frequencies in total. The 18 open-shell diatomic molecules are: NO, $CO^+$, $O_2^+$, CN, OF, $N_2^+$, $F_2^+$, CH, OH, $HF^+$, CF, BO, $NH^+$, $C_2^-$, $BH^+$, LiO, NF, and NH. For the rest of this section, we will denote the closed shell subset by Lochan-CS, and the open shell subset by Lochan-OS. The second data set, constructed by Bozkaya and coworkers[137], contains 5 open shell molecules that suffer from artifactual symmetry-breaking problems: cis-$H_2O_2^+$, trans-$H_2O_2^+$, $LiO_2$, $C_3^+$, and $NO_2$. It uses

Figure 3.2: Methanol cation radical $CH_3OH^{+\cdot}$ C-O stretch energy surface and C-H bond length profile, computed with MP2, $\kappa$-OOMP2, and CCSD.

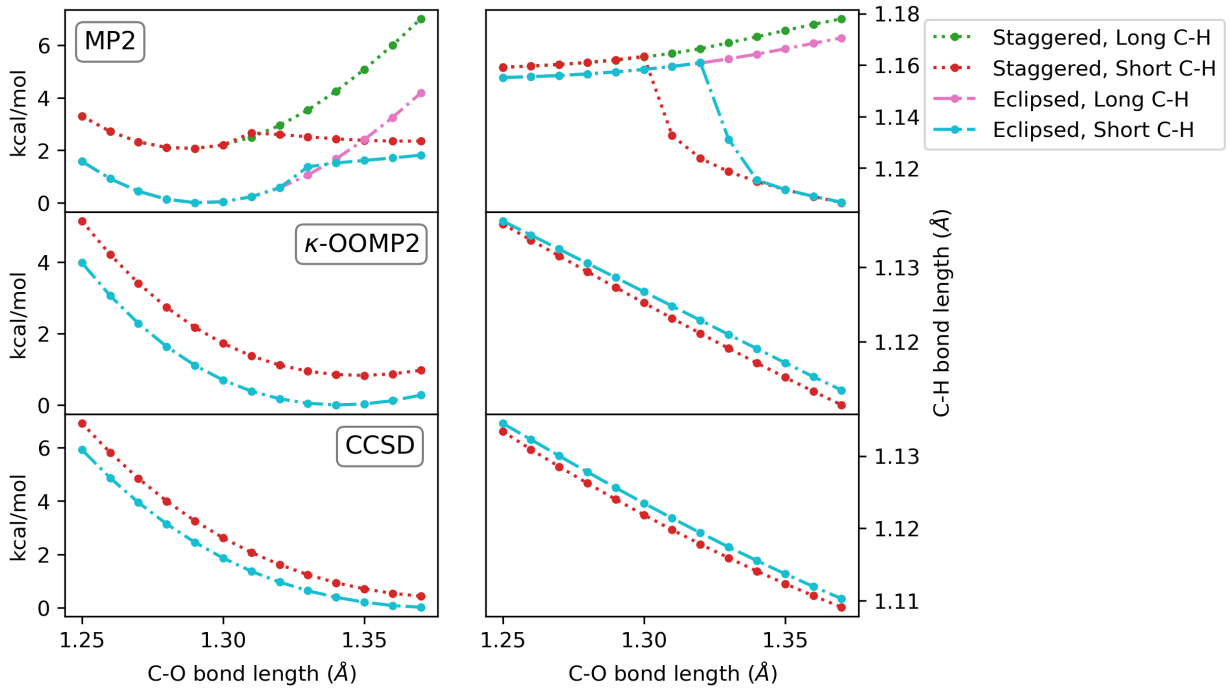



Figure 3.3: Graphical depiction of $E_4N_4$ in cage configuration (left) and $E_8^{2+}$ (right). Blue spheres represent nitrogen atoms, yellow spheres represent chalcogen atoms. Dashed lines represent E•••E distances in question

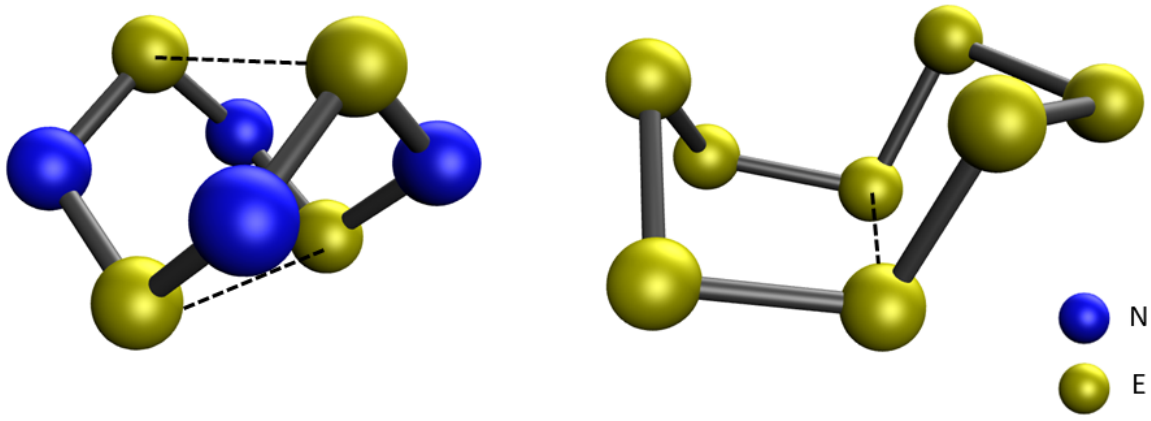


Table 3.1: $E_4N_4$ and $E_8^{2+}$ bond lengths and angles with various methods, optimized with the cc-pVTZ basis set.

| | HF | | $\kappa-$OOMP2 | | OOMP2 | MP2 | CCSD(T) |
|---|---|---|---|---|---|---|---|
| | | $\kappa$ =1.1 | 1.45 | 2 | | | |
| $S_4N_4$ S•••S | 2.451 | 2.477 | 2.525 | 2.599 | 3.048 | 3.047 | 2.676 |
| $Se_4N_4$ Se•••Se | 2.649 | 2.641 | 2.677 | 2.737 | 2.85 | 3.16 | 2.811 |
| $S_8^{2+}$ S•••S | 2.251 | 2.401 | 2.513 | 2.644 | 3.17 | 3.344 | 2.899 |
| $Se_8^{2+}$ Se•••Se | 2.492 | 2.544 | 2.604 | 2.695 | 3.103 | 3.518 | 2.933 |
| RMSE | | | | | | | |
| E•••E | 0.8314 | 0.6840 | 0.5459 | 0.3648 | 0.4923 | 0.8943 | |
| Bond length | 0.0888 | 0.1632 | 0.1376 | 0.1141 | 0.1094 | 0.0773 | |
| Angle | 8.71 | 6.33 | 7.55 | 9.08 | 9.50 | 12.72 | |

CCSD(T)/cc-pCVQZ vibrational frequencies as reference. We denote this data set by Bozkaya. We present statistical results in Table 3.2.

First, considering the Lochan-CS set, we observe that MP2 significantly improves upon HF: RMSE reduced by more than half from 177.57 ppm to 73.83 ppm, MAE reduced from 156.22 ppm to 41 ppm. We note that the RMSRE in MP2 is dominated by the error in $Li_2$, removing which would give RMSRE of 0.021. Orbital optimization did not lead to much improvement, with RMSE reduced slightly to 64.88 ppm and MAE increased slightly to 45.6 ppm. Regularization on MP2 causes higher errors. With $\kappa$ = 1.1, RMSE increased to 84.96 ppm, and MAE increased to 73.63 ppm. This is expected because MP2 errors were already lower than that of HF, regularization then would not lead to much improvement, if at all. In the same vein, regularization on OOMP2 with both $\kappa$ = 1.1 and 1.45 deteriorated all statistical error measures. This result can be observed visually in Figure 3.4, where experimental vibrational frequencies are plotted against calculated vibrational frequencies. The OOMP2, $\kappa$-OOMP2, and $\kappa$-MP2 data points lie outside the MP2 data points.

Moving to open shell molecules, for the Lochan-OS set, as expected MP2 performs quite poorly, giving much larger errors than the Lochan-CS set. Both RMSE and ME are almost 3 times of HF's. Applying orbital optimization led to significant improvement. RMSE and MAE halfed, from 560.60 ppm to 256.59 ppm and from 384.49 ppm to 163.61 ppm respectively. ME reduced significantly from 372.09 ppm to -77.79 ppm. However RMSE is still above that of HF. Regularization with $\kappa$ = 1.1 on MP2 only led to improvment in ME, from 372.09 ppm to 75.87 ppm. There were no improvements in RMSE and MAE. On the other hand, regularization with both $\kappa$ = 1.1 and 1.45 on OOMP2 led to improvements. Particularly, with $\kappa$ = 1.45, RMSE reduced from 256.59 ppm to 128.22 ppm, which is below that of HF. MAE also reduced from 163.61 ppm to 105.26 ppm. These results can be observed visually in Figure 3.5. MP2 and $\kappa$-MP2 data points lie significantly outside the diagonal, mostly on the right, signifying over estimation over experimental values. In contrast, $\kappa$-OOMP2 data points cluster much closer to the diagonal.

A similar trend shows up in the Bozkaya set. MP2 performed even more poorly than for the Lochan-OS set, dominated by a particularly large MaxE for the $6^{th}$ normal mode in cis-$H_2O_2^+$.

$\kappa$-OOMP2 with $\kappa$ = 1.45 greatly reduced errors. Interestingly, compared to the Lochan-OS set, regularization alone did lead to improved errors. This can be partially attributed to $\kappa$-MP2 resolving the large error in the $6^th$ normal mode in cis-$H_2O_2^+$. $\kappa$-MP2 also had lower errors than those of HF. Figure 3.6 shows visually the improvement of $\kappa$-MP2 and $\kappa$-OOMP2. It also shows OOMP2 underestimates vibrational frequencies for this data set.

Overall, $\kappa$-OOMP2 significantly improved the accuracy of open shell moleucles vibrational frequencies. Regarding regularization strength, $\kappa$-OOMP2 leans towards weaker regularization for vibrational frequency calculations. $\kappa = 1.6$ gives best performance in the two open-shell data sets, beating all other methods; for the closed shell set, even weaker regularization of $\kappa$ = 2 was preferred, giving the closest results to experimental data.

Figure 3.4: Graphical representation of the distribution of errors in vibrational frequencies of the Lochan-CS data set. Values referenced against experimental data.

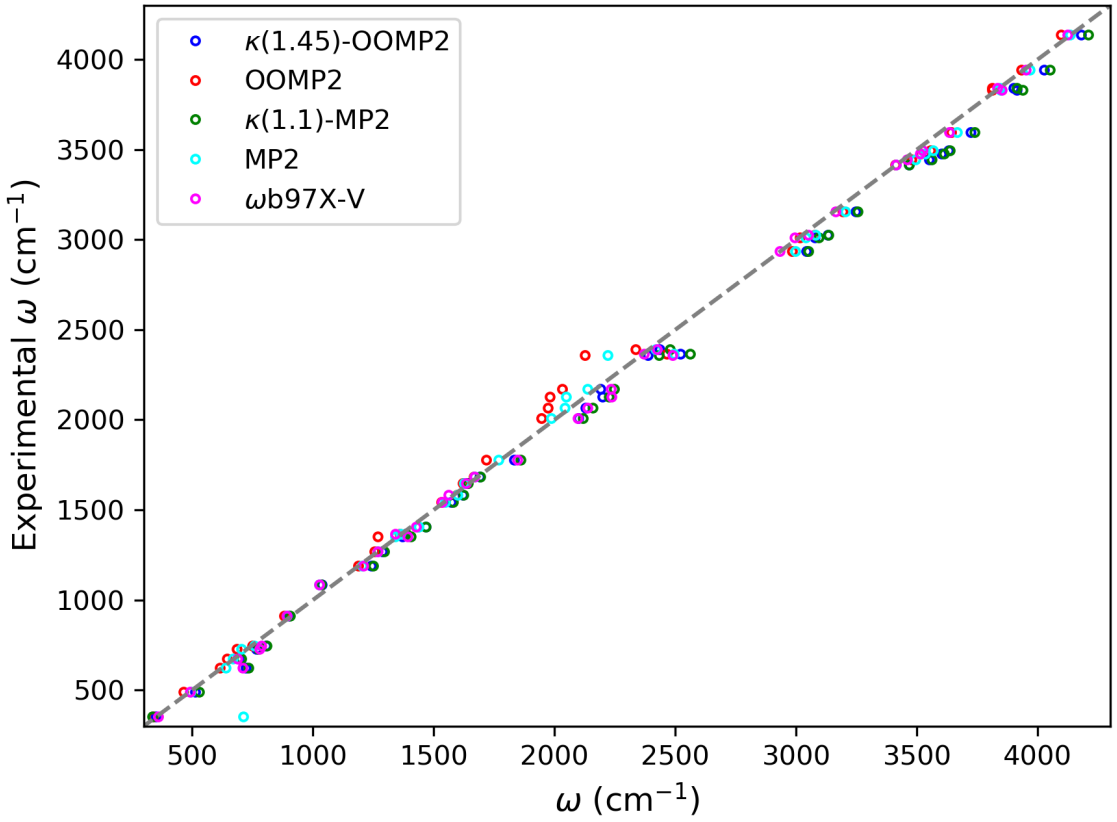
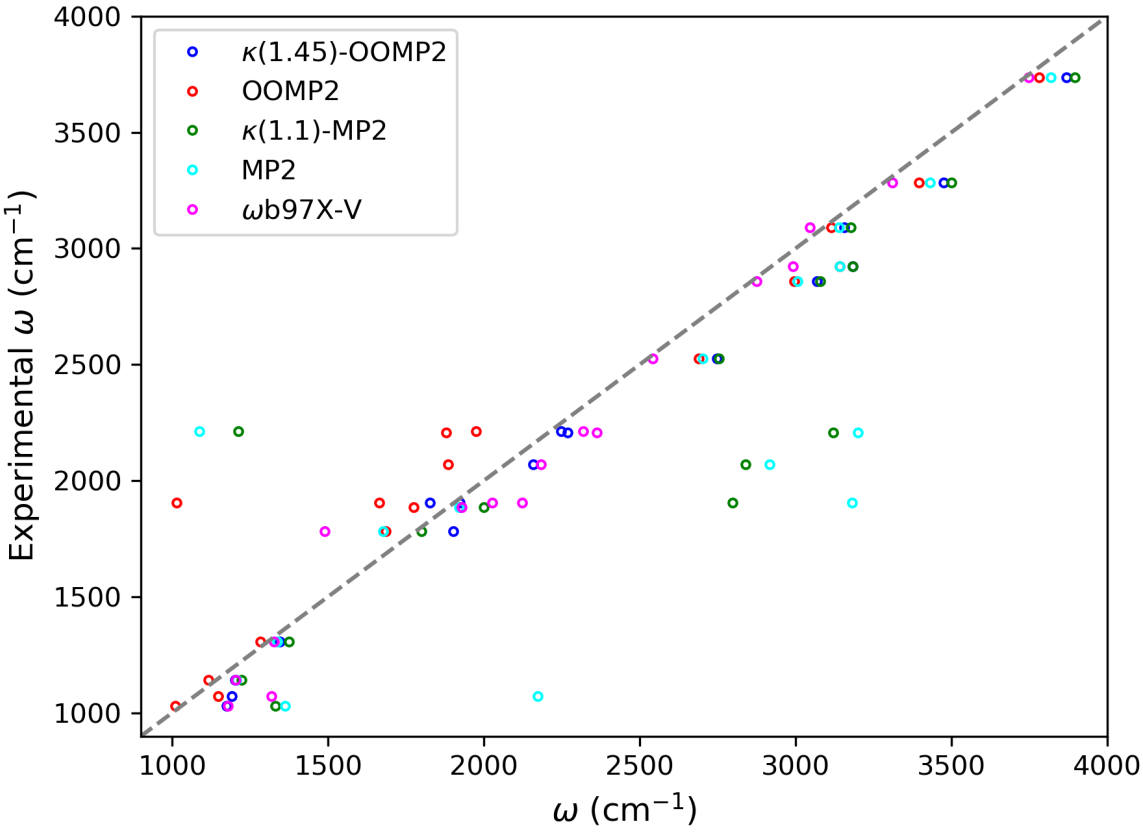


Table 3.2: RMSE, RMSRE, ME, MRE, MAE, MARE, MaxE for the Lochan closed shell and open shell subsets and the Bozkaya dataset. RMSE, ME, MAE, and MaxE presented with units of $cm^{-1}$

| | | $\kappa$-OOMP2 | | | | |
|---|---|---|---|---|---|---|
| | HF | $\kappa$ = 1.1 | 1.45 | OO-MP2 | $\kappa$1.1-MP2 | MP2 |
| Lochan-CS | | | | | | |
| RMSE | 177.57 | 111.57 | 70.61 | 64.88 | 84.96 | 73.83 |
| RMSRE | 0.101 | 0.063 | 0.041 | 0.034 | 0.049 | 0.168 |
| ME | 155.39 | 94.87 | 54.83 | -21.68 | 69.40 | 15.79 |
| MRE | 0.082 | 0.047 | 0.028 | -0.017 | 0.034 | 0.027 |
| MAE | 156.22 | 97.94 | 58.67 | 45.6 | 73.63 | 41 |
| MARE | 0.085 | 0.052 | 0.031 | 0.026 | 0.040 | 0.043 |
| MaxE | 367.56 | 204.28 | 153.01 | -233.57 | 193.35 | 361.13 |
| MaxE normal mode | $N_2$ | BH | BH | $N_2$ | BH | $Li_2$ |
| Lochan-OS | | | | | | |
| RMSE | 194.81 | 183.62 | 128.22 | 256.59 | 559.6 | 560.60 |
| RMSRE | 0.095 | 0.102 | 0.063 | 0.129 | 0.304 | 0.350 |
| ME | 125.41 | 169.03 | 96.69 | -77.79 | 75.87 | 372.09 |
| MRE | 0.06 | 0.088 | 0.047 | -0.045 | 0.021 | 0.213 |
| MAE | 168.95 | 169.03 | 105.26 | 163.61 | 413.87 | 384.49 |
| MARE | 0.082 | 0.088 | 0.052 | 0.08 | 0.219 | 0.220 |
| MaxE | 315.1 | 295.89 | 261.21 | -889.72 | -1470.65 | 1275.0 |
| MaxE normal mode | $O_2^+$ | $NH^+$ | $NH^+$ | $O_2^+$ | $O_2^+$ | NO |
| Bozkaya | | | | | | |
| RMSE | 235.00 | 142.83 | 91.09 | 158.85 | 186.71 | 840.23 |
| RMSRE | 0.227 | 0.146 | 0.093 | 0.133 | 0.185 | 0.296 |
| ME | 186.15 | 90.21 | 29.81 | -113.12 | 134.05 | 341.77 |
| MRE | 0.143 | 0.075 | 0.027 | -0.085 | 0.109 | 0.156 |
| MAE | 202.43 | 118.52 | 69.44 | 130.56 | 150.6 | 383.56 |
| MARE | 0.164 | 0.103 | 0.062 | 0.106 | 0.125 | 0.19 |
| MaxE | 565.56 | 320.94 | 235.78 | 311.37 | 462.96 | 2941.75 |
| MaxE normal mode | $C_3^+$,1 | $C_3^+$,1 | $C_3^+$,2 | $NO_2$,2 | $C_3^+$,1 | cis-$H_2O_2^+$,6 |

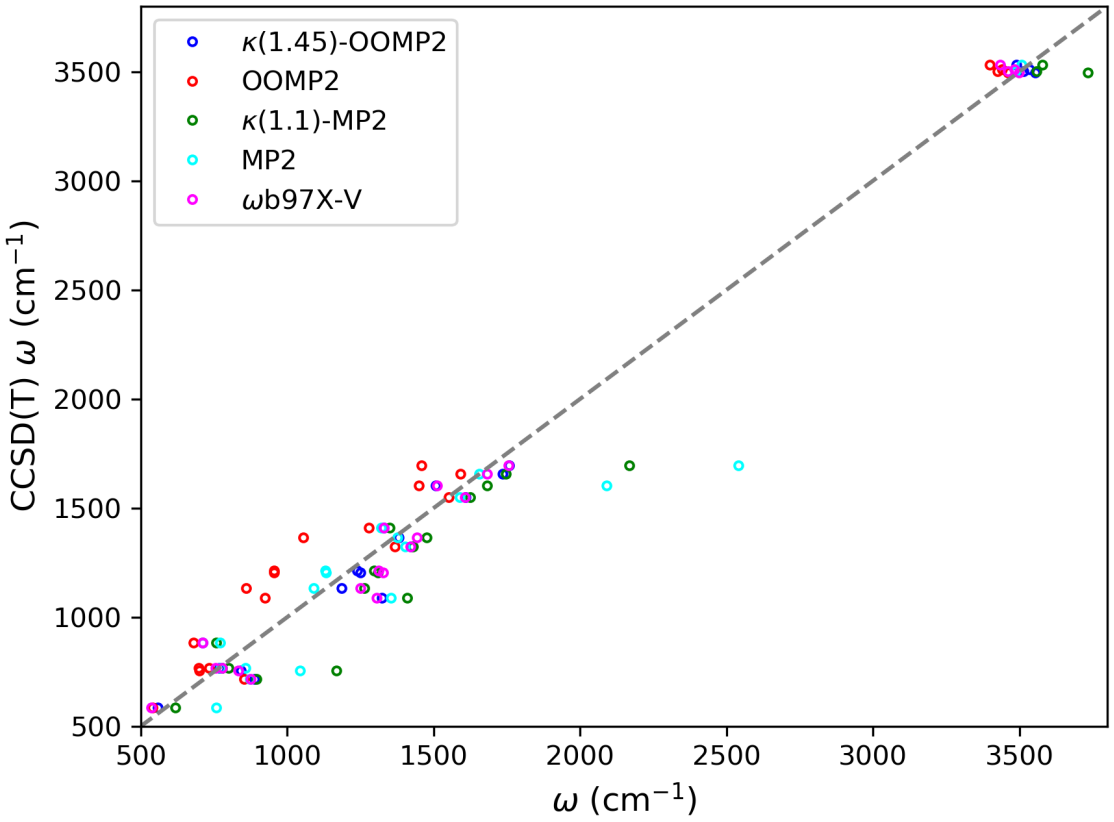
Figure 3.5: Graphical representation of the distribution of errors in vibrational frequencies of the Lochan-OS data set. Values referenced against experimental data.




## $C_{60}$ fullerene singlet-triplet gap

Fullerenes are a class of cage like molecules consisting solely of carbon. It has been a topic of great interest[157]. From the standpoint of symmetry breaking at the mean-field level, an interesting feature of fullerenes is that they often exhibit complex generalized HF (cGHF) solutions[158]. In particular, Jimenez-Hoyos and coworkers interpreted this result as a suggestion that $C_{60}$ fullerene has polyradicaloid character. This is in contrast to experimental observations where $C_{60}$ fullerenes were quite stable and electron paramagnetic resonance silent[159], suggesting closed shell character. Some of us has previously investigated in detail the symmetry breaking and correlation behaviour of $C_{20}$, $C_{36}$, and $C_{60}$[160]. It was concluded that $C_{60}$ is not strongly correlated, the symmetry breaking of cGHF is artificial, and that UMP2 overestimates the singlet-triplet gap by more than a factor of 2, compared to experiment. Interestingly, $\kappa$-UOOMP2 with $\kappa$=1.45 predicted a slightly larger singlet-triplet gap than UHF, both larger than the experimental value. In addition, with $\kappa > 1.05$, all symmetries that were broken at the HF level is restored. On the other hand. $C_{36}$ and the $I_h$ isomer

Figure 3.6: Graphical representation of the distribution of errors in vibrational frequencies of the Bozkaya data set. Values referenced against CCSD(T)/cc-pCVQZ




of $C_{20}$ is strongly correlated.

Since $C_{60}$ is not strongly correlated, for the purposing of assessing the geometric behaviour of a single reference method, we investigate the adiabatic singlet-triplet gap of $C_{60}$ fullerene. We hope to see whether adiabetic treatment will resolve the single-triple gap over estimation.

In Table 3.3 we present the adiabatic $C_{60}$ singlet-triplet gap computed with $\kappa$-UOOMP2 with various $\kappa$ values. The cc-pVDZ basis set and the corresponding auxiliary basis set was used. The singlet-triplet gap predicted by $\kappa = 2$ (38.75 kcal mol$^{-1}$) is remarkably close to the experimental single-triplet gap (36.95 $\pm$ 0.02 kcal mol$^{-1}$), with only 1.8 kcal mol$^{-1}$ difference. This error gets larger as regularization strengthens with increasing $\kappa$. It appears for $C_{60}$ single-triplet gap, weaker regularization than $\kappa = 1.45$ is preferred. We also note that, compared to the vertical gap of 49.23 kcal mol$^{-1}$ found previously for $\kappa$-OOMP2 with $\kappa = 1.45$, the adiabetic gap decreased by around 4 kcal mol$^{-1}$.

Table 3.3: $C_{60}$ Fullerene singlet-triplet gap $\Delta E_{S-T}$ predicted by $\kappa$-UOOMP2 at various $\kappa$ values with the cc-pVDZ basis set. (kcal mol$^{-1}$)

| | $\kappa$-UOOMP2 | | | Experiment |
|---|---|---|---|---|
| | 1.1 | 1.45 | 2.0 | |
| $\Delta E_{S-T}$ | 49.66 | 44.93 | 38.75 | $36.95 \pm 0.02$ |

## 3.5 Conclusions

We derived and implemented $\kappa$ regularized OOMP2 nuclear gradients and investigated its performance over a few systems of different sizes. We found that $\kappa$-OOMP2 was able to resolve most of the artifacts arising from the over-ionic nature of HF orbitals. Compared to the optimal $\kappa$ of 1.45 found for thermochemistry, weaker regularization produced smaller errors for the systems studied in this work.

# Chapter 4

# Mechanism of Solvent Water Oxygen Incorporation into the Oxygenated Products of CO Reduction over Cu

The incorporation of oxygen from solvent water into the multicarbon oxygenates produced during the electrochemical reduction of CO over Cu has recently been reported.[161] This process was observed by conducting $C^{16}O$ reduction in $H_2{}^{18}O$-based electrolyte and measuring the isotopic composition of the resulting liquid-phase products by GC-MS. A significant fraction of the acetate, ethanol, and n-propanol produced were found to contain $^{18}O$, suggesting that the oxygen from solvent water was incorporated into the multicarbon oxygenated products. The authors explained their observations by proposing a new reaction mechanism, which they called Grotthuss chain ethynyl (C-CH) concerted hydrolysis. The proposed mechanism involves the simultaneous transfer of $H^+$ to the C end and $OH^-$ to the CH end of an ethynyl intermediate by water molecules in a Grotthuss chain. The calculated energy barrier for this process was +0.81 eV, which is significantly higher than the barriers for the reduction of ethynyl to either ethylene (+0.60 eV) or ethanol (+0.67 eV) reported previously by the authors. [161, 162] Thus, the proposed mechanism is not energetically favorable despite its consistency with the experimental observations. As shown below, a simpler explanation for these observations that is also consistent with the theoretical models and previously proposed mechanisms can be provided by considering the known chemistry of aldehydes.

While the authors considered the possibility of isotopic scrambling between the solvent water and observed reaction products, they failed to account for the possibility of isotopic scrambling with acetaldehyde, an intermediate in the reduction of CO and $CO_2$ to ethanol over Cu.[163] The hypothesis that acetaldehyde is an intermediate of $CO_2$ reduction is supported by numerous experimental observations. Acetaldehyde reduction over Cu yields ethanol. [164, 165] Furthermore, the concentration of acetaldehyde in the bulk electrolyte rapidly plateaus during CO reduction over Cu, suggesting that it is in equilibrium with the electrode surface.[166] Finally, acetaldehyde is found in relatively high abundance in the immediate vicinity of a Cu cathode during $CO_2$ reduction despite being present only in trace quantities in the bulk electrolyte.[167] These observations support the conclusion that ethanol is produced during CO and $CO_2$ reduction via the reduction of

acetaldehyde. Therefore, it is necessary to consider the possibility that the incorporation of oxygen from solvent water into the multicarbon oxygenated products occurs by isotopic scrambling of a transient acetaldehyde intermediate.

Acetaldehyde is known to undergo hydration in aqueous solutions to form ethane-1,1-diol, which results in isotopic scrambling with solvent water. [168, 169] Based on the thermodynamics for the hydration reaction, $\sim$55% of the acetaldehyde in solution is expected to be hydrated at room temperature. The electrochemical reduction of this acetaldehyde would then yield ethanol containing oxygen derived from solvent water, as observed by Lum *et al*. Interestingly, the authors observed that $\sim$65% of the ethanol produced during $C^{16}O$ reduction contained $^{18}O$ from the solvent water, in close agreement with the fraction of acetaldehyde that is hydrated at equilibrium.

We have carried out density functional theory (DFT) electronic structure calculations that model the interconversion of acetaldehyde with its hydrate to support this hypothesis. The calculations were performed with Q-Chem [170] using the $\omega$B97M-V functional,[171] a highly accurate hybrid meta-GGA for both thermodynamics and kinetics, [172, 173] the def2-TZVPPD basis set,[174] and the SMD aqueous solvation model.[175] For the transition state determination, the freezing string method was used to provide a rough reaction path,[176] followed by a transition state search initiating from the maximum of the reaction path. Thermal corrections from the molecular motions to the free energy were calculated using the standard rigid-rotor harmonic-oscillator approximation. The results, presented in Fig. 4.1, show that under neutral pH the free energy barrier ($\Delta G^{\ddagger}$) for the equilibrium between acetaldehyde ($CH_3CHO$) and ethane-1,1-diol ($CH_3CH(OH)_2$) is reduced as the number of explicit water molecules included in the calculation is increased, illustrating catalysis via a proton shuttle mechanism. The adiabatic free energy difference ($\Delta G_{\text{reac}}$) is consistent with the experimental equilibrium constant of close to unity.

Under the realistic conditions of $CO_2$ reduction, the pH in the vicinity of the Cu cathode increases substantially due to concentration polarization.[177] Under such conditions, a hydroxide-catalyzed pathway for hydration of acetaldehyde opens up, as shown in Fig. 4.2. The acetaldehyde is first attacked by a hydroxide ion through the hydroxide shuttle to yield the very strong conjugate base of ethane-1,1-diol (calculated $pK_a$ of 12.1), ethane-1,1-diolate ($CH_3CH(OH)O^-$). This reaction has a relatively low free energy barrier (10.49 kcal/mol), allowing rapid interconversion between acetaldehyde and ethane-1,1-diolate. Furthermore, this calculated activation barrier is in close agreement with what has been reported experimentally (10.75 kcal/mol).18 Ethane-1,1-diolate is then protonated by water to yield ethane-1,1-diol, regenerating the hydroxide ion. Clearly, acetaldehyde does not have to leave the vicinity of the working electrode for incorporation of oxygen from solvent to occur on short time scales.

The incorporation of oxygen from solvent water into the oxygenated products of CO reduction has already been hypothesized to occur via the isotopic scrambling of an acetaldehyde intermediate.[169] However, this mechanism does not explain the incorporation of oxygen from solvent water into other oxygenated products, such as acetate and n-propanol. Propionaldehyde can be reduced to n-propanol and also undergoes hydration reactions in aqueous solutions with hydroxide catalyzed activation barriers similar to those observed over acetaldehyde. [164, 178] Furthermore, acetaldehyde has recently been hypothesized to be an intermediate in the formation of propionaldehyde over Cu.[167] Thus, the incorporation of oxygen from solvent water into n-propanol can also

be explained in terms of isotopic scrambling of aldehyde intermediates (either acetaldehyde or propionaldehyde) due to reversible hydration.

It is more difficult to explain the incorporation of oxygen from solvent water into acetate anions, because little is known about the mechanism of acetate formation over Cu. It has been hypothesized that acetate forms via a "Cannizzaro-type" disproportionation of acetaldehyde within the relatively alkaline conditions found near the cathode surface and that this process is initiated by the hydration of acetaldehyde. [179, 180] If this reasoning is correct, then all the acetate produced should contain exactly one oxygen atom derived from the solvent water, in complete agreement with the experimental observations. [161, 169] In conclusion, explanation of the observation that the multicarbon oxygenates produced during CO reduction contain oxygen derived from solvent water does not require a novel reaction mechanism. These observations can be explained via the isotopic scrambling of carbonyl-containing intermediate reaction products that are reduced further to oxygenated products. Consequently, the claims that "all previous mechanisms for the formation of oxygenates require reexamination"[161] and that "$H_2O$ is the dominant source of O for the formation of these [oxygenated] products"[161] must be regarded with caution.

Figure 4.1: Calculated barrier height, $\Delta G^{\ddagger}$ (left scale, black) and adiabatic free energy difference, $\Delta G_{\text{reac}}$ (right scale, red) for the acetaldehyde hydration reaction to ethane-1,1-diol. Geometries shown are the calculated transition structures for each number of water molecules, n($H_2O$). The trend in $\Delta G^{\ddagger}$ shows the importance of a proton shuttle mechanism in reducing the barrier.

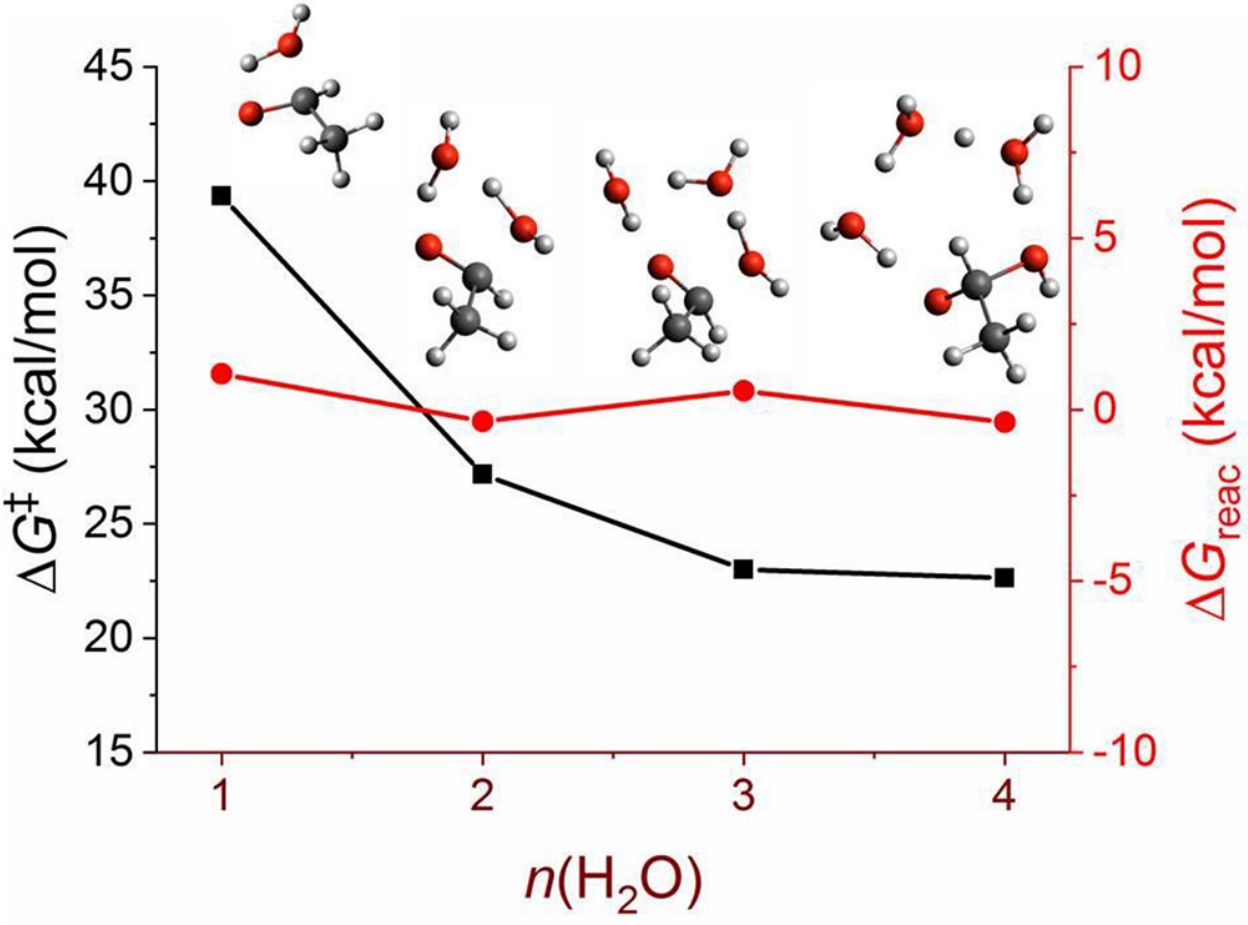
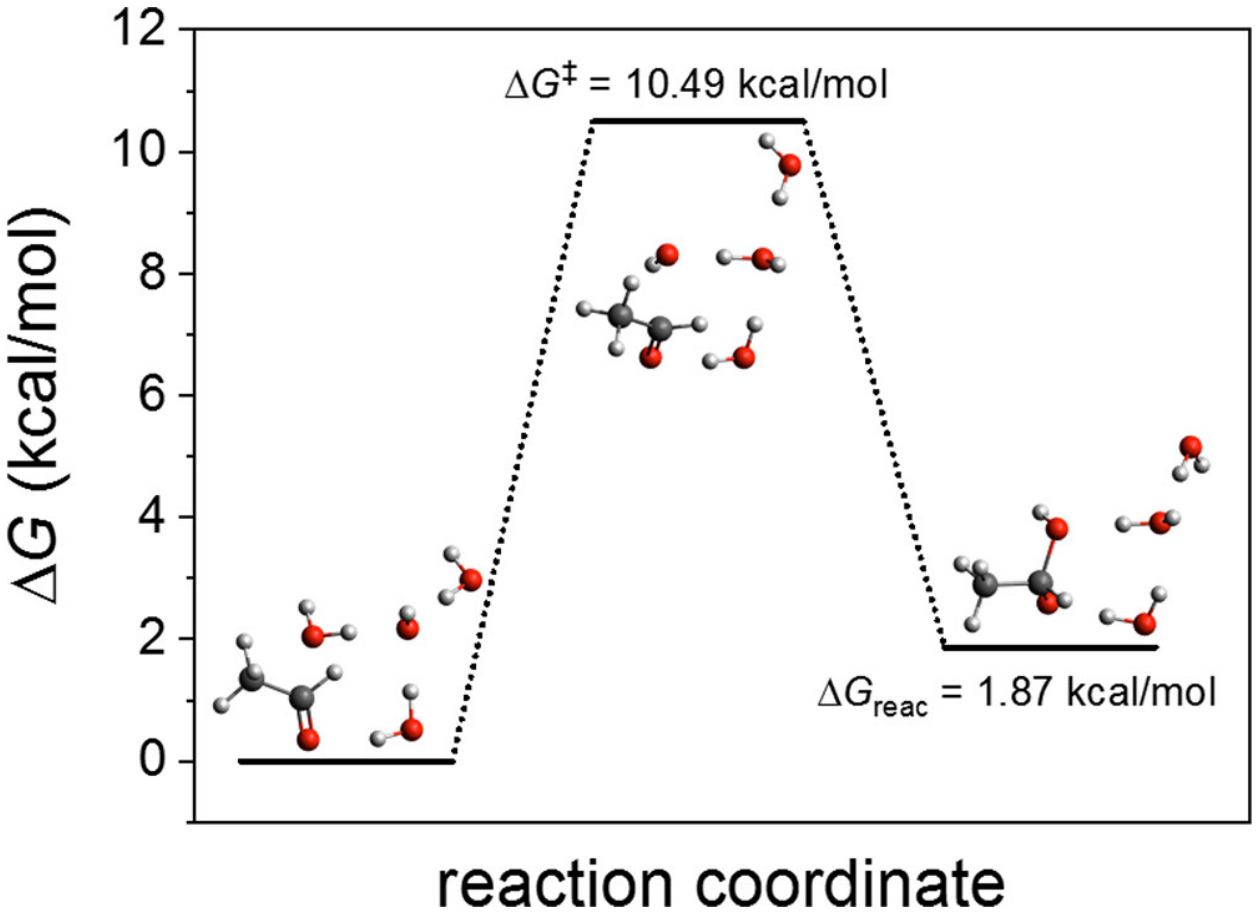


Figure 4.2: Calculated free energy barrier height, $\Delta G^{\ddagger}$, and adiabatic free energy difference, $\Delta G_{\text{reac}}$, for the hydroxide-catalyzed acetaldehyde hydration reaction under a high pH condition in the vicinity of the Cu cathode. Geometries shown are the calculated structures for the complexes of the reactant, transition state, and product. Three explicit water molecules are included in these geometries.



# Bibliography


(1) Schrödinger, E. *Ann. Phys.* **1926**, *79*, 361.

(2) Roothaan, C. C. J. New Developments in Molecular Orbital Theory. *Rev. Mod. Phys.* **1951**, *23*, 69–89.

(3) Purvis, G. D.; Bartlett, R. J. *J. Chem. Phys.* **1982**, *76*, 1910.

(4) Scuseria, G. E.; Janssen, C. L.; Schaefer Henry F., I. An efficient reformulation of the closed-shell coupled cluster single and double excitation (CCSD) equations. **1988**, *89*, 7382–7387.

(5) Noga, J.; Bartlett, R. J. The full CCSDT model for molecular electronic structure. *Journal of Chemical Physics* **1987**, *86*, 7041–7050.

(6) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. A fifth-order perturbation comparison of electron correlation theories. *Chem. Phys. Lett.* **1989**, *157*, 479.

(7) Møller, C.; Plesset, M. S. Note on an approximation treatment for many-electron systems. *Phys. Rev.* **1934**, *46*, 618.

(8) Kvasnička, V.; Laurinc, V.; Biskupič, S. Wigner's (2n + 1) rule in MBPT. *Molecular Physics* **1980**, *39*, 143–161.

(9) Lee, J.; Head-Gordon, M. Regularized orbital-optimized second-order Møller–Plesset perturbation theory: A reliable fifth-order-scaling electron correlation model with orbital energy dependent regularizers. *J. Chem. Theory Comput.* **2018**, *14*, 5203–5219.

(10) Stück, D.; Head-Gordon, M. Regularized orbital-optimized second-order perturbation theory. *J. Chem. Phys.* **2013**, *139*, 244109.

(11) Razban, R. M.; Stück, D.; Head-Gordon, M. Addressing first derivative discontinuities in orbital-optimised opposite-spin scaled second-order perturbation theory with regularisation. *Mol. Phys.* **2017**, *115*, 21029.

(12) Nobes, R. H.; Pople, J. A.; Radom, L.; Handy, N. C.; Knowles, P. J. Slow convergence of the møller-plesset perturbation series: the dissociation energy of hydrogen cyanide and the electron affinity of the cyano radical. *Chem. Phys. Lett.* **1987**, *138*, 481–485.

(13) Hylleraas, E. A. Über den Grundterm der Zweielektronenprobleme von H, He, Li+, Be++ usw. *Zeitschrift für Physik* **1930**, *65*, 209–225.

(14) Neese, F.; Schwabe, T.; Kossmann, S.; Schirmer, B.; Grimme, S. Assessment of orbital-optimized, spin-component scaled second-order many-body perturbation theory for thermochemistry and kinetics. *J. Chem. Theory Comput.* **2009**, *5*, 3060–3073.

(15) Hohenberg, P.; Kohn, W. Inhomogeneous electron gas. *Phys. Rev. B* **1964**, *136*, B864–B871.

(16) Thomas, L. H. The calculation of atomic fields. *Mathematical Proceedings of the Cambridge Philosophical Society* **1927**, *23*, 542–548.

(17) Fermi, E. A Statistical Method for the Determination of Some Atomic Properties and the Application of this Method to the Theory of the Periodic System of Elements. **1975**.

(18) Hung, L.; Carter, E. A. Accurate simulations of metals at the mesoscale: Explicit treatment of 1 million atoms with quantum mechanics. *Chemical Physics Letters* **2009**, *475*, 163–170.

(19) Shin, I.; Ramasubramaniam, A.; Huang, C.; Hung, L.; Carter, E. Orbital-free density functional theory simulations of dislocations in aluminum. *Philosophical Magazine* **2009**, *89*, Funding Information: We are grateful to the Office of Naval Research for funding this work and Princeton University for supercomputing time. L. Hung thanks the National Defense Science and Engineering Graduate Fellowship program. We also wish to thank Drs. Yue Qi and Louis Hector for suggesting this study. Copyright: Copyright 2012 Elsevier B.V., All rights reserved., 3195–3213.

(20) Goerigk, L.; Hansen, A.; Bauer, C.; Ehrlich, S.; Najibi, A.; Grimme, S. A look at the density functional theory zoo with the advanced GMTKN55 database for general main group thermochemistry, kinetics and noncovalent interactions. *Phys. Chem. Chem. Phys.* **2017**, *19*, 32184–32215.

(21) Perdew, J. P.; Ruzsinszky, A.; Tao, J.; Staroverov, V. N.; Scuseria, G. E.; Csonka, G. I. *J. Chem. Phys.* **2005**, *123*, 062201.

(22) Grimme, S. Accurate description of van der Waals complexes by density functional theory including empirical corrections. *Journal of Computational Chemistry* **2004**, *25*, 1463–1473.

(23) Grimme, S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.* **2006**, *27*, 1787–1789.

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

(25) Vydrov, O. A.; Van Voorhis, T. Nonlocal van der Waals density functional: The simpler the better. *J. Chem. Phys.* **2010**, *133*, 244103.

(26) Mardirossian, N.; Head-Gordon, M. Mapping the genome of meta-generalized gradient approximation density functionals: The search for B97M-V. *J. Chem. Phys.* **2015**, *142*, 074111.

(27) Mardirossian, N.; Head-Gordon, M. $\omega$B97X-V: A 10-parameter, range-separated hybrid, generalized gradient approximation density functional with nonlocal correlation, designed by a survival-of-the-fittest strategy. *Phys. Chem. Chem. Phys.* **2014**, *16*, 9904–9924.

(28) Mardirossian, N.; Head-Gordon, M. $\omega$B97M-V: A combinatorially optimized, range-separated hybrid, meta-GGA density functional with VV10 nonlocal correlation. *J. Chem. Phys.* **2016**, *144*, 214110.

(29) Mardirossian, N.; Head-Gordon, M. Thirty years of density functional theory in computational chemistry: An overview and extensive assessment of 200 density functionals. *Mol. Phys.* **2017**, *115*, 2315–2372.

(30) Grimme, S. Semiempirical hybrid density functional with perturbative second-order correction. *J. Chem. Phys.* **2006**, *124*, 034108.

(31) Mardirossian, N.; Head-Gordon, M. Survival of the most transferable at the top of Jacob's ladder: Defining and testing the $\omega$B97M(2) double hybrid density functional. *J. Chem. Phys.* **2018**, *148*, 241736.

(32) Gil, R. R. Constitutional, configurational, and conformational analysis of small organic molecules on the basis of NMR residual dipolar couplings. *Angew. Chem. Int. Ed.* **2011**, *50*, 7222–7224.

(33) Becette, O. B.; Zong, G.; Chen, B.; Taiwo, K. M.; Case, D. A.; Dayie, T. K. Solution NMR readily reveals distinct structural folds and interactions in doubly 13C-and 19F-labeled RNAs. *Sci. Adv.* **2020**, *6*, eabc6572.

(34) Krivdin, L. B. Computational 1H NMR: Part 1. Theoretical background. *Magn. Reson. Chem.* **2019**, *57*, 897–914.

(35) Krivdin, L. B. Computational protocols for calculating 13C NMR chemical shifts. *Prog. Nucl. Magn. Reson. Spectrosc.* **2019**, *112*, 103–156.

(36) Gauss, J. Effects of electron correlation in the calculation of nuclear magnetic resonance chemical shifts. *J. Chem. Phys.* **1993**, *99*, 3629–3643.

(37) Helgaker, T.; Jaszuński, M.; Ruud, K. Ab Initio Methods for the Calculation of NMR Shielding and Indirect Spin-Spin Coupling Constants. *Chem. Rev.* **1999**, *99*, 293–352.

(38) Casabianca, L. B.; de Dios, A. C. Ab initio calculations of NMR chemical shifts. *The Journal of Chemical Physics* **2008**, *128*, 052201.

(39) Grimblat, N.; Sarotti, A. M. Computational chemistry to the rescue: Modern toolboxes for the assignment of complex molecules by GIAO NMR calculations. *Chemistry–A European Journal* **2016**, *22*, 12246–12261.

(40) Wolinski, K.; Hinton, J. F.; Pulay, P. *J. Am. Chem. Soc.* **1990**, *112*, 8251.

(41) Auer, A. A.; Gauss, J.; Stanton, J. F. Quantitative prediction of gas-phase 13C nuclear magnetic shielding constants. *J. Chem. Phys.* **2003**, *128*, 10407–10417.

(42) Teale, A. M.; Lutnæs, O. B.; Helgaker, T.; Tozer, D. J.; Gauss, J. Benchmarking density-functional theory calculations of NMR shielding constants and spin–rotation constants using accurate coupled-cluster calculations. *J. Chem. Phys* **2013**, *138*, 024111.

(43) Grayce, C. J.; Harris, R. A. Magnetic-field density-functional theory. *Phys. Rev. A* **1994**, *50*, 3089–3095.

(44) Tellgren, E. I.; Kvaal, S.; Sagvolden, E.; Ekström, U.; Teale, A. M.; Helgaker, T. Choice of basic variables in current-density-functional theory. *Phys. Rev. A* **2012**, *86*, 062506.

(45) Furness, J. W.; Verbeke, J.; Tellgren, E. I.; Stopkowicz, S.; Ekström, U.; Helgaker, T.; Teale, A. M. Current density functional theory using meta-generalized gradient exchange-correlation functionals. *Journal of chemical theory and computation* **2015**, *11*, 4169–4181.

(46) Tao, J.; Mo, Y. *Phys. Rev. Lett.* **2016**, *117*, 073001.

(47) Reimann, S.; Borgoo, A.; Austad, J.; Tellgren, E. I.; Teale, A. M.; Helgaker, T.; Stopkowicz, S. Kohn–Sham energy decomposition for molecules in a magnetic field. *Molecular Physics* **2019**, *117*, 97–109.

(48) Flaig, D.; Maurer, M.; Hanni, M.; Braunger, K.; Kick, L.; Thubauville, M.; Ochsenfeld, C. Benchmarking hydrogen and carbon NMR chemical shifts at HF, DFT, and MP2 levels. *J. Chem. Theory Comput.* **2014**, *10*, 572–578.

(49) Irons, T. J. P.; Zemen, J.; Teale, A. M. Efficient Calculation of Molecular Integrals over London Atomic Orbitals. *J. Chem. Theory Comput.* **2017**, *13*, 3636–3649.

(50) Stoychev, G. L.; Auer, A. A.; Izsak, R.; Neese, F. Self-consistent field calculation of nuclear magnetic resonance chemical shielding constants using gauge-including atomic orbitals and approximate two-electron integrals. *J. Chem. Theory Comput.* **2018**, *14*, 619–637.

(51) Schattenberg, C. J.; Kaupp, M. Extended Benchmark Set of Main-Group Nuclear Shielding Constants and NMR Chemical Shifts and Its Use to Evaluate Modern DFT Methods. *J. Chem. Theory Comput.* **2021**, *17*, 7602–7621.

(52) Liang, J.; Wang, Z.; Li, J.; Wong, J.; Liu, X.; Ganoe, B.; Head-Gordon, T.; Head-Gordon, M. Efficient Calculation of NMR Shielding Constants Using Composite Method Approximations and Locally Dense Basis Sets. *J. Chem. Theory Comput.* **2023**.

(53) Ochsenfeld, C.; Kussmann, J.; Koziol, F. Ab initio NMR spectra for molecular systems with a thousand and more atoms: A linear-scaling method. *Angew. Chem. Int. Ed.* **2004**, *43*, 4485–4489.

(54) Kussmann, J.; Ochsenfeld, C. *J. Chem. Phys.* **2007**, *127*, 054103.

(55) Kussmann, J.; Ochsenfeld, C. *J. Chem. Phys.* **2007**, *127*, 204103.

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

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

(58) Gauss, J. Analytic second derivatives for the full coupled-cluster singles, doubles, and triples model: Nuclear magnetic shielding constants for BH, HF, CO, N 2, N 2 O, and O 3. *J. Chem. Phys.* **2002**, *116*, 4773–4776.

(59) Reid, D. M.; Collins, M. A. Approximating CCSD (T) nuclear magnetic shielding calculations using composite methods. *J. Chem. Theory Comput.* **2015**, *11*, 5177–5181.

(60) Prochnow, E.; Auer, A. A. Quantitative prediction of gas-phase N15 and P31 nuclear magnetic shielding constants. *The Journal of Chemical Physics* **2010**, *132*, 064109.

(61) Harding, M. E.; Lenhart, M.; Auer, A. A.; Gauss, J. Quantitative prediction of gas-phase F19 nuclear magnetic shielding constants. *The Journal of Chemical Physics* **2008**, *128*, 244111.

(62) Gauss, J.; Stanton, J. F. In *Advances in Chemical Physics*; John Wiley Sons, Ltd: 2002; Chapter 6, pp 355–422.

(63) Handy, N. C.; Schaefer III, H. F. On the evaluation of analytic energy derivatives for correlated wave functions. *J. Chem. Phys.* **1984**, *81*, 5031–5033.

(64) Gauss, J.; Werner, H.-J. NMR chemical shift calculations within local correlation methods: the GIAO-LMP2 approach. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2083–2090.

(65) Loibl, S.; Schütz, M. NMR shielding tensors for density fitted local second-order Moller-Plesset perturbation theory using gauge including atomic orbitals. *J. Chem. Phys.* **2012**, *137*, 084107.

(66) Stoychev, G. L.; Auer, A. A.; Gauss, J.; Neese, F. DLPNO-MP2 second derivatives for the computation of polarizabilities and NMR shieldings. *The Journal of Chemical Physics* **2021**, *154*, 164110.

(67) Almlöf, J. Elimination of energy denominators in Møller—Plesset perturbation theory by a Laplace transform approach. *Chem. Phys. Lett.* **1991**, *181*, 319–320.

(68) Häser, M.; Ahlrichs, R.; Baron, H. P.; Weiss, P.; Horn, H. *Theor. Chem. Acc.* **1992**, *83*, 455.

(69) Ayala, P. Y.; Scuseria, G. E. Linear scaling second-order Moller–Plesset theory in the atomic orbital basis for large molecular systems. *The Journal of chemical physics* **1999**, *110*, 3660–3671.

(70) Maurer, S. A.; Lambrecht, D. S.; Flaig, D.; Ochsenfeld, C. Distance-dependent Schwarz-based integral estimates for two-electron integrals: Reliable tightness vs. rigorous upper bounds. *The Journal of Chemical Physics* **2012**, *136*, 144107.

(71) Maurer, S. A.; Lambrecht, D. S.; Kussmann, J.; Ochsenfeld, C. Efficient distance-including integral screening in linear-scaling Møller-Plesset perturbation theory. *The Journal of Chemical Physics* **2013**, *138*, 014101.

(72) Glasbrenner, M.; Vogler, S.; Ochsenfeld, C. Efficient low-scaling computation of NMR shieldings at the second-order Møller–Plesset perturbation theory level with Cholesky-decomposed densities and an attenuated Coulomb metric. *The Journal of Chemical Physics* **2021**, *155*, 224107.

(73) Blaschke, S.; Stopkowicz, S. Cholesky decomposition of complex two-electron integrals over GIAOs: Efficient MP2 computations for large molecules in strong magnetic fields. *The Journal of Chemical Physics* **2022**, *156*, 044115.

(74) Grimme, S. Improved second-order Møller–Plesset perturbation theory by separate scaling of parallel-and antiparallel-spin pair correlation energies. *J. Chem. Phys.* **2003**, *118*, 9095–9102.

(75) Jung, Y.; Heine, T.; Schleyer, P.; Head-Gordon, M. Aromaticity of four-membered-ring 6 pi-electron systems: N2S2 and Li2C4H4. *J. Am. Chem. Soc.* **2004**, *126*, 3132–3138.

(76) Stoychev, G. L.; Auer, A. A.; Neese, F. Efficient and accurate prediction of nuclear magnetic resonance shielding tensors with double-hybrid density functional theory. *J. Chem. Theory Comput.* **2018**, *14*, 4756–4771.

(77) Yan, W.; Xu, X. Accurate Prediction of Nuclear Magnetic Resonance Parameters via the XYG3 Type of Doubly Hybrid Density Functionals. *Journal of Chemical Theory and Computation* **2022**, *18*, PMID: 35467852, 2931–2946.

(78) Stück, D.; Head-Gordon, M. Regularized orbital-optimized second-order perturbation theory. *J. Chem. Phys.* **2013**, *139*, 244109.

(79) Lee, J.; Head-Gordon, M. Regularized Orbital-Optimized Second-Order Møller-Plesset Perturbation Theory : A Reliable Fifth-Order Scaling Electron Correlation Model with Orbital Energy Dependent Regularizers. *J. Chem. Theory Comput.* **2018**, *14*, 5203–5219.

(80) Shee, J.; Loipersberger, M.; Rettig, A.; Lee, J.; Head-Gordon, M. 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. *J. Phys. Chem. Lett.* **2021**, *12*, 12084–12097.

(81) Nguyen, B. D.; Chen, G. P.; Agee, M. M.; Burow, A. M.; Tang, M. P.; Furche, F. Divergence of many-body perturbation theory for noncovalent interactions of large molecules. *J. Chem. Theory Comput.* **2020**, *16*, 2258–2273.

(82) Pitoňák, M.; Neogrády, P.; Černý, J.; Grimme, S.; Hobza, P. Scaled MP3 Non-Covalent Interaction Energies Agree Closely with Accurate CCSD(T) Benchmark Data. *ChemPhysChem* **2009**, *10*, 282–289.

(83) Riley, K. E.; Řezáč, J.; Hobza, P. The performance of MP2.5 and MP2.X methods for nonequilibrium geometries of molecular complexes. *Phys. Chem. Chem. Phys.* **2012**, *14*, 13187–13193.

(84) Sedlak, R.; Riley, K. E.; Řezáč, J.; Pitoňák, M.; Hobza, P. MP2.5 and MP2.X: Approaching CCSD(T) Quality Description of Noncovalent Interaction at the Cost of a Single CCSD Iteration. *ChemPhysChem* **2013**, *14*, 698–707.

(85) Bertels, L. W.; Lee, J.; Head-Gordon, M. Third-Order Møller–Plesset Perturbation Theory Made Useful? Choice of Orbitals and Scaling Greatly Improves Accuracy for Thermochemistry, Kinetics, and Intermolecular Interactions. *J. Phys. Chem. Lett.* **2019**, *10*, 4170–4176.

(86) Rettig, A.; Hait, D.; Bertels, L. W.; Head-Gordon, M. Third-Order Møller–Plesset Theory Made More Useful? The Role of Density Functional Theory Orbitals. *J. Chem. Theory Comput.* **2020**, *16*, 7473–7489.

(87) Loipersberger, M.; Cabral, D. G. A.; Chu, D. B. K.; Head-Gordon, M. Mechanistic insights into Co and Fe quaterpyridine-based $CO_2$ reduction catalysts: Metal–ligand orbital interaction as the key driving force for distinct pathways. *J. Am. Chem. Soc.* **2021**, *143*, 744–763.

(88) Helgaker, T.; Coriani, S.; Jørgensen, P.; Kristensen, K.; Olsen, J.; Ruud, K. Recent Advances in Wave Function-Based Methods of Molecular-Property Calculations. *Chem. Rev.* **2012**, *112*, 543–631.

(89) Stauch, T.; Ganoe, B.; Wong, J.; Lee, J.; Rettig, A.; Liang, J.; Li, J.; Epifanovsky, E.; Head-Gordon, T.; Head-Gordon, M. Molecular magnetisabilities computed via finite fields: assessing alternatives to MP2 and revisiting magnetic exaltations in aromatic and antiaromatic species. *Molecular Physics* **2021**, *119*, PMID: 35264815, e1990426.

(90) Tellgren, E. I.; Teale, A. M.; Furness, J. W.; Lange, K. K.; Ekström, U.; Helgaker, T. Non-perturbative calculation of molecular magnetic properties within current-density functional theory. *J. Chem. Phys.* **2014**, *140*, 034101.

(91) Stopkowicz, S.; Gauss, J.; Lange, K. K.; Tellgren, E. I.; Helgaker, T. Coupled-cluster theory for atoms and molecules in strong magnetic fields. *J. Chem. Phys.* **2015**, *143*, 074110.

(92) Sen, S.; Tellgren, E. I. Benchmarking density functional approximations for diamagnetic and paramagnetic molecules in nonuniform magnetic fields. *Journal of chemical theory and computation* **2021**, *17*, 1480–1496.

(93) Lutnæs, O. B.; Teale, A. M.; Helgaker, T.; Tozer, D. J.; Ruud, K.; Gauss, J. Benchmarking density-functional-theory calculations of rotational g tensors and magnetizabilities using accurate coupled-cluster calculations. *J. Chem. Phys.* **2009**, *131*, 144104.

(94) Vaara, J. Theory and computation of nuclear magnetic resonance parameters. *Phys. Chem. Chem. Phys.* **2007**, *9*, 5399–5418.

(95) London, F. Théorie Quantique des Courants Interatomiques Dans les Combinaisons Aromatiques. *J. Phys. Radium* **1937**, *8*, 397–409.

(96) Fornberg, B. Generation of Finite Difference Formulas on Arbitrarily Spaced Grids. *Math. Comput.* **1988**, *51*, 699–706.

(97) Helgaker, T.; Jørgensen, P. An electronic Hamiltonian for origin independent calculations of magnetic properties. *J. Chem. Phys.* **1991**, *95*, 2595–2601.

(98) Lee, J.; Head-Gordon, M. Two single-reference approaches to singlet biradicaloid problems: Complex, restricted orbitals and approximate spin-projection combined with regularized orbital-optimized Møller-Plesset perturbation theory. *J. Chem. Phys.* **2019**, *150*, 244106.

(99) Lee, J.; Head-Gordon, M. Distinguishing artificial and essential symmetry breaking in a single determinant: Approach and application to the C 60, C 36, and C 20 fullerenes. *Phys. Chem. Chem. Phys.* **2019**, *21*, 4763–4778.

(100) Razban, R. M.; Stück, D.; Head-Gordon, M. Addressing first derivative discontinuities in orbital-optimised opposite-spin scaled second-order perturbation theory with regularisation. *Mol. Phys.* **2017**, *115*, 2102–2109.

(101) Hohenstein, E. G.; Sherrill, C. D. Wavefunction methods for noncovalent interactions. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 304–326.

(102) Handy, N.; Knowles, P.; Somasundram, K. On the convergence of the Møller-Plesset perturbation series. *Theor. Chem. Acc.* **1985**, *68*, 87–100.

(103) Alberts, I. L.; Handy, N. C. Mo/ller–Plesset third order calculations with large basis sets. *The Journal of chemical physics* **1988**, *89*, 2107–2115.

(104) Epifanovsky, E.; Gilbert, A. T.; Feng, X.; Lee, J.; Mao, Y.; Mardirossian, N.; Pokhilko, P.; White, A. F.; Coons, M. P.; Dempwolff, A. L., et al. Software for the frontiers of quantum chemistry: An overview of developments in the Q-Chem 5 package. *J. Chem. Phys.* **2021**, *155*, 084801.

(105) Lee, J.; Bertels, L. W.; Small, D. W.; Head-Gordon, M. Kohn-sham density functional theory with complex, spin-restricted orbitals: Accessing a new class of densities without the symmetry dilemma. *Phys. Rev. Lett.* **2019**, *123*, 113001.

(106) Valatin, J. G. Generalized Hartree-Fock Method. *Phys. Rev.* **1961**, *122*, 1012–1020.

(107) Stanton, J. F.; Gauss, J.; Cheng, L.; Harding, M. E.; Matthews, D. A.; Szalay, P. G. CFOUR, a quantum chemical program package. For the current version, see http://www.cfour.de.

(108) Auer, A. A.; Gauss, J. Orbital instabilities and spin-symmetry breaking in coupled-cluster calculations of indirect nuclear spin-spin coupling constants. *Chem. Phys.* **2009**, *356*, 7–13.

(109) Lehtola, S.; Dimitrova, M.; Fliegl, H.; Sundholm, D. Benchmarking Magnetizabilities with Recent Density Functionals. *J. Chem. Theory Comput.* **2021**, DOI: 10.1021/acs.jctc.0c01190.

(110) De Oliveira, M. T.; Alves, J. M.; Braga, A. A.; Wilson, D. J.; Barboza, C. A. Do double-hybrid exchange–correlation functionals provide accurate chemical shifts? A benchmark assessment for proton NMR. *J. Chem. Theory Comput.* **2021**, *17*, 6876–6885.

(111) Rettig, A.; Shee, J.; Lee, J.; Head-Gordon, M. Revisiting the Orbital Energy-Dependent Regularization of Orbital-Optimized Second-Order Møller–Plesset Theory. *Journal of Chemical Theory and Computation* **2022**, *18*, 5382–5392.

(112) Schlegel, H. B. Exploring potential energy surfaces for chemical reactions: an overview of some practical methods. *J. Comput. Chem.* **2003**, *24*, 1514–1527.

(113) Riley, K. E.; Op't Holt, B. T.; Merz, K. M. Critical assessment of the performance of density functional methods for several atomic and molecular properties. *J. Chem. Theory Comput.* **2007**, *3*, 407–433.

(114) Neidinger, R. D. Introduction to automatic differentiation and MATLAB object-oriented programming. *SIAM Rev.* **2010**, *52*, 545–563.

(115) Margossian, C. C. A review of automatic differentiation and its efficient implementation. *WiREs: Rev. Data Min. Knowl. Discov.* **2019**, *9*, e1305.

(116) Tamayo-Mendoza, T.; Kreisbeck, C.; Lindh, R.; Aspuru-Guzik, A. Automatic differentiation in quantum chemistry with applications to fully variational Hartree–Fock. *ACS Cent. Sci.* **2018**, *4*, 559–566.

(117) Abbott, A. S.; Abbott, B. Z.; Turney, J. M.; Schaefer III, H. F. Arbitrary-order derivatives of quantum chemical methods via automatic differentiation. *The Journal of Physical Chemistry Letters* **2021**, *12*, 3232–3239.

(118) Cremer, D. Moller-Plesset perturbation theory: from small molecule methods to methods for thousands of atoms. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2011**, *1*, 509–530.

(119) Head-Gordon, M.; Pople, J. A.; Frisch, M. J. MP2 energy evaluation by direct methods. *Chem. Phys. Lett.* **1988**, *153*, 503.

(120) Sæbø, S.; Almlöf, J. Avoiding the integral storage bottleneck in LCAO calculations of electron correlation. *Chem. Phys. Lett.* **1989**, *154*, 83–89.

(121) Weigend, F.; Häser, M.; Patzelt, H.; Ahlrichs, R. RI-MP2: optimized auxiliary basis sets and demonstration of efficiency. *Chem. Phys. Lett.* **1998**, *294*, 143.

(122) Weigend, F. Efficient use of the correlation consistent basis sets in resolution of the identity MP2 calculations. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4285.

(123) Werner, H.-J.; Knizia, G.; Krause, C.; Schwilk, M.; Dornbach, M. Scalable Electron Correlation Methods I.: PNO-LMP2 with Linear Scaling in the Molecular Size and Near-Inverse-Linear Scaling in the Number of Processors. *J. Chem. Theory Comput.* **2015**, *11*, 484–507.

(124) Pinski, P.; Riplinger, C.; Valeev, E. F.; Neese, F. Sparse maps—A systematic infrastructure for reduced-scaling electronic structure methods. I. An efficient and simple linear scaling local MP2 method that uses an intermediate basis of pair natural orbitals. *J. Chem. Phys.* **2015**, *143*, 034108.

(125) Helgaker, T.; Klopper, W.; Tew, D. P. Quantitative quantum chemistry. *Mol. Phys.* **2008**, *106*, 2107–2143.

(126) Fanourgakis, G. S.; Apra, E.; Xantheas, S. S. 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. *J. Chem. Phys.* **2004**, *121*, 2655–2663.

(127) Yoo, S.; Apra, E.; Zeng, X. C.; Xantheas, S. S. High-level ab initio electronic structure calculations of water clusters (H2O) 16 and (H2O) 17: A new global minimum for (H2O) 16. *J. Phys. Chem. Lett.* **2010**, *1*, 3122–3127.

(128) Rakshit, A.; Bandyopadhyay, P.; Heindel, J. P.; Xantheas, S. S. Atlas of putative minima and low-lying energy networks of water clusters n = 3–25. *The Journal of Chemical Physics* **2019**, *151*, 214307.

(129) Lao, K. U.; Schäffer, R.; Jansen, G.; Herbert, J. M. Accurate description of intermolecular interactions involving ions using symmetry-adapted perturbation theory. *J. Chem. Theory Comput.* **2015**, *11*, 2473–2486.

(130) Lao, K. U.; Herbert, J. M. A simple correction for nonadditive dispersion within extended symmetry-adapted perturbation theory (XSAPT). *J. Chem. Theory Comput.* **2018**, *14*, 5128–5142.

(131) Zahn, S.; MacFarlane, D. R.; Izgorodina, E. I. Assessment of Kohn–Sham density functional theory and Møller–Plesset perturbation theory for ionic liquids. *Phys. Chem. Chem. Phys.* **2013**, *15*, 13664–13675.

(132) Izgorodina, E. I.; Seeger, Z. L.; Scarborough, D. L.; Tan, S. Y. Quantum chemical methods for the prediction of energetic, physical, and spectroscopic properties of ionic liquids. *Chem. Rev.* **2017**, *117*, 6696–6754.

(133) Byrd, E. F.; Sherrill, C. D.; Head-Gordon, M. The theoretical prediction of molecular radical species: a systematic study of equilibrium geometries and harmonic vibrational frequencies. *J. Phys. Chem. A* **2001**, *105*, 9736–9747.

(134) Lochan, R. C.; Head-Gordon, M. Orbital-optimized opposite-spin scaled second-order correlation: An economical method to improve the description of open-shell molecules. *J. Chem. Phys.* **2007**, *126*, 164101.

(135) Kossmann, S.; Neese, F. Correlated ab initio spin densities for larger molecules: orbital-optimized spin-component-scaled MP2 method. *J. Phys. Chem. A* **2010**, *114*, 11768–11781.

(136) Bozkaya, U.; Turney, J. M.; Yamaguchi, Y.; Schaefer III, H. F.; Sherrill, C. D. Quadratically convergent algorithm for orbital optimization in the orbital-optimized coupled-cluster doubles method and in orbital-optimized second-order Møller-Plesset perturbation theory. *J. Chem. Phys.* **2011**, *135*, 104103.

(137) Bozkaya, U.; Sherrill, C. D. Analytic energy gradients for the orbital-optimized second-order Møller–Plesset perturbation theory. *The Journal of Chemical Physics* **2013**, *138*, 184103.

(138) Bozkaya, U. Orbital-optimized second-order perturbation theory with density-fitting and Cholesky decomposition approximations: An efficient implementation. *J. Chem. Theory Comput.* **2014**, *10*, 2371–2378.

(139) Weigend, F.; Häser, M. RI-MP2: first derivatives and global consistency. *Theor. Chem. Acc.* **1997**, *97*, 331–340.

(140) Head-Gordon, M. An improved semidirect MP2 gradient method. *Mol. Phys.* **1999**, *96*, 673.

(141) Aikens, C. M.; Webb, S. P.; Bell, R. L.; Fletcher, G. D.; Schmidt, M. W.; Gordon, M. S. A derivation of the frozen-orbital unrestricted open-shell and restricted closed-shell second-order perturbation theory analytic gradient expressions. *Theor. Chem. Acc.* **2003**, *110*, 233–253.

(142) DiStasio, Jr., R. A.; Steele, R. P.; Rhee, Y. M.; Shao, Y.; Head-Gordon, M. An improved algorithm for analytical gradient evaluation in resolution-of-the-identity second-order Møller-Plesset perturbation theory: Application to alanine tetrapeptide conformational analysis. *J. Comput. Chem.* **2007**, *28*, 839.

(143) Kurlancheek, W.; Lochan, R.; Lawler, K.; Head-Gordon, M. Exploring the competition between localization and delocalization of the neutral soliton defect in polyenyl chains with the orbital optimized second order opposite spin method. *The Journal of Chemical Physics* **2012**, *136*, 054113.

(144) Bozkaya, U.; Sherrill, C. D. Analytic energy gradients for the orbital-optimized second-order Møller–Plesset perturbation theory. *J. Chem. Phys.* **2013**, *138*, 184103.

(145) Lee, J.; Head-Gordon, M. Distinguishing artificial and essential symmetry breaking in a single determinant: Approach and application to the $C_{60}$, $C_{36}$, and $C_{20}$ fullerenes. *Phys. Chem. Chem. Phys.* **2019**, *21*, 4763–4768.

(146) Loipersberger, M.; Bertels, L. W.; Lee, J.; Head-Gordon, M. 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. *J. Chem. Theory Comput.* **2021**, *17*, 5582–5599.

(147) Evangelista, F. A. A driven similarity renormalization group approach to quantum many-body problems. *J. Chem. Phys.* **2014**, *141*, 054109.

(148) Keller, E.; Tsatsoulis, T.; Reuter, K.; Margraf, J. T. Regularized second-order correlation methods for extended systems. *J. Chem. Phys.* **2022**, *156*, 024106.

(149) Lee, J.; Head-Gordon, M. Two single-reference approaches to singlet biradicaloid problems: Complex, restricted orbitals and approximate spin-projection combined with regularized orbital-optimized Møller-Plesset perturbation theory. *J. Chem. Phys.* **2019**, *150*, 244106.

(150) DiStasio, Jr., R. A.; Steele, R. P.; Head-Gordon, M. *Mol. Phys.* **2007**, *105*, 2731.

(151) Karton, A.; Daon, S.; Martin, J. M. W4-11: A high-confidence benchmark dataset for computational thermochemistry derived from first-principles W4 data. *Chemical Physics Letters* **2011**, *510*, 165–178.

(152) Gauld, J. W.; Glukhovtsev, M. N.; Radom, L. The structure of the methanol radical cation: an artificially short C-O bond with MP2 theory. *Chemical Physics Letters* **1996**, *262*, 187–193.

(153) Mosley, J. D.; Young, J. W.; Huang, M.; McCoy, A. B.; Duncan, M. A. Infrared spectroscopy of the methanol cation and its methylene-oxonium isomer. *The Journal of Chemical Physics* **2015**, *142*, 114301.

(154) Moilanen, J.; Karttunen, A. J.; Tuononen, H. M.; Chivers, T. The Nature of Transannular Interactions in E4N4 and E82+ (E = S, Se). *Journal of Chemical Theory and Computation* **2012**, *8*, PMID: 26605588, 4249–4258.

(155) Helgaker, T.; Jørgensen, P.; Olsen, J., *Molecular Electronic-Structure Theory*; John Wiley & Sons: 2000.

(156) Lochan, R. C.; Head-Gordon, M. Orbital-optimized opposite-spin scaled second-order correlation: An economical method to improve the description of open-shell molecules. *J. Chem. Phys.* **2007**, *126*, 164101.

(157) Fowler, P.; Manolopoulos, D., *An Atlas of Fullerenes*; Dover books on chemistry; Dover Publications: 2007.

(158) Jiménez-Hoyos, C. A.; Rodrıguez-Guzmán, R.; Scuseria, G. E. Polyradical character and spin frustration in fullerene molecules: An ab initio non-collinear Hartree–Fock study. *J. Phys. Chem. A* **2014**, *118*, 9925–9940.

(159) Paul, P.; Kim, K.-C.; Sun, D.; Boyd, P. D.; Reed, C. A. Artifacts in the electron paramagnetic resonance spectra of C60 fullerene ions: inevitable C120O impurity. *Journal of the American Chemical Society* **2002**, *124*, 4394–4401.

(160) Lee, J.; Head-Gordon, M. Distinguishing artificial and essential symmetry breaking in a single determinant: approach and application to the C60, C36, and C20 fullerenes. *Phys. Chem. Chem. Phys.* **2019**, *21*, 4763–4778.

(161) Lum, Y.; Cheng, T.; Goddard III, W. A.; Ager, J. W. Electrochemical CO reduction builds solvent water into oxygenate products. *Journal of the American Chemical Society* **2018**, *140*, 9337–9340.

(162) Cheng, T.; Xiao, H.; Goddard III, W. A. Full atomistic reaction mechanism with kinetics for CO reduction on Cu (100) from ab initio molecular dynamics free-energy calculations at 298 K. *Proceedings of the National Academy of Sciences* **2017**, *114*, 1795–1800.

(163) Garza, A. J.; Bell, A. T.; Head-Gordon, M. Mechanism of CO2 reduction at copper surfaces: pathways to C2 products. *Acs Catalysis* **2018**, *8*, 1490–1499.

(164) Hori, Y.; Takahashi, R.; Yoshinami, Y.; Murata, A. Electrochemical reduction of CO at a copper electrode. *The Journal of Physical Chemistry B* **1997**, *101*, 7075–7081.

(165) Schouten, K.; Kwon, Y.; Van Der Ham, C.; Qin, Z.; Koper, M. A new mechanism for the selectivity to C 1 and C 2 species in the electrochemical reduction of carbon dioxide on copper electrodes. *Chemical Science* **2011**, *2*, 1902–1909.

(166) Bertheussen, E.; Verdaguer-Casadevall, A.; Ravasio, D.; Montoya, J. H.; Trimarco, D. B.; Roy, C.; Meier, S.; Wendland, J.; Nørskov, J. K.; Stephens, I. E., et al. Acetaldehyde as an intermediate in the electroreduction of carbon monoxide to ethanol on oxide-derived copper. *Angewandte Chemie* **2016**, *128*, 1472–1476.

(167) Clark, E. L.; Bell, A. T. Direct observation of the local reaction environment during the electrochemical reduction of CO2. *Journal of the American Chemical Society* **2018**, *140*, 7012–7020.

(168) Greenzaid, P.; Luz, Z.; Samuel, D. A nuclear magnetic resonance study of the reversible hydration of aliphatic aldehydes and ketones. I. Oxygen-17 and proton spectra and equilibrium constants. *Journal of the American Chemical Society* **1967**, *89*, 749–756.

(169) Jouny, M.; Luc, W.; Jiao, F. High-rate electroreduction of carbon monoxide to multi-carbon products. *Nature Catalysis* **2018**, *1*, 748–755.

(170) Shao, Y.; Gan, Z.; Epifanovsky, E.; Gilbert, A. T.; Wormit, M.; Kussmann, J.; Lange, A. W.; Behn, A.; Deng, J.; Feng, X., et al. Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Molecular Physics* **2015**, *113*, 184–215.

(171) Mardirossian, N.; Head-Gordon, M. $\omega$ B97M-V: A combinatorially optimized, range-separated hybrid, meta-GGA density functional with VV10 nonlocal correlation. *The Journal of chemical physics* **2016**, *144*, 214110.

(172) Mardirossian, N.; Head-Gordon, M. Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals. *Molecular Physics* **2017**, *115*, 2315–2372.

(173) Najibi, A.; Goerigk, L. The nonlocal kernel in van der Waals density functionals as an additive correction: An extensive analysis with special emphasis on the B97M-V and $\omega$B97M-V approaches. *Journal of Chemical Theory and Computation* **2018**, *14*, 5725–5738.

(174) Rappoport, D.; Furche, F. Property-optimized Gaussian basis sets for molecular response calculations. *The Journal of chemical physics* **2010**, *133*, 134105.

(175) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *The Journal of Physical Chemistry B* **2009**, *113*, 6378–6396.

(176) Behn, A.; Zimmerman, P. M.; Bell, A. T.; Head-Gordon, M. Efficient exploration of reaction paths via a freezing string method. *The Journal of chemical physics* **2011**, *135*, 224108.

(177) Singh, M. R.; Clark, E. L.; Bell, A. T. Effects of electrolyte, catalyst, and membrane composition and operating conditions on the performance of solar-driven electrochemical reduction of carbon dioxide. *Physical Chemistry Chemical Physics* **2015**, *17*, 18924–18936.

(178) Guthrie, J. P. Hydration of carbonyl compounds, an analysis in terms of multidimensional marcus theory. *Journal of the American Chemical Society* **2000**, *122*, 5529–5538.

(179) Birdja, Y. Y.; Koper, M. T. The importance of cannizzaro-type reactions during electrocatalytic reduction of carbon dioxide. *Journal of the American Chemical Society* **2017**, *139*, 2030–2034.

(180) Clark, E. L.; Hahn, C.; Jaramillo, T. F.; Bell, A. T. Electrochemical CO2 reduction over compressively strained CuAg surface alloys with enhanced multi-carbon oxygenate selectivity. *Journal of the American Chemical Society* **2017**, *139*, 15848–15857.

# Appendix A: GIAO matrix element of the paramagnetic spin-orbit coupling

To arrive at a convenient expression for matrix elements of the paramagnetic spin-orbit coupling operator, Eq. 2.9, in the complex GIAO basis associated with a finite applied $\mathrm{B}$-field and induced magnetic moment $\mathbf{m}_K$, we use the following shorthand notation for the charge-less three center field integral and the charge-less three center nuclear attraction integral for center $\mathbf{K}$:

$$F^x(l'+1) = \left\langle \omega_\mu | \frac{x_K}{|\mathbf{r}-\mathbf{R}_K|^3} | \omega_\nu(l'+1) \right\rangle \tag{1}$$

$$I(l'+1) = \left\langle \omega_\mu | |\mathbf{r}-\mathbf{R}_K|^{-1} | \omega_\nu(l'+1) \right\rangle \tag{2}$$

Here, $l'$ is the polynomial power associated with the $x$-component of the Gaussian orbital that is a part of the GIAO $\omega_\nu$. Similarly, $m'$ and $n'$ are used for the $y$- and $z$-components, respectively. $l$, $m$, $n$, refers to the GIAO $\omega_\mu$. Hence, $F^x(l'+1)$ is the field integral in the $x$-direction between GIAOs $\omega_\mu$ and $\omega_\nu$ and center K with increased orbital angular momentum ($x$-component) of $\omega_\nu$; similarly $I(l'+1)$ is the nuclear attraction integral between GIAOs $\omega_\mu$ and $\omega_\nu$ and center K with increased orbital angular momentum ($x$-component) of $\omega_\nu$.

## Field Integral approach

Expanding as sum of charge-less three center field integrals, we arrive at the following expression for matrix elements of the paramagnetic spin-orbit operator given in Eq. 2.9:

$$\begin{aligned} h_{\mu\nu}^{\text{PSO},x} = &- i\alpha^2 \sum_K m_{K,x} \times \\ &\Big[ n' F^y(n'-1) - \frac{i}{2}\chi_{\nu,z} F^y - 2\beta F^y(n'+1) \\ &- m' F^z(m'-1) + \frac{i}{2}\chi_{\nu,y} F^z + 2\beta F^z(m'+1) \Big] \end{aligned} \tag{3}$$

$$\begin{aligned} h_{\mu\nu}^{\text{PSO},y} = &- i\alpha^2 \sum_K m_{K,y} \times \\ &\Big[ l' F^z(l'-1) - \frac{i}{2}\chi_{\nu,x} F^z - 2\beta F^z(l'+1) \\ &- n' F^x(n'-1) + \frac{i}{2}\chi_{\nu,z} F^x + 2\beta F^x(n'+1) \Big] \end{aligned} \tag{4}$$

$$
\begin{aligned}
h_{\mu\nu}^{\text{PSO},z} = &- i\alpha^2 \sum_K m_{K,z} \times \\
&\Big[ m' F^x (m'-1) - \frac{i}{2} \chi_{\nu,y} F^x - 2\beta F^x (m'+1) \\
&- l' F^y (l'-1) + \frac{i}{2} \chi_{\nu,x} F^y + 2\beta F^y (l'+1) \Big]
\end{aligned} \tag{5}
$$

## Nuclear Attraction Integral approach

Alternatively, we can expand into sum of nuclear attraction integrals by explicitly treating the $|r^{-3}|$ operator. We employ the following equation:

$$
\int dx \frac{x_K}{|r_K|^3} \omega_\mu \omega_\nu = \int dx \frac{1}{|r_K|} \frac{\partial \omega_\mu}{\partial x} \omega_\nu + \int dx \frac{1}{|r_K|} \omega_\nu \frac{\partial \omega_\nu}{\partial x} \tag{6}
$$

After some straightforward but cumbersome algebra we arrive at the following expression for matrix elements of the diamagnetic shieldings operator given in Eq. 2.9:

$$
\begin{aligned}
h_{\mu\nu}^{\text{PSO,x}} = &- i\alpha^2 \sum_K m_{K,x} \big[ 4\alpha\beta I(m+1, n'+1) - 2\beta m I(m-1, n'+1) - i\beta \chi_{\mu,y} I(n'+1) \\
&- 2\alpha n' I(m+1, n'-1) + m n' I(m-1, n'-1) + \frac{i}{2} n' \chi_{\mu,y} I(n'-1) \\
&+ i\alpha \chi_{\nu,z} I(m+1) - \frac{i}{2} m \chi_{\nu,z} I(m-1) + \frac{1}{4} \left( \chi_{\mu,y} \chi_{\nu,z} - \chi_{\mu,z} \chi_{\nu,y} \right) I \\
&- 4\alpha\beta I(n+1, m'+1) + 2\beta n I(n-1, m'+1) + i\beta \chi_{\mu,z} I(m'+1) \\
&+ 2\alpha m' I(n+1, m'-1) - n m' I(n-1, m'-1) - \frac{i}{2} m' \chi_{\mu,z} I(m'-1) \\
&- i\alpha \chi_{\nu,z} I(l+1) + \frac{i}{2} l \chi_{\nu,z} I(l-1) \big]
\end{aligned} \tag{7}
$$

$$
\begin{aligned}
h_{\mu\nu}^{\text{PSO,y}} = &\frac{-i\alpha^2}{2} \sum_K m_{K,y} \big[ 4\alpha\beta I(n+1, l'+1) - 2\beta n I(n-1, l'+1) - i\beta \chi_{\mu,z} I(l'+1) \\
&- 2\alpha l' I(n+1, l'-1) + n l' I(n-1, l'-1) + \frac{i}{2} l' \chi_{\mu,z} I(l'-1) \\
&+ i\alpha \chi_{\nu,x} I(l+1) - \frac{i}{2} n \chi_{\nu,x} I(n-1) + \frac{1}{4} \left( \chi_{\mu,x} \chi_{\nu,z} - \chi_{\mu,z} \chi_{\nu,x} \right) I \\
&- 4\alpha\beta I(l+1, n'+1) + 2\beta n' I(l-1, n'+1) + i\beta \chi_{\mu,z} I(m'+1) \\
&+ 2\alpha m' I(n+1, m'-1) - n m' I(n-1, m'-1) - \frac{i}{2} m' \chi_{\mu,z} I(m'-1) \\
&- \frac{i}{2} \alpha \chi_{\nu,y} I(n+1) + \frac{i}{2} n \chi_{\nu,y} I(n-1) \big]
\end{aligned} \tag{8}
$$

$$
\begin{aligned}
h_{\mu\nu}^{\mathrm{PSO},z} =& \frac{-i\alpha^2}{2}\sum_K m_{K,z}\big[4\alpha\beta I(m+1,n'+1) - 2\beta m I(m-1,n'+1) - i\beta\chi_{\mu,y} I(n'+1) \\
& - 2\alpha n' I(m+1,n'-1) + m n' I(m-1,n'-1) + \frac{i}{2} n'\chi_{\mu,y} I(n'-1) \\
& + \frac{i}{2}\alpha\chi_{\nu,z} I(m+1) - \frac{i}{2} m\chi_{\nu,z} I(m-1) + \frac{1}{4}\left(\chi_{\mu,y}\chi_{\nu,z} - \chi_{\mu,z}\chi_{\nu,y}\right) I \\
& - 4\alpha\beta I(n+1,m'+1) + 2\beta n I(n-1,m'+1) + i\beta\chi_{\mu,z} I(m'+1) \\
& + 2\alpha m' I(n+1,m'-1) - n m' I(n-1,m'-1) - \frac{i}{2} m'\chi_{\mu,z} I(m'-1) \\
& - \frac{i}{2}\alpha\chi_{\nu,y} I(n+1) + \frac{i}{2} n\chi_{\nu,y} I(n-1)\big] \qquad (9)
\end{aligned}
$$

# Appendix B: GIAO matrix elements of the diamagnetic shielding

## Field Integral approach

The following are the expressions for matrix elements of the diamagnetic shieldings operator given in Eq. 2.8:

$$
\begin{aligned}
h_{\mu\nu}^{\mathrm{DS},x} =& \frac{\alpha^2}{2}\sum_K \{(B_y m_{K,y} + B_z m_{K,z}) F^x(l'+1) - B_x m_{K,y} F^x(m'+1) - B_x m_{K,z} F^x(n'+1) \qquad (10) \\
& + [(B_y m_{K,y} + B_z m_{K,z}) R_{\nu,x} - B_x m_{K,y} R_{\nu,y} - B_x m_{K,z} R_{\nu,z}] F^x)\} \qquad (11) \\
& \qquad (12)
\end{aligned}
$$

$$
\begin{aligned}
h_{\mu\nu}^{\mathrm{DS},y} =& \frac{\alpha^2}{2}\sum_K \{(B_x m_{K,x} + B_z m_{K,z}) F^y(m'+1) - B_y m_{K,x} F^y(l'+1) - B_y m_{K,z} F^y(n'+1) \qquad (13) \\
& + [(B_x m_{K,x} + B_z m_{K,z}) R_{\nu,y} - B_y m_{K,x} R_{\nu,x} - B_y m_{K,z} R_{\nu,z}] F^y)\} \qquad (14) \\
& \qquad (15)
\end{aligned}
$$

$$
\begin{aligned}
h_{\mu\nu}^{\mathrm{DS},z} =& \frac{\alpha^2}{2}\sum_K \{(B_x m_{K,x} + B_y m_{K,y}) F^z(n'+1) - B_x m_{K,z} F^z(l'+1) - B_z m_{K,y} F^z(m'+1) \qquad (16) \\
& + [(B_x m_{K,x} + B_y m_{K,y}) R_{\nu,z} - B_z m_{K,x} R_{\nu,x} - B_z m_{K,y} R_{\nu,y}] F^z)\} \qquad (17) \\
& \qquad (18)
\end{aligned}
$$

In the above expressions, $\alpha$,$\beta$ are the Gaussian exponents of the primitive GIAO's $\mu$,$\nu$ and $\boldsymbol{\chi}_\nu$ is the vector

$$\boldsymbol{\chi}_\nu = \mathbf{B} \times \mathbf{R}_\nu \tag{19}$$

Table .1: HF, MP2, $\kappa$-MP2, MP3, MP2.5, CCSD, CCSD(T) Isotropic NMR shielding values, with the aug-cc-pVTZ basis set, in units of ppm.

| Molecule | Nuclei | HF | MP2 | $\kappa$1.1-MP2 | $\kappa$1.45-MP2 | $\kappa$2-MP2 | MP3 | MP2.5 | CCSD | CCSD(T) |
|---|---|---|---|---|---|---|---|---|---|---|
| AlF | Al | 581.51 | 572.98 | 577.33 | 576.54 | 575.98 | 580.05 | 578.88 | 575.94 | 574.08 |
| AlF | F | 235.94 | 219.45 | 238.00 | 231.37 | 225.32 | 226.34 | 223.54 | 231.11 | 221.67 |
| $C_2H_4$ | C | 63.60 | 75.62 | 70.73 | 72.48 | 74.31 | 73.16 | 74.86 | 75.30 | 76.89 |
| $C_2H_4$ | H | 26.27 | 26.12 | 26.12 | 26.12 | 26.12 | 26.18 | 26.15 | 26.29 | 26.24 |
| $C_3H_4$ | C | 195.35 | 196.33 | 196.45 | 196.63 | 196.67 | 194.73 | 195.53 | 195.65 | 195.32 |
| $C_3H_4$ | C | 75.02 | 91.86 | 85.51 | 87.81 | 90.05 | 86.13 | 89.00 | 88.82 | 90.43 |
| $C_3H_4$ | H | 24.17 | 24.48 | 24.31 | 24.34 | 24.37 | 24.48 | 24.48 | 24.61 | 24.57 |
| $C_3H_4$ | H | 31.00 | 30.81 | 30.93 | 30.89 | 30.85 | 30.86 | 30.83 | 30.89 | 30.83 |
| $CH_2O$ | O | -427.82 | -311.87 | -308.17 | -304.45 | -304.27 | -396.15 | -352.40 | -361.53 | -357.70 |
| $CH_2O$ | C | -2.51 | 11.75 | 6.55 | 8.69 | 10.72 | 7.96 | 10.41 | 10.09 | 11.44 |
| $CH_2O$ | H | 22.53 | 22.12 | 22.36 | 22.28 | 22.21 | 22.27 | 22.21 | 22.30 | 22.12 |
| $CH_3F$ | C | 128.42 | 126.06 | 126.90 | 126.73 | 126.57 | 128.49 | 127.46 | 128.14 | 127.47 |
| $CH_3F$ | F | 486.13 | 488.10 | 489.63 | 489.09 | 488.52 | 482.88 | 485.43 | 483.10 | 482.16 |
| $CH_3F$ | H | 27.97 | 27.41 | 27.60 | 27.54 | 27.48 | 27.66 | 27.55 | 27.62 | 27.52 |
| $CH_4$ | C | 197.23 | 203.42 | 201.98 | 202.74 | 203.29 | 201.73 | 202.68 | 201.29 | 201.72 |
| $CH_4$ | H | 31.63 | 31.42 | 31.51 | 31.48 | 31.47 | 31.48 | 31.45 | 31.50 | 31.47 |
| CO | C | -20.62 | 18.43 | 6.08 | 11.19 | 15.65 | 4.51 | 11.98 | 7.87 | 12.67 |
| CO | O | -82.51 | -36.82 | -38.59 | -36.93 | -36.20 | -58.10 | -47.10 | -47.09 | -44.00 |
| FCCH | C | 178.48 | 185.47 | 182.91 | 184.10 | 185.14 | 181.70 | 183.97 | 182.45 | 183.57 |
| FCCH | C | 104.17 | 102.63 | 103.18 | 103.11 | 103.04 | 104.68 | 103.96 | 105.78 | 105.24 |
| FCCH | H | 30.61 | 30.62 | 30.62 | 30.60 | 30.63 | 30.68 | 30.65 | 30.69 | 30.69 |
| FCCH | F | 428.53 | 431.44 | 433.70 | 432.84 | 432.15 | 424.76 | 428.28 | 427.59 | 424.97 |
| FCN | F | 377.91 | 385.04 | 387.30 | 386.46 | 385.66 | 377.57 | 381.30 | 379.89 | 376.10 |
| FCN | C | 79.20 | 87.33 | 84.24 | 85.42 | 86.69 | 85.58 | 86.83 | 87.27 | 88.35 |
| FCN | N | 96.59 | 139.39 | 128.03 | 133.24 | 137.59 | 116.26 | 128.58 | 121.62 | 126.13 |
| $H_2C_2O$ | C | 191.62 | 202.03 | 198.87 | 200.37 | 201.09 | 196.14 | 199.30 | 195.97 | 197.09 |
| $H_2C_2O$ | C | -9.63 | 6.73 | 2.36 | 4.62 | 5.46 | -0.49 | 3.57 | 1.16 | 2.75 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $H_2C_2O$ | O | -20.23 | 38.40 | 33.24 | 36.91 | 39.03 | -10.92 | 14.33 | -0.21 | 4.58 |
| $H_2C_2O$ | H | 29.46 | 29.48 | 29.51 | 29.52 | 29.55 | 29.40 | 29.43 | 29.49 | 29.46 |
| $H_2O$ | O | 328.58 | 346.21 | 344.37 | 345.41 | 346.04 | 337.58 | 341.96 | 337.60 | 338.46 |
| $H_2O$ | H | 30.78 | 30.82 | 30.78 | 30.81 | 30.82 | 31.04 | 30.93 | 31.06 | 31.08 |
| $H_2S$ | S | 736.43 | 779.99 | 760.97 | 768.75 | 775.97 | 769.61 | 776.04 | 760.44 | 763.73 |
| $H_2S$ | H | 30.73 | 30.70 | 30.77 | 30.76 | 30.77 | 30.87 | 30.81 | 30.88 | 30.88 |
| $H_4C_2O$ | O | 378.59 | 373.80 | 378.49 | 377.10 | 375.09 | 364.17 | 368.98 | 367.71 | 365.37 |
| $H_4C_2O$ | C | 158.18 | 156.98 | 157.86 | 157.81 | 157.59 | 158.20 | 157.59 | 158.33 | 157.76 |
| $H_4C_2O$ | H | 29.79 | 29.26 | 29.49 | 29.44 | 29.36 | 29.49 | 29.37 | 29.43 | 29.34 |
| HCN | H | 29.28 | 29.07 | 29.20 | 29.16 | 29.11 | 29.22 | 29.15 | 29.21 | 29.14 |
| HCN | C | 74.30 | 92.51 | 84.88 | 87.60 | 90.29 | 86.46 | 89.88 | 88.71 | 90.92 |
| HCN | N | -44.18 | 9.57 | -4.02 | 1.92 | 6.95 | -15.24 | -2.18 | -7.56 | -3.52 |
| HCP | H | 30.18 | 29.57 | 29.93 | 29.80 | 29.68 | 29.84 | 29.70 | 29.83 | 29.70 |
| HCP | C | 17.11 | 50.33 | 35.73 | 39.83 | 44.28 | 35.85 | 43.77 | 41.12 | 45.36 |
| HCP | P | 355.04 | 419.03 | 369.68 | 380.38 | 395.44 | 395.21 | 410.52 | 402.68 | 412.83 |
| HF | H | 28.42 | 28.99 | 28.80 | 28.91 | 28.97 | 29.18 | 29.09 | 29.21 | 29.28 |
| HF | F | 414.35 | 425.29 | 425.66 | 425.67 | 425.53 | 419.03 | 422.19 | 419.41 | 419.84 |
| HFCO | O | -121.01 | -48.96 | -50.13 | -46.70 | -45.57 | -104.33 | -75.58 | -86.99 | -80.27 |
| HFCO | C | 37.92 | 47.85 | 44.91 | 46.20 | 47.39 | 44.54 | 46.60 | 46.07 | 47.39 |
| HFCO | F | 190.35 | 170.28 | 183.11 | 177.87 | 173.53 | 177.19 | 173.87 | 179.76 | 170.33 |
| HFCO | H | 24.47 | 23.97 | 24.22 | 24.14 | 24.06 | 24.23 | 24.13 | 24.17 | 24.03 |
| HOF | O | -128.97 | -39.27 | -56.95 | -46.78 | -40.56 | -41.78 | -40.14 | -52.67 | -53.43 |
| HOF | H | 19.25 | 20.12 | 19.90 | 20.05 | 20.13 | 20.70 | 20.41 | 20.27 | 20.09 |
| HOF | F | 292.25 | 203.09 | 241.07 | 227.22 | 214.62 | 197.91 | 200.65 | 210.20 | 199.51 |
| LiF | Li | 91.55 | 90.31 | 90.98 | 90.84 | 90.68 | 91.13 | 90.88 | 90.57 | 90.39 |
| LiF | F | 392.65 | 383.63 | 394.17 | 391.28 | 388.69 | 390.80 | 388.14 | 387.50 | 384.71 |
| LiH | H | 26.54 | 26.62 | 26.95 | 26.90 | 26.81 | 26.66 | 26.67 | 26.52 | 26.51 |
| LiH | Li | 90.10 | 89.99 | 90.13 | 90.15 | 90.18 | 90.40 | 90.38 | 90.03 | 90.03 |
| $N_2$ | N | -106.97 | -32.14 | -52.13 | -43.65 | -36.69 | -62.01 | -46.66 | -55.30 | -49.59 |
| $N_2O$ | N | 66.85 | 136.97 | 117.06 | 125.18 | 132.04 | 100.80 | 119.32 | 106.48 | 114.11 |
| $N_2O$ | N | -28.23 | 38.13 | 17.86 | 25.00 | 31.62 | 5.61 | 22.22 | 12.25 | 20.30 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $N_2O$ | O | 178.07 | 224.14 | 217.83 | 222.34 | 224.14 | 199.24 | 211.81 | 205.11 | 205.93 |
| $NH_3$ | N | 263.81 | 277.53 | 274.70 | 276.06 | 277.05 | 271.94 | 274.84 | 271.26 | 272.14 |
| $NH_3$ | H | 31.78 | 31.60 | 31.67 | 31.66 | 31.63 | 31.77 | 31.69 | 31.77 | 31.75 |
| OCS | O | 80.84 | 118.46 | 114.52 | 115.22 | 118.00 | 96.45 | 107.81 | 102.24 | 104.98 |
| OCS | C | 12.78 | 45.75 | 33.28 | 36.58 | 41.54 | 30.42 | 38.57 | 32.40 | 38.60 |
| OCS | S | 798.54 | 833.88 | 821.74 | 828.52 | 832.81 | 814.50 | 826.00 | 813.18 | 814.63 |
| $OF_2$ | O | -432.61 | -442.77 | -428.41 | -433.84 | -434.88 | -400.91 | -421.58 | -393.75 | -420.45 |
| $OF_2$ | F | 28.67 | -1.65 | 26.45 | 16.60 | 8.53 | -2.53 | -1.84 | 3.19 | -10.37 |
| PN | N | -499.37 | -231.62 | -317.89 | -287.00 | -260.10 | -385.09 | -306.45 | -345.26 | -324.24 |
| PN | P | -83.57 | 158.63 | 31.19 | 64.17 | 103.17 | 40.11 | 104.77 | 68.34 | 93.97 |
| $SO_2$ | S | -342.20 | -57.94 | -153.82 | -119.20 | -85.95 | -174.38 | -110.93 | -159.58 | -122.27 |
| $SO_2$ | O | -343.14 | -182.77 | -190.56 | -185.40 | -178.34 | -276.89 | -228.92 | -241.44 | -233.88 |