# Expanding the Toolkit for Optical Rotation Calculations: Development of Interpretive Models and Extension to Solid State Systems



Ty Balduf

B.S. Chemistry/Mathematics, Saint Ambrose University, 2016

Submitted to the graduate degree program in the Department of Chemistry and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Committee members

Marco Caricato, Chairperson

James Blakemore

Brian Laird

Ward Thompson

Hartwin Peelaers

Date defended: March 29, 2022

The Dissertation Committee for Ty Balduf certifies
that this is the approved version of the following dissertation :

Expanding the Toolkit for Optical Rotation Calculations: Development of Interpretive Models and Extension to Solid State Systems

Marco Caricato, Chairperson

Date approved: March 31, 2022

# Abstract

The measurement of optical rotation (OR) is a long standing approach for characterizing chiral systems. The ability to predict optical rotation theoretically without resorting to expensive *ab initio* simulations would greatly aid in absolute configuration assignment of chiral species and would help develop a more chemically intuitive sense of the physical processes that drive this phenomenon. However, it has proven challenging to develop predictive models for OR, not only due to the difficulty of computing the intrinsic OR arising from the structure of a molecule, but also the need to disentangle it from optical activity induced by the surrounding environment. We have worked to unravel these competing effects by developing models to decompose the intrinsic OR into more interpretable structural contributions and by extending quantum mechanical methods to compute OR to solid state systems, allowing for a more thorough analysis of environmental effects.

We have worked to extend the $\tilde{S}$ analysis framework, which decomposes the OR into occupied-virtual orbital pair contributions, to make it more generally applicable. We calculated OR for a small test of molecules in both the modified velocity gauge (MVG) and length gauge (LG) in order to determine how the physical interpretation provided by $\tilde{S}$ analysis is affected by the gauge. We found that the distribution of $\tilde{S}$ contributions was consistent across different gauges. The $\tilde{S}$ values from different gauges also converged at a similar rate to the total OR when summing contributions from largest in magnitude to smallest. Examining the electric and magnetic vectors associated with the largest magnitude transitions, we found that each gauge gave a consistent interpretation of the dominant physical processes driving the OR.

We have also worked to remove the unphysical origin dependence from computed $\tilde{S}$ values. To this end, we derived two origin invariant MVG formulations of $\tilde{S}$: $\tilde{S}^{\mathrm{Avg}}$, which averages $\tilde{S}$ computed with an electric (MVG-E) and magnetic (MVG-M) perturbation $\tilde{S}$ to cancel out their opposite sign, equal magnitude origin dependence, and $\tilde{S}^{\mathrm{Hemi}}$, which eliminates the origin depen-

dence by contracting together "hemi-perturbed" densities formed via a Cholesky decomposition of the response matrix. We tested these methods on two small organic molecules and confirmed that $\tilde{S}$ remained unchanged at a displaced origin. While further work is needed to develop associated electric and magnetic vectors for these origin invariant $\tilde{S}$, they can already be used to detect significant deviation due to the chosen origin.

As a first step towards developing quantum mechanical methods to compute the OR of solids, we have created a benchmark set of helical chains of diatomic molecules, for which the OR can be computed using existing molecular methods. We probed the effect of cell size, cell spacing, and helical orientation on the computed OR. We found that even for these simple models, convergence to the macroscopic limit was slow with system size and the effect of helical orientation could not be readily reproduced using the semi-empirical Kirkwood model. Our results made clear the need for a periodic, quantum mechanical methods to compute the full OR tensor, including both the electric dipole-magnetic dipole and electric dipole-electric quadrupole polarizability contributions.

To aid in the development of highly accurate methods for computing solid state OR, we compared MVG and origin invariant length gauge [LG(OI)] calculations of the full OR tensor using coupled cluster with single and double excitations (CCSD) and Density Functional Theory (DFT) methods. We found that DFT and CCSD components of the OR tensor were well correlated, though DFT could significantly overestimate these values depending on the chosen functional. CCSD MVG and LG(OI) tensor components were highly correlated, but the extent of correlation showed a high dependence on the inclusion/exclusion of outliers. Our results suggest that some care must still be taken with the choice of gauge, as for some cases CCSD may not provide a sufficient treatment of electron correlation to achieve gauge invariance.

We have derived translation invariant forms for the magnetic dipole and electric quadrupole in order to implement the periodic, full OR tensor into Gaussian. While a periodic form of the magnetic dipole has recently been published, a derivation of the periodic electric quadrupole is reported for the first time in this work. Once the implementation is complete, we aim to distinguish intrinsic and environmental influences on the OR via simulations of molecular crystals.

# Acknowledgements

I want to first thank my advisor, Dr. Marco Caricato, for his mentorship throughout my PhD. Starting out, it was tough for me to imagine developing as a scientist to the point where I could formulate novel scientific questions and publish the resulting research, but you challenged me to step out of my comfort zone and develop the confidence to carry out independent research. You allowed me the opportunity to travel to present my work and see the wide breadth of our field. You also helped me to learn the joys and pains of developing scientific software, which has set me towards an exciting start to my scientific career.

I would also like to thank our research group. To the senior students that came before me (Tal, Amy, Matt, Sijin) you really helped me to get my footing here at KU. Moving down to Kansas and starting computational research was a big change for me and it was made exponentially easier by how inviting and helpful you were with getting me integrated into the group. To the students I have gotten to collaborate with or mentor (Kaihua, Isaac, Taylor, Brian, Katherine, Tasmin, Delaney, and many others), I hope I left some sound advice and just the right amount of broken code for you. It has been great to get to know you all, in person and through many, many Zoom calls.

To my family, I can't thank you enough for helping to keep me sane throughout this whole process. It was tough to move away to Kansas, but I always knew I was in your thoughts. Visits with you really helped me to relax and step away from research for awhile. I know you didn't necessarily understand everything I was working on (hopefully I have gotten closer to understanding it myself), but it really bolstered my confidence to see that you were proud of me.

None of this would be possible without my lovely wife, Mary. From the very beginning of our marriage, you did whatever you could to support me through graduate school, even moving away to Kansas the day after our wedding. Through all the good and bad life/world events since then, you have been there for me and helped me get through them. There were stressful times and points

where I really didn't believe I could ever finish my doctorate, but it always put my mind at ease to know that you were with me and believed in me. Over these past 6 years (thank you again for your patience with me), we have made new friends, traveled, and started our own small family. With our time in Lawrence coming to end, I can't wait to see what the years to come have in store for us. I love you Mary ♡

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

The ability to detect and control the chirality of molecules has a wide array of applications across chemical and biological disciplines, most notably drug design/analysis,[1–3] asymmetric catalysis,[4–8] and chiral supramolecular self-assembly.[9–12] The chiral absolute configuration or "handedness" of a molecule has a substantial effect on how it interacts with other chiral molecules and, more generally, how it behaves in a chiral environment. Various forms of chromatography exploit these chiral interactions in order to separate mixtures and quantify the enantiomeric excess in a sample.[13] The earliest attempt at separating chiral species by Louis Pasteur in 1848 relied on visual inspection to distinguish crystal polymorphs preferentially formed by a specific enantiomer.[14,15] Optical techniques for probing chirality have advanced considerably since then: X-ray crystallography can provide high resolution structures of crystallizable chiral molecules,[16,17] while vibrational/electronic circular dichroism (VCD/ECD) allow for *in situ* spectroscopic analysis of the nuclear/electronic structure.[18–20] Characterizing enantiomers experimentally requires a careful separation, typically either by chiral chromatography or chiral controlled crystallization.[21,22] However, crystallization can be a laborious and cost-intensive process, and many molecules can not be reliably crystallized. While chromatographic separation and subsequent characterization using chiroptical spectroscopy is generally possible, theoretical calculations are required to assign the absolute configuration of each species without a crystal structure.

The measurement of optical rotation (OR), the change in polarization of light as it passes through a chiral medium, was one of the first techniques developed to spectroscopically characterize chiral systems.[23–25] This change in polarization can be ascribed to differential refraction of left and right circularly polarized light, as the chiral arrangement of the molecule causes it to interact

differently with the chiral components of the electromagnetic field, see Figure 1.1. Despite the long-standing use of OR measurements and the conceptual simplicity of this phenomenon, it has proven challenging to formulate a general, chemically intuitive model of how molecular structure affects the observed OR. Given a pair of enantiomers, it is not obvious from their structure what the magnitude of the OR will be, or even which direction each enantiomer will rotate impinging light. Probing this relationship is further complicated by the fact that the observed OR does not typically arise from a single, isolated molecule. Most OR measurements are performed in solution phase, where the observed effect is not just the result of the minimum energy structure, but a conformational average of thermally excited structures, along with the solvent-specific chiral solvation shells that form around them. More generally, any condensed phase measurement of the OR (or VCD/ECD) of a compound is a result of a nonadditive combination of its intrinsic OR and OR induced through intermolecular interactions with the surrounding chiral environment. The difficulty of disentangling these effects experimentally has spurred efforts to computationally model chiroptical spectroscopy.[26–32]

To better understand the underlying, intrinsic OR of a molecule, methods have been developed to decompose molecular optical rotation into more interpretable contributions arising from structural features. Polavarapu et al. tested the applicability of the semi-empirical Kirkwood model, which expresses the OR in terms of bond polarizabilities and anisotropies.[33,34] A similar effort was made by Beratan et al., which split the OR tensor into atomic and bond contributions.[35] Kahr et al. showed that the optical activity of conjugated hydrocarbons could be rationalized using semi-empirical Huckel theory.[36] Several schemes have been developed to assign OR contributions to individual orbitals or orbital pairs. Grimme et al.'s approach generates so called natural response orbitals using a transformation derived from the Singular Value Decomposition (SVD) of the perturbed density.[37] Autschbach et al. developed a method which localizes the orbitals onto lone pairs and bonds, expressing the OR as a sum of individual orbital contributions.[38] Our group's $\tilde{S}$ analysis is a similar approach to Autschbach, but expresses the OR in terms of occupied-virtual *ia* orbital pair contributions, $\tilde{S}_{ia}$.[39] While the initial implementation of $\tilde{S}$ analysis employed the canonical

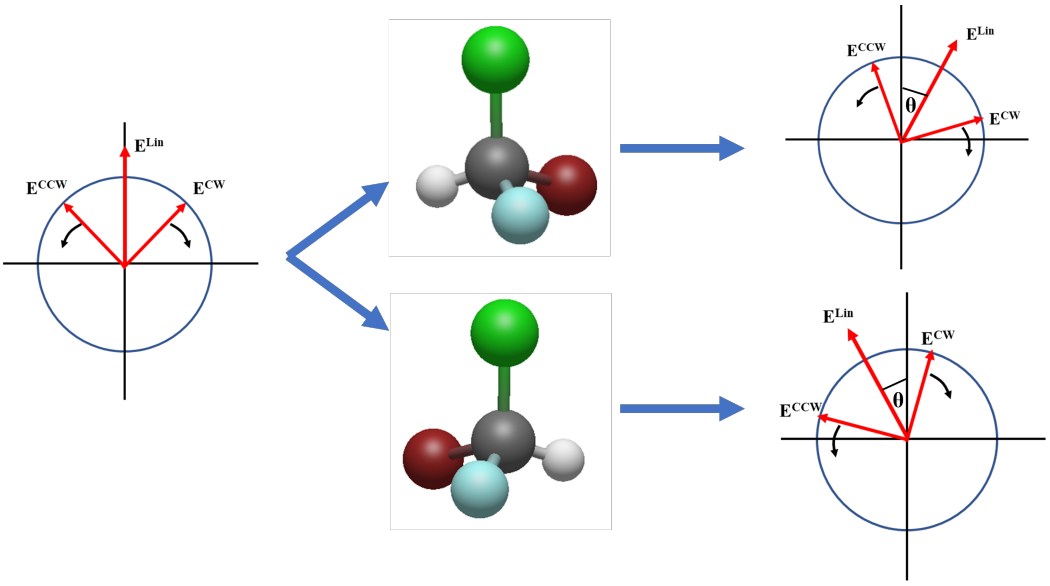


Figure 1.1: Schematic representation of an optical rotation measurement. On the left, incident plane polarized light $E^{\mathrm{Lin}}$ is formed from a combination of clockwise $E^{\mathrm{CW}}$ and counterclockwise $E^{\mathrm{CCW}}$ polarized light. After passing through a chiral sample, the plane of polarization of the exiting light is rotated by $\theta$ degrees due to differential refraction of the circularly polarized components. The direction of rotation is determined by which enantiomer is present in the sample while the magnitude depends on the sample concentration and the path length through which the light travels.

MOs, the method is also applicable with orbital localization or any other unitary transformation within the occupied or virtual MO subspaces.[40]

A number of different attempts have been made to characterize how the environment influences the OR. The Vaccaro group has used cavity ring-down polarimetry (CRPD) to measure OR in the gas phase, removing solvent effects from the measurements.[30] A combined experimental and computational study found that the shift in OR due to solvation was substantial, regardless of the type of solvent (e.g. polar/nonpolar), and that an implicit solvation model could not quantitatively reproduce the experimental shifts.[41] However, with currently feasible combinations of model chemistry and system size, even incorporating explicit solvation has proven insufficient to reproduce experimental OR.[42–44] The Kahr group has experimentally studied the optical activity of chiral crystals.[45] Unlike in solution, the orientation of a chiral molecule in a solid is fixed, allowing measurements of the whole OR tensor. The rigidity of a solid also fixes the environment around a molecule, simplifying the process of replicating it in a simulation. Computational results

for the OR of solids have been limited to a small number of semi-empirical calculations based on classical polarizability theory.[46–48] In first-principles calculation, optical rotation is derived by computing the Buckingham-Dunn OR tensor, which is formed via a combination of the electric dipole-magnetic dipole and electric dipole-electric quadrupole polarizabilities of the molecule.[49] While standard methods are available to compute these quantities for molecules, only recently have *periodic*, first-principles calculations of OR been implemented and thus far only the electric dipole-magnetic dipole component is included.[50]

Our goals with this work are (1) to refine existing computational tools for structural interpretation of the intrinsic OR and (2) to develop methods for fully first-principles, periodic calculations of OR to enable further characterization of the influence of environmental effects. This work is organized as follows. Before delving into each project, in Chapter 2 we first briefly review the requisite theory underpinning OR (i.e. the choice of electromagnetic gauge and the linear response formalism for evaluating polarizabilities), as well as periodic boundary condition (PBC) electronic structure theory. We then present our research in two parts: In Part I, we describe our efforts to extend and refine the existing $\tilde{S}$ analysis framework, while in Part II, we recount our progress towards an implementation of the full OR tensor for general periodic systems. We conclude with a brief summary of our results and a discussion of how we expect chiroptical research to progress in the years to come.

Part I begins with a discussion of our new implementations of the $\tilde{S}$ method using the modified velocity gauge for both electric and magnetic perturbations, in Chapter 3. The $\tilde{S}$ analysis approach was initial developed using the length gauge, but it is not immediately obvious if the interpretation provided by this framework is dependent on the choice of gauge. We constructed a test set of molecules previously analyzed with the length gauge (magnetic perturbation) implementation of $\tilde{S}$ and compared each method to determine whether they provided the same picture of what underlying physical processes are inducing OR.

A potential problem with the $\tilde{S}$ framework is that the computed values have an unphysical dependence on the coordinate origin, potentially distorting the physical interpretation. Chapter

4 presents two different approaches for removing this origin dependence: one which averages the MVG $\tilde{S}$ computed with an electric and a magnetic perturbation and one which forms the $\tilde{S}$ values using "hemi-perturbed densities" obtained by splitting the response matrix with a Cholesky decomposition. We test these methods using calculations at a displaced origin and discuss their relative benefits and drawbacks as a means of analyzing the OR.

In Part II, Chapter 5 showcases a benchmark set of long helical chains of diatomic molecules that we use as a model to study solid state optical rotation using existing molecular electronic structure methods. We tested the convergence of the OR towards the macroscopic limit with system size and assessed whether a semi-empirical method could adequately describe the dependence of the OR on the helical orientation. For each of these models, we examined the contribution of both the electric dipole-magnetic dipole and the electric dipole-electric quadrupole polarizability to the full OR tensor. The electric dipole-electric quadrupole polarizability is typically not computed when simulating isotropic systems, but may have a significant impact on the OR tensor components of an oriented system.

In Chapter 6, we compare MVG and LG(OI) methods for computing an origin independent, full optical rotation tensor using coupled cluster with single and double excitations (CCSD) and density functional theory (DFT) methods. Since the full OR tensor is not typically computed for isotropic systems, it is unclear how the level of theory and chosen gauge affect these calculations. We computed OR for a test set of 22 small/medium sized organic molecules and assessed the correlation of OR tensor elements between CCSD and DFT methods, as well as between MVG and LG(OI) CCSD calculations.

In Chapter 7, we derive expressions for a translation invariant magnetic dipole and electric quadrupole in order to compute OR for periodic systems. We found that our derivation of the periodic magnetic dipole agrees with the recently published derivation by Rerát and Kirtman.[50] Our derivation of the periodic electric quadrupole is, to the best of our knowledge, the first for this property. We discuss details of the implementation in GAUSSIAN and review the results of OR calculations on a simple 1D hydrogen peroxide chain.

# Chapter 2

# Theory

Optical rotation is the differential refraction of plane polarized light by a chiral sample. Formally, OR is a result of the Buckingham/Dunn optical activity tensor **B**, which is defined as:[49,51,52]

$$B_{\alpha\beta} = \frac{1}{2}\left[\beta_{\alpha\beta} + \beta_{\beta\alpha} + \frac{1}{3}\sum_{\delta,\gamma}\left(\varepsilon_{\alpha\gamma\delta}A_{\gamma,\delta\beta} + \varepsilon_{\beta\gamma\delta}A_{\gamma,\delta\alpha}\right)\right] \tag{2.1}$$

where $\boldsymbol{\varepsilon}$ is the Levi-Civita operator, $\boldsymbol{\beta}$ is the electric dipole-magnetic dipole polarizability tensor, and **A** is the electric dipole-electric quadrupole polarizability tensor. In the length gauge,

$$\beta^{L}_{\alpha\beta} = 2\sum_{j\neq 0}\mathrm{Im}\frac{\langle\psi_0|\mu^{L}_{\alpha}|\psi_j\rangle\langle\psi_j|m_\beta|\psi_0\rangle}{\omega_j^2-\omega^2} = \mathrm{Im}\langle\!\langle \mu^{L}_{\alpha};m_\beta\rangle\!\rangle_\omega \tag{2.2}$$

$$A^{L}_{\alpha,\beta\gamma} = 2\sum_{j\neq 0}\omega_j\mathrm{Re}\frac{\langle\psi_0|\mu^{L}_{\alpha}|\psi_j\rangle\langle\psi_j|\Theta^{L}_{\beta\gamma}|\psi_0\rangle}{\omega_j^2-\omega^2} = \mathrm{Re}\langle\!\langle \mu^{L}_{\alpha};\Theta^{L}_{\beta\gamma}\rangle\!\rangle_\omega \tag{2.3}$$

while in the velocity gauge

$$\beta^{V}_{\alpha\beta} = \frac{2}{\omega}\sum_{j\neq 0}\mathrm{Re}\frac{\langle\psi_0|\mu^{V}_{\alpha}|\psi_j\rangle\langle\psi_j|m_\beta|\psi_0\rangle}{\omega_j^2-\omega^2} = \mathrm{Re}\langle\!\langle \mu^{V}_{\alpha};m_\beta\rangle\!\rangle_\omega \tag{2.4}$$

$$A^{V}_{\alpha,\beta\gamma} = \frac{2}{\omega}\sum_{j\neq 0}\omega_j\mathrm{Im}\frac{\langle\psi_0|\mu^{V}_{\alpha}|\psi_j\rangle\langle\psi_j|\Theta^{V}_{\beta\gamma}|\psi_0\rangle}{\omega_j^2-\omega^2} = \mathrm{Im}\langle\!\langle \mu^{V}_{\alpha};\Theta^{V}_{\beta\gamma}\rangle\!\rangle_\omega \tag{2.5}$$

where $\langle\!\langle\cdot;\cdot\rangle\!\rangle$ is the response function of the enclosed operators. The distinction between the length and velocity gauges will be discussed in more detail below.

Throughout this thesis, we use atomic units unless otherwise specified and denote tensors over

Cartesian indices using bold font and matrices over an orbital basis with an underline. Since there has been some confusion in the literature regarding the use of the $\mathbf{G}'$ symbol to indicate the electric dipole-magnetic dipole polarizability with or without the $\omega^{-1}$ factor, we use the notation $\boldsymbol{\beta} = -\omega^{-1}\mathbf{G}'$.[51] Greek indices denote Cartesian coordinates, $\omega$ is the frequency of the incident electromagnetic radiation while $|\psi_j\rangle$ and $\omega_j$ are the $j^{\text{th}}$ excited state wave function and excitation frequency, respectively. These definitions for $\boldsymbol{\beta}$ and $\mathbf{A}$ are valid for non-resonant optical activity ($\omega_j \neq \omega$) calculations; resonant optical activity is discussed in greater detail elsewhere.[29,49,53,54]

For isotropic media, the observed optical rotation is commonly reported as a normalized quantity in units of deg [dm (g/mL)]$^{-1}$, known as specific rotation:

$$[\alpha]_\omega = \frac{(72\times 10^6)\hbar^2 N_A \omega^2}{3c^2 m_e^2 M}\mathrm{Tr}(\mathbf{B}) \tag{2.6}$$

where $\hbar$ is the reduced Planck's constant (J s), $N_A$ is Avogadro's number, $c$ is the speed of light (m/s), $m_e$ is the electron rest mass (kg), and $M$ is the molecular mass (amu). Note that Eq. 2.6 depends on $\omega$ both through the constant prefactor on the right hand side as well through $\mathbf{B}$. For an oriented system, the $\frac{1}{3}\mathrm{Tr}(\mathbf{B})$ factor should be replaced by $\beta_{\mathbf{n}}$, the average of the components of the $\mathbf{B}$ tensor that are perpendicular to the direction of the beam.[53] By defining a chiroptical response tensor $\boldsymbol{\mathcal{B}}$:[51]

$$\mathcal{B}_{\alpha\beta} = \frac{1}{2}\left[\mathrm{Tr}(\mathbf{B})\delta_{\alpha\beta} - B_{\alpha\beta}\right] \tag{2.7}$$

the OR parameter for a light beam in the direction of the unit vector $\mathbf{n}$ is:

$$\beta_{\mathbf{n}} = \mathbf{n}^\dagger \boldsymbol{\mathcal{B}}\mathbf{n} \tag{2.8}$$

By defining this alternative chiroptical response tensor, one can easily compute the specific rotation for light traveling in an arbitrary direction relative to the oriented material.[51,53]

## 2.1 Gauge Theory

In order to simulate optical rotation, one must, at least implicitly, define a representation for light in the calculation. Maxwell's equations allow one to represent light as an electric field **E** and magnetic field **B**.[55,56] Alternatively, Maxwell's equations can be expressed in terms of a scalar potential $\phi$ and a vector potential **A** (not to be confused with the dipole-quadrupole polarizability in Eq. 2.3). However, there is some amount of ambiguity in the choice of $\phi$ and **A**; given any twice differentiable function of time and position $f(r,t)$, we can define

$$\begin{aligned} \phi' &= \phi - \frac{\partial f}{\partial t} \\ \mathbf{A}' &= \mathbf{A} + \nabla f \end{aligned} \tag{2.9}$$

such that $\phi'$ and $\mathbf{A}'$ produce the same observable electromagnetic field.[57,58] Different choices of the function $f$ are referred to as gauges. Gauges are typically defined either by imposing constraints on $\phi$ and/or **A** or by directly specifying a gauge transformation function $f(r,t)$ from a known gauge. For example, the so-called velocity dipole gauge is defined by the condition

$$\nabla \cdot \mathbf{A}^V = -\frac{1}{\mathbf{v}^2}\frac{\partial \phi^V}{\partial t} \tag{2.10}$$

where **v** is the velocity and the $V$ superscript denotes the velocity gauge.[56,59] The length dipole gauge can then be defined using the gauge transformation

$$f(r,t)^{V\to L} = -\mathbf{r}\cdot\mathbf{A}^V \tag{2.11}$$

where the superscript $V \to L$ labels this as the transformation from the velocity to the length gauge.[56,59]

Both the length and velocity gauges are used extensively in OR calculations. The multipole operators in Eqs. 2.2-2.3 are expressed in the length (superscript $L$) and velocity (superscript $V$)

dipole gauges as:

$$\begin{aligned} \boldsymbol{\mu}^L &= -\boldsymbol{r} \\ \boldsymbol{m}^L &= \frac{i}{2}(\boldsymbol{r} \times \nabla) \\ \Theta^L_{\beta\gamma} &= -\frac{1}{2}(3r_\beta r_\gamma - \delta_{\beta\gamma} r^2) \end{aligned} \tag{2.12}$$

$$\begin{aligned} \boldsymbol{\mu}^V &= -\boldsymbol{p} \\ \boldsymbol{m}^V &= \frac{i}{2}(\boldsymbol{r} \times \nabla) \\ \Theta^V_{\beta\gamma} &= -\frac{1}{2}\left[(3r_\beta p_\gamma + 3p_\beta r_\gamma) - \delta_{\beta\gamma}(\boldsymbol{r}\cdot\boldsymbol{p} + \boldsymbol{r}\cdot\boldsymbol{p})\right] \end{aligned} \tag{2.13}$$

which are respectively the electric dipole, the magnetic dipole, and the traceless electric quadrupole operators, with the position $\boldsymbol{r}$, momentum $\boldsymbol{p}$, and gradient $\nabla$ operators implicitly summed over all the electrons of the molecule.

It is not immediately obvious why the length and velocity forms of these operators should produce the same optical rotation. The equivalence of these approaches can be more readily seen using the response function identity[29,60]

$$\omega\langle\langle R;S\rangle\rangle_\omega = \langle\langle [R,H];S\rangle\rangle_\omega + \langle [R,S]\rangle \tag{2.14}$$

By taking $R = \boldsymbol{r}$, $S = \boldsymbol{m}$ and noting that $[\boldsymbol{r},H] = i\boldsymbol{p}$ and $[\boldsymbol{r},\boldsymbol{m}] = 0$, we see that $\boldsymbol{\beta}$ computed in the length and velocity gauge should be equivalent. The same identity can be used to show that $\mathbf{A}$ is equivalent in the velocity and length gauges.

This response function relationship will hold for an exact calculation, but also for a variational electronic structure method (e.g. Hartree-Fock, DFT) in a complete basis.[29,58,60] However, not only will OR computed in the length and velocity gauges generally differ when using a finite basis or a nonvariational method (e.g. coupled cluster), but each gauge also introduces unphysical computational artifacts that must be corrected to obtain reasonable results. OR computed in the length gauge has an unphysical dependence on the coordinate origin arising from the magnetic dipole. This is typically corrected by using gauge including atomic orbitals (also referred to as

London orbitals), which include a magnetic field dependent phase factor for each basis function that exactly counteracts the origin dependent term.[61] More recently, the origin independent length gauge [LG(OI)] was developed, which transforms the OR tensor to an orientation where it is origin invariant.[62] This method is described in more detail in Chapter 6. The velocity gauge, while origin invariant even for approximate electronic structure calculations, exhibits nonzero OR for a static perturbation. The so called modified velocity gauge (MVG) corrects for this by simply subtracting this static OR from that computed at the desired frequency.[60]

## 2.2 Linear Response Formalism

While the electric dipole-magnetic dipole and electric dipole-electric quadrupole polarizabilities, $\boldsymbol{\beta}$ and $\mathbf{A}$, are formally defined using the sum-over-states (SOS) expressions in Eqs. 2.2-2.5, these are rarely used in practice, as the sums converge very slowly and non-monotonically.[63] A much more computationally efficient route to computing these tensors is to employ linear response. Linear response allows us to determine a variationally optimal set of wavefunction parameters upon applying a perturbation to some initial wavefunction. From these wavefunction parameters, we can calculate the frequency dependent polarizability of any given property in response to a linear perturbation. The following derivation closely follows that given by McWeeny.[64] To determine these parameters, one starts from the Frenkel variational principle[65]

$$\langle \delta\Psi | \mathcal{P}\left(H - i\hbar\frac{\partial}{\partial t}\right)|\Psi\rangle + \text{complex conjugate} = 0 \tag{2.15}$$

where $\Psi$ is a normalized self-consistent field (SCF) wavefunction, $\delta\Psi$ is some arbitrary variation of that wavefunction, and $\mathcal{P}$ is a projector onto the orthogonal complement of $\Psi$

$$\mathcal{P} = 1 - |\Psi\rangle\langle\Psi| \tag{2.16}$$

which allows us to consider only variations that are orthogonal to $\Psi$. The time dependent Hamiltonian can be partitioned as

$$H(t) = H_0 + H'(t) = H_0 + F(t)R \tag{2.17}$$

where $H_0$ is time independent and $H'(t)$ is a time dependent perturbation, expressed as field strength $F(t)$ times a perturbation operator $R$. $\Psi$ depends on a vector of parameters $\underline{\rho}$, which in turn can be partitioned as

$$\underline{\rho}(t) = \underline{\rho}_0 + \underline{d}(t) \tag{2.18}$$

where $\underline{\rho}_0$ are the optimal parameters in the absence of the perturbation and $\underline{d}$ is the time dependent deviation in response to the perturbation. Using this partitioning, we can expand $\Psi$ and $\delta\Psi$ to first order in $\underline{d}$ as[64]

$$\begin{aligned} |\Psi\rangle &= |\Psi_0\rangle + |(\nabla\Psi)_0\rangle\underline{d} \\ |\delta\Psi\rangle &= \left[|(\nabla\Psi)_0\rangle + |(\nabla\nabla\Psi)_0\rangle\underline{d}\right]\delta\underline{\rho} \end{aligned} \tag{2.19}$$

Where $\nabla$ is the gradient with respect to the parameters. Substituting these expressions into Eq. 2.15 and keeping only the terms linear in the perturbation, the variational condition can now be expressed as

$$\nabla\underline{H'} + \underline{M}\underline{d} + \underline{Q}\underline{d}^* = i\hbar\dot{\underline{d}} \tag{2.20}$$

where we have collected several terms into matrices as shown

$$\begin{aligned} \underline{H'} &= \langle(\Psi\nabla)|\mathcal{P}H'|\Psi\rangle \\ \underline{M} &= \langle(\Psi\nabla)|\mathcal{P}H|(\nabla\Psi)\rangle \\ \underline{Q} &= \langle(\Psi\nabla\nabla)|\mathcal{P}H|\Psi\rangle \end{aligned} \tag{2.21}$$

Converting Eq. 2.20 to the frequency domain, we can rewrite the deviation vector for a particular perturbation frequency $\omega$ as

$$\underline{d} = \underline{X}e^{-i\omega t} + \underline{Y}^* e^{i\omega t} \tag{2.22}$$

where $\underline{X}$ and $\underline{Y}$ are amplitude vectors, each multiplied by a phase factor. Inserting this into Eq. 2.20 and rearranging gives

$$\left[\begin{pmatrix} \underline{M} & \underline{Q} \\ \underline{Q}^* & \underline{M}^* \end{pmatrix} - \omega \begin{pmatrix} \underline{1} & \underline{0} \\ \underline{0} & -\underline{1} \end{pmatrix}\right]\begin{pmatrix} \underline{X} \\ \underline{Y} \end{pmatrix} = \underline{\Omega}\begin{pmatrix} \underline{X} \\ \underline{Y} \end{pmatrix} = \begin{pmatrix} \nabla \underline{R} \\ \underline{R} \nabla \end{pmatrix} \tag{2.23}$$

where $\underline{1}$ and $\underline{0}$ are identity and zero matrices, respectively and $^*$ denotes the complex conjugate. After solving this system of equations for $\underline{X}$ and $\underline{Y}$, which we can think of as perturbed densities with respect to the perturbation $R$, we can form the frequency dependent polarizability $\Pi_{RS}$ of the property $S$ perturbed by $R$ via the contraction

$$\Pi_{RS} = \begin{pmatrix} \nabla \underline{S} \\ \underline{S} \nabla \end{pmatrix}^{\dagger} \begin{pmatrix} \underline{X}_R \\ \underline{Y}_R \end{pmatrix} \tag{2.24}$$

Letting $R = \mu$, we can compute $\boldsymbol{\beta}$ by letting $S = m$ and $\mathbf{A}$ by letting $S = \Theta$.

For a general SCF wavefunction, we can write the matrix elements more explicitly as

$$Q_{ia,jb} = (aj|ib) - c_{HF}(ab|ij) - (aj|w|ib) \tag{2.25}$$

$$M_{ia,jb} = \delta_{ij}\delta_{ab}(\varepsilon_a - \varepsilon_i) + Q_{ia,jb} \tag{2.26}$$

$$R_{ia} = (i|R|a) \tag{2.27}$$

with $i, j$ denoting occupied orbitals and $a, b$ denoting virtual orbitals, $w$ the exchange-correlation kernel (0 in the case of HF), and $c_{HF}$ the percentage of Hartree-Fock exchange.[64,66–68] Practically, it is generally too cost-intensive to form the entire matrix $\underline{\Omega}$ and invert it to solve for $\underline{X}$ and $\underline{Y}$. Instead, we can iteratively form the inverse using the direct inversion in the iterative subspace (DIIS) procedure.[68,69] We can write Eq. 2.23 more compactly as $\underline{\Omega}\tilde{\underline{X}} = \tilde{\underline{R}}$. Inverting and expressing $\underline{\Omega}^{-1}$ as a series expansion

$$\tilde{\underline{X}} = (\underline{1} + (\underline{1} - \underline{\Omega}) + (\underline{1} - \underline{\Omega})^2 + \ldots)\tilde{\underline{R}} \tag{2.28}$$

This suggests that by forming affordable matrix-vector products $\tilde{\underline{X}}_n = (\underline{1} - \underline{\Omega})\tilde{\underline{X}}_{n-1}$ with $\tilde{\underline{X}}_0 = \tilde{\underline{R}}$ and

orthogonalizing $\underline{\tilde{X}}_n$ to the set of previous vectors $\{\underline{\tilde{X}}_0, \underline{\tilde{X}}_1, ..., \underline{\tilde{X}}_{n-1}\}$ until the norm of $\underline{\tilde{X}}_n$ converges to zero, we can obtain $\underline{\tilde{X}}$ via a rapidly convergent sum

$$\underline{\tilde{X}} = \sum_{i=0}^{n-1} \alpha_i \underline{\tilde{X}}_i \tag{2.29}$$

where the $\alpha_i$ can be determined from a simple $n \times n$ system of equations.

While we have only described how to evaluate the response functions for SCF methods, response theory is also used for coupled cluster calculations, which we perform in Chapter 6.[70,71] The derivation of coupled cluster linear response equations is described in more detail elsewhere.[72,73]

## 2.3 Periodic Electronic Structure Theory

In order to practically model crystalline systems, we must efficiently incorporate information about periodicity into electronic structure methods. We can represent a crystalline system in real space as a unit cell repeated along as many as three periodic axes, which we can collect as columns of a matrix $\boldsymbol{D}$. We can assign a set of atomic basis functions $\{\chi_\nu | \nu = 1, ..., \Omega\}$ to each cell and label the basis functions of different cells with a vector $\boldsymbol{l}$, denoting the number of increments along each periodic axis up to $\boldsymbol{L} = (L_1, L_2, L_3)$. We can obtain a more convenient representation by converting to so called crystal orbitals

$$|q\boldsymbol{k}\rangle = \sum_{l_1=-L_1}^{L_1} \sum_{l_2=-L_2}^{L_2} \sum_{l_3=-L_3}^{L_3} e^{i\boldsymbol{k}\cdot(\boldsymbol{D}\boldsymbol{l})} \sum_{\nu=1}^{\Omega} C_{q\nu}(\boldsymbol{k}) \chi_\nu^{\boldsymbol{l}}(\boldsymbol{r}) \tag{2.30}$$

where $|q\boldsymbol{k}\rangle$ is the $q^{\text{th}}$ Bloch function at the point $\boldsymbol{k}$ in reciprocal space and $\underline{C}$ is the matrix of molecular orbital coefficients for that point.[68,74] To simplify the notation, we will write the triple sum over cell indices as a single sum over the multi-index $\boldsymbol{l}$

$$\sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}} = \sum_{l_1=-L_1}^{L_1} \sum_{l_2=-L_2}^{L_2} \sum_{l_3=-L_3}^{L_3} \tag{2.31}$$

Working in reciprocal space allows us to more naturally treat the periodicity of the system.

Instead of solving the SCF equation over a large supercell in real-space, we can instead separately solve much smaller SCF equations for each k-point[68,75,76]

$$\underline{F}(\boldsymbol{k})\underline{C}(\boldsymbol{k}) = \underline{S}(\boldsymbol{k})\underline{C}(\boldsymbol{k})\underline{\varepsilon}(\boldsymbol{k}) \tag{2.32}$$

where the Fock matrix $\underline{F}$ and overlap matrix $\underline{S}$ in reciprocal space are obtained via Fourier transforms

$$F_{\mu\nu}(\boldsymbol{k}) = \sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}} e^{i\boldsymbol{k}\cdot(\boldsymbol{D}\boldsymbol{l})} F_{\mu\nu}^{\boldsymbol{0}\boldsymbol{l}} \tag{2.33}$$

$$S_{\mu\nu}(\boldsymbol{k}) = \sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}} e^{i\boldsymbol{k}\cdot(\boldsymbol{D}\boldsymbol{l})} S_{\mu\nu}^{\boldsymbol{0}\boldsymbol{l}} \tag{2.34}$$

A complication arises when trying to model properties that depend on the position operator $\boldsymbol{r}$, specifically the electric dipole and electric dipole-electric dipole polarizability, as this operator is not translation invariant. For weakly conducting systems under the influence of weak electric fields, we can represent $\boldsymbol{r}$ in a translation invariant form[77]

$$\boldsymbol{r} \simeq i\exp(i\boldsymbol{k}\cdot\boldsymbol{r})\nabla_{\boldsymbol{k}}\exp(-i\boldsymbol{k}\cdot\boldsymbol{r}) \tag{2.35}$$

The matrix elements of the electric dipole using this transformed operator take the form

$$\begin{aligned}\hat{\boldsymbol{\mu}}_{p\boldsymbol{k},q\boldsymbol{k}} &= i\langle p\boldsymbol{k}|e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\nabla_{\boldsymbol{k}}e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}|q\boldsymbol{k}\rangle \\ &= i\langle u_{p\boldsymbol{k}}|\nabla_{\boldsymbol{k}}|u_{q\boldsymbol{k}}\rangle\end{aligned} \tag{2.36}$$

where

$$u_{q\boldsymbol{k}}(\boldsymbol{r}) = e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}|q\boldsymbol{k}\rangle \tag{2.37}$$

is the cell periodic part of the Bloch function and $\hat{\underline{\boldsymbol{\mu}}}$ is the periodic dipole matrix. To express this periodic matrix in terms of atomic orbital matrix elements so it can be evaluated, we start by

applying the operator in Eq. 2.36 to the ket to obtain

$$i\nabla_{\boldsymbol{k}}|u_{q\boldsymbol{k}}\rangle = \Bigg[\boldsymbol{r}\sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}} e^{i\boldsymbol{k}\cdot(\boldsymbol{D}\boldsymbol{l}-\boldsymbol{r})}\sum_{\nu=1}^{\Omega} C_{q\nu}(\boldsymbol{k})\chi_{\nu}^{\boldsymbol{l}}(\boldsymbol{r}) \tag{2.38a}$$

$$-\sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}}(\boldsymbol{D}\boldsymbol{l})e^{i\boldsymbol{k}\cdot(\boldsymbol{D}\boldsymbol{l}-\boldsymbol{r})}\sum_{\nu=1}^{\Omega} C_{q\nu}(\boldsymbol{k})\chi_{\nu}^{\boldsymbol{l}}(\boldsymbol{r}) \tag{2.38b}$$

$$i\sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}} e^{i\boldsymbol{k}\cdot(\boldsymbol{D}\boldsymbol{l}-\boldsymbol{r})}\sum_{\nu=1}^{\Omega}\frac{\partial C_{q\nu}(\boldsymbol{k})}{\partial \boldsymbol{k}}\chi_{\nu}^{\boldsymbol{l}}(\boldsymbol{r})\Bigg] \tag{2.38c}$$

We can now express the periodic electric dipole as

$$\hat{\underline{\boldsymbol{\mu}}} = \Bigg[\underline{C}^{\dagger}\underline{\boldsymbol{\mu}}\underline{C} \tag{2.39a}$$

$$+i\underline{C}^{\dagger}(\nabla_{\boldsymbol{k}}\underline{S})\underline{C} \tag{2.39b}$$

$$+i\underline{C}^{\dagger}\nabla_{\boldsymbol{k}}\underline{C}\Bigg] \tag{2.39c}$$

where Eq. 2.39a-c result from applying the bra to the terms Eq. 2.38a-c respectively. To explicitly evaluate the term 2.39c, we define $\nabla_{\boldsymbol{k}}\underline{C} = \underline{C}\underline{\boldsymbol{U}}$, i.e. $\underline{\boldsymbol{U}}$ is the matrix that transforms the molecular orbital coefficients to their gradient with respect to $\boldsymbol{k}$. We can obtain an explicit form for $\underline{\boldsymbol{U}}$ by applying $\nabla_{\boldsymbol{k}}$ to the SCF equation (Eq. 2.32) and the orthonormality condition for the molecular orbital coefficients,[66,68] which results in

$$\begin{aligned}\boldsymbol{U}_{pp} &= -\frac{1}{2}\nabla_{\boldsymbol{k}}S_{pp}\\ \boldsymbol{U}_{pq} &= \frac{(\nabla_{\boldsymbol{k}}F)_{pq}-(\nabla_{\boldsymbol{k}}S)_{pq}\varepsilon_q}{\varepsilon_q-\varepsilon_p}\end{aligned} \tag{2.40}$$

An undesirable trait of Eq. 2.39 is that neither $\underline{\boldsymbol{\mu}}$ nor $i\underline{\boldsymbol{U}}$ are Hermitian, as this would allow us to only store symmetry inequivalent matrix elements rather than the whole matrices. Inspecting a matrix element of $\underline{\boldsymbol{\mu}}$, we can see that

$$\begin{aligned}
\boldsymbol{\mu}_{\mu\nu}^{\mathbf{0}\boldsymbol{l}} &= \langle \mu_{\mathbf{0}} | \boldsymbol{r} e^{-i\boldsymbol{D}\boldsymbol{l}\cdot\boldsymbol{p}} | \nu_{\mathbf{0}} \rangle \\
&= \langle \mu_{-\boldsymbol{l}} | \boldsymbol{r} | \nu_{\mathbf{0}} \rangle + \langle \mu_{\mathbf{0}} | [\boldsymbol{r}, e^{-i\boldsymbol{D}\boldsymbol{l}\cdot\boldsymbol{p}}] | \nu_{\mathbf{0}} \rangle \\
&= \langle \mu_{-\boldsymbol{l}} | \boldsymbol{r} | \nu_{\mathbf{0}} \rangle + \boldsymbol{D}\boldsymbol{l} \langle \mu_{-\boldsymbol{l}} | \nu_{\mathbf{0}} \rangle \\
&= \boldsymbol{\mu}_{\nu\mu}^{\mathbf{0}-\boldsymbol{l}} + \boldsymbol{D}\boldsymbol{l} S_{\nu\mu}^{\mathbf{0}-\boldsymbol{l}}
\end{aligned} \tag{2.41}$$

The second step is the result of $\boldsymbol{r} e^{-i\boldsymbol{D}\boldsymbol{l}\cdot\boldsymbol{p}} = e^{-i\boldsymbol{D}\boldsymbol{l}\cdot\boldsymbol{p}} \boldsymbol{r} + [\boldsymbol{r}, e^{-i\boldsymbol{D}\boldsymbol{l}\cdot\boldsymbol{p}}]$ where $[\cdot,\cdot]$ is the commutator. The third step evaluates the commutator in the second term using the relationship $[\boldsymbol{r}, f(\boldsymbol{p})] = i f'(\boldsymbol{p})$, which holds for any function of momentum $f(\boldsymbol{p})$ expressed in atomic units.[78] We can form a Hermitian matrix by averaging each matrix element with its corresponding adjoint element

$$\tilde{\boldsymbol{\mu}}_{\mu\nu}^{\mathbf{0}\boldsymbol{l}} = \frac{1}{2}\left(\boldsymbol{\mu}_{\mu\nu}^{\mathbf{0}\boldsymbol{l}} + \boldsymbol{\mu}_{\nu\mu}^{\mathbf{0}-\boldsymbol{l}}\right) = \boldsymbol{\mu}_{\mu\nu}^{\mathbf{0}\boldsymbol{l}} + \frac{1}{2}\boldsymbol{D}\boldsymbol{l} S_{\nu\mu}^{\mathbf{0}-\boldsymbol{l}} = (\tilde{\boldsymbol{\mu}}_{\nu\mu}^{\mathbf{0}-\boldsymbol{l}})^* \tag{2.42}$$

Note that

$$i\nabla_{\boldsymbol{k}} S_{\mu\nu} = \nabla_{\boldsymbol{k}} \sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}} e^{i\boldsymbol{k}\cdot\boldsymbol{D}\boldsymbol{l}} S_{\mu\nu}^{\mathbf{0}\boldsymbol{l}} = \sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}} e^{i\boldsymbol{k}\cdot\boldsymbol{D}\boldsymbol{l}} (\boldsymbol{D}\boldsymbol{l}) S_{\mu\nu}^{\mathbf{0}\boldsymbol{l}} \tag{2.43}$$

so we can combine the term 2.39a and half of the term 2.39b to form $\underline{\tilde{\boldsymbol{\mu}}}$. We can then add the remaining half of the term 2.39b to the term 2.39c to form the Hermitian matrix $i\underline{\tilde{\boldsymbol{U}}}$. Altogether, this allows us to rewrite Eq. 2.39 more simply in terms of these Hermitian matrices

$$\hat{\underline{\boldsymbol{\mu}}} = \underline{C}^{\dagger} \underline{\tilde{\boldsymbol{\mu}}} \underline{C} + i\underline{\tilde{\boldsymbol{U}}} \tag{2.44}$$

Using this periodic form of the dipole operator, it is possible to compute a periodic electric dipole polarizability using the linear response techniques described in section 2.2.[68] Evaluating other periodic polarizabilities, such as $\boldsymbol{\beta}$ and $\mathbf{A}$, requires the additional step of converting the magnetic dipole and electric quadrupole to a translation invariant form. This process is described in more detail in Chapter 7.

# Part I

# Optical Rotation Analysis Methods

# Chapter 3

# Gauge Dependence of the $\tilde{S}$ Molecular Orbital Space Decomposition of Optical Rotation

(This work taken with the permission of Ty Balduf and Marco Caricato from *J. Phys. Chem. A* **2021**, *125*, 4976-4985.[79] Supporting information is available online.)

## 3.1 Introduction

Detecting and controlling molecular and supra-molecular chirality is an active area of research with applications ranging from pharmaceutical development[2] and organic chiral catalysis[7,8,80] to nanostructure assembly[81,82] and chiral light generation.[83,84] The measurement of optical rotation (OR), the change in orientation of plane polarized light impinging on a chiral sample, is a foundational technique for absolute configuration assignment and chiral sensing.[6,26,85,86] Despite the long standing use of OR measurements, the relationship between the structural features of a compound and its observed OR remains unclear.

In recent years, a number of theoretical schemes have been developed to decompose the OR into electronic contributions in an effort to understand how the molecular shape affects the OR. Wiberg et al. have used the sum-over-states formulation to describe OR contributions in terms of movement of charge density in excited state transitions.[31] Polavarapu et al. used modern electronic structure methods to test the semiempirical Kirkwood model, which expresses the OR as a sum of bond polarizabilities and anisotropies.[33,34] Beratan et al. developed a Mulliken-like partitioning to assign OR contributions to individual atoms and functional groups of a molecule.[35] Autschbach et al. studied how OR could be split into molecular orbital (MO) contributions.[38]

We previously proposed a similar method to Autschbach et al. where instead of assigning contributions to individual MOs, we instead decompose the OR into contributions from occupied to virtual MO transitions.[39] We refer to these contributions as $\tilde{S}_{ia}$ values, where $i/a$ is the index of an occupied/virtual MO pair. While a molecule can have many possible transitions, the OR can often be described qualitatively using just the transitions with the largest $\tilde{S}_{ia}$ values. In this way, the induced OR can be ascribed to the movement of electronic density during these transitions. These shifts in the density can be tied to particular functional groups of the molecule, offering chemical intuition that can be applied more generally to other systems. We have used this approach to explain differences in OR between molecular conformers,[87] as well to determine the effect of functionalization.[40]

In these prior studies, $\tilde{S}_{ia}$ was expressed in the length gauge (LG), which is a commonly used form for OR calculations. In principle, the choice of gauge should not affect the OR, but this is only true for variational electronic structure methods using an infinite basis; thus, practical calculations of OR exhibit gauge dependence.[72] Even ignoring these differences in the total OR, there is no guarantee that $\tilde{S}_{ia}$ contributions and their corresponding physical interpretations remain the same if a different gauge is used to define $\tilde{S}$. To address this ambiguity, here we develop two new definitions of $\tilde{S}_{ia}$ with the modified velocity gauge (MVG) and apply them to a subset of the molecules from these prior studies. By comparing among these different definitions, we can determine if they produce consistent physical interpretations of the OR.

This chapter is organized as follows. Section 3.2 gives a brief description of the various definitions that can be used to calculate optical rotation and its MO space decomposition, $\tilde{S}$. Section 3.3 defines $\hat{S}$ values, a normalized version of $\tilde{S}$, and describes the model chemistry used in our electronic structure calculations. In Section 3.4, we compute $\hat{S}$ values for a small test set of molecules and analyze how the values change with the choice of gauge/perturbation. We conclude with a discussion of the use of $\tilde{S}$ as an interpretive tool and whether insights from this method are dependent on the choice of definition in Section 3.5.

## 3.2 Theory

The specific rotation $[\alpha]_\omega$ (deg dm$^{-1}$ (g/mL)$^{-1}$) induced by light of frequency $\omega$ impinging on an isotropically dispersed chiral sample can be calculated as shown in Eq. 2.6, where $\boldsymbol{\beta}$ is defined by the sum over states expression in Eq. 2.2. Eq. 2.2 is not used in practice, as the sum over excited states is slow to converge.[63] Instead, $\boldsymbol{\beta}$ is typically evaluated via linear response (LR) theory, which for self consistent field (SCF) methods takes the form:[28,64,88]

$$\begin{aligned}\beta_{\alpha\beta} &= \frac{1}{c_\omega}\sum_{i}^{N_{\text{occ}}}\sum_{a}^{N_{\text{virt}}}\langle\phi_i|\mu_\alpha|\phi_a\rangle\langle\phi_a|X^{+}_{m_\beta}|\phi_i\rangle \\ &= \frac{1}{c_\omega}\sum_{i}^{N_{\text{occ}}}\sum_{a}^{N_{\text{virt}}}\langle\phi_i|X^{-}_{\mu_\alpha}|\phi_a\rangle\langle\phi_a|m_\beta|\phi_i\rangle\end{aligned} \tag{3.1}$$

where $\langle\phi_a|X^{\pm}_{y_\alpha}|\phi_i\rangle$ represents the density perturbed by the $y$ field (either electric or magnetic) expressed in the MO basis, and $c_\omega = \omega$ for LG or $c_\omega = \omega^2$ for MVG (see below). Typically, the expressions in Eq. 3.1 are actually calculated in atomic orbital basis (AO), but the MO basis is more convenient here for clarity's sake. The perturbed density is computed by solving Eqs. 2.23 and 2.29. This changes the problem from determining a large number of excited states to finding a self-consistent response of the density to the applied perturbation. Either type of perturbed density can be determined iteratively by solving the LR-SCF equations starting from the ground state wavefunction; as shown in Eq. 3.1, the electric and magnetic perturbations are equivalent and will result in the same values for $\boldsymbol{\beta}$.

Electric dipole matrix elements in principle satisfy the hypervirial relationship (in a.u.):[89]

$$\langle\phi_a|\mu_\alpha|\phi_i\rangle = \frac{i}{\omega_{ia}}\langle\phi_a|p_\alpha|\phi_i\rangle \tag{3.2}$$

where $\boldsymbol{p}$ is the momentum operator. However, this relationship does not hold for approximate electronic structure methods with finite basis sets and $\boldsymbol{\beta}$ will differ depending on the gauge used to model the perturbation.[58] Optical rotation calculations are generally done in the length gauge (LG), which uses the form of the dipoles on the left-hand side of Eq. 3.2, or the modified velocity gauge

(MVG), which uses the form on the right-hand side. Length gauge calculations of OR are not origin invariant, which is typically corrected by using gauge-including atomic orbitals (GIAOs) for SCF methods.[61] However, using GIAOs introduces extra terms into the expression for $\boldsymbol{\beta}$ in Eqs. 3.1-2.23. These terms can be incorporated into the magnetic perturbed density via the LR equations in Eq. 2.23, while they remain separate for the electric perturbed density, thus breaking the symmetry between the two sides of Eq. 3.1.[90] On the other hand, the OR in the velocity gauge is origin independent for any electronic structure method, but it contains a spurious static field contribution that needs to be evaluated and subtracted out explicitly (hence the name "modified" velocity gauge).[60]

Recently, we developed a configuration space analysis of the $\boldsymbol{\beta}$ tensor as a tool to determine the underlying electronic processes that induce optical rotation in chiral molecules.[39,87] We defined a rotatory strength in molecular orbital space, $\tilde{S}_{ia}$, resulting from an occupied to virtual ($i \rightarrow a$) MO transition, which can be written in the length gauge as:

$$\tilde{S}_{ia}^{LG-M} = \mathrm{Im}\left[\langle i|\boldsymbol{\mu}|a\rangle \cdot \langle a|\boldsymbol{X}_{\boldsymbol{m}}^{+}|i\rangle\right] \tag{3.3}$$

such that:

$$\sum_{ia} \tilde{S}_{ia}^{LG-M} = \omega \mathrm{Tr}(\boldsymbol{\beta}^{LG}) \tag{3.4}$$

The insight behind this approach is that it allows the optical rotation to be expressed as a sum of contributions from individual transitions: By determining which transitions make a large contribution to the OR, one can make qualitative predictions about how particular changes to the molecular geometry or the electronic density will affect the total OR. The $\tilde{S}$ values in the LG are defined in terms of the magnetic perturbed density in Eq. 3.3 because GIAOs will introduce additional terms for the electric perturbation, as was the case for $\boldsymbol{\beta}$. Incorporating these terms into an $\tilde{S}_{ia}$ definition is nontrivial, so we forgo consideration of the LG-electric perturbed density for the remainder of this paper.

In this work, we propose two new definitions of the $\tilde{S}$ rotatory strength using the modified

velocity gauge:

$$\tilde{S}_{ia}^{MVG-M} = \frac{1}{\omega}\mathrm{Re}\left[\langle i|\boldsymbol{p}|a\rangle \cdot \langle a|\boldsymbol{X}_{\boldsymbol{m}}^{+}|i\rangle\right] \tag{3.5a}$$

$$\tilde{S}_{ia}^{MVG-E} = \frac{1}{\omega}\mathrm{Re}\left[\langle i|\boldsymbol{X}_{\boldsymbol{p}}^{-}|a\rangle \cdot \langle a|\boldsymbol{m}|i\rangle\right] \tag{3.5b}$$

such that:

$$\sum_{ia}\tilde{S}_{ia}^{MVG-M} = \sum_{ia}\tilde{S}_{ia}^{MVG-E} = \omega\mathrm{Tr}(\boldsymbol{\beta}^{MVG}) \tag{3.6}$$

The unphysical static term is already subtracted out in the expressions in Eq. 3.5. These two definitions provide the same MVG OR once all contributions are included, but each individual value is different. We can define two separate expressions for $\tilde{S}$ with this choice of gauge because there is no complication due to GIAOs. As for the original $\tilde{S}$ definition in Eq. 3.3, the computational cost to evaluate all $\tilde{S}_{ia}$ values in Eq. 3.5 is negligible compared to that of the SCF and LR equations. In fact, evaluating $\tilde{S}$ costs $6N^3$ for the AO→MO transformation of the Cartesian components of the dipole integrals and of the perturbed density matrices.

The main goal of this work is to test whether these three $\tilde{S}$ definitions (LG-M in Eq. 3.3, and MVG-M/MVG-E in Eq. 3.5) provide a similar qualitative picture of the OR for chiral molecules in terms of constituent one-electron transitions or whether one definition is preferable to the others. We also test the quantitative convergence of each definition, as we have previously shown that small fractions of the total number of transitions can be used to approximate the specific rotation. Demonstrating this for other definitions of $\tilde{S}$ would aid in the development of reduced cost methods to compute OR.[91]

## 3.3 Computational Procedure

To analyze how the choice of gauge and perturbed density affects the $\tilde{S}$ values, we compiled a test set of molecules from previous studies based on the length gauge with the magnetic-perturbed density (LG-M) definition of $\tilde{S}$ in Eq. 3.3, depicted in Figure 3.1. The optimized geometries were

taken as-is from these prior studies (or references therein) and are recorded here in Tables S1-S9 of the Supporting Information (SI).[30,31,40,92] For each of these molecules, we calculated the total optical rotation and $\tilde{S}$ using the MVG-M and MVG-E definitions in Eq. 3.5 and compared them with previous LG-M results. Calculations were performed with the B3LYP/aug-cc-pVDZ model chemistry[93,94] with a 589.3 nm perturbation wavelength in a development version of the GAUSSIAN suite of programs.[95] Note that while this model chemistry has been shown to be reasonable for calculating OR, we make no effort here to assess its accuracy. Rather, we seek to compare the physical interpretation of $\tilde{S}$ with each perturbed density, regardless of how accurately the OR is described using this method.

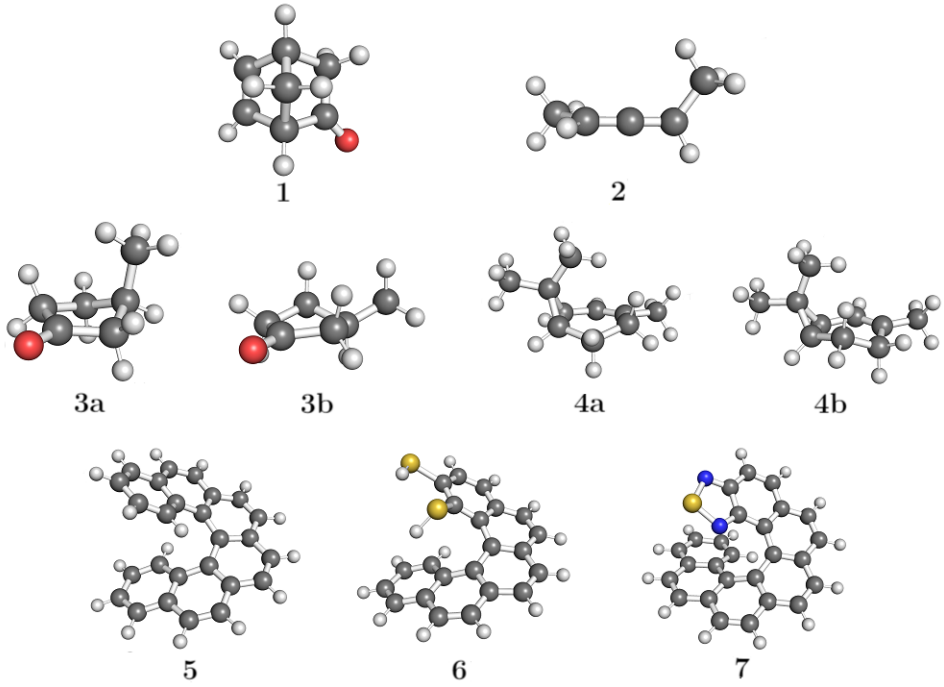


Figure 3.1: Structure of the molecules in the test set: (**1**) (1S,4S)-norbornenone, (**2**) P-(2,3)-pentadiene, (**3a**,**3b**) axial and equatorial (R)-(+)-3-methycyclopentanone, (**4a**,**4b**) A and B conformers of (S)-(+)-2-carene, (**5**) [6]helicene, (**6**) dithiol[5]helicene, (**7**) benzothiadiazole[6]helicene. Atoms are colored as follows: C (gray), H (white), O (red), N (blue), S (yellow).

The overall OR does not change much for different perturbations, see Table S10 in the SI. However, to ensure that differences in the total OR are not influencing comparisons of $\tilde{S}$, we define

a normalized configuration rotatory strength, $\hat{S}$:

$$\hat{S}_{ia} = \frac{\tilde{S}_{ia}}{\sum_{jb} \tilde{S}_{jb}} = \frac{\tilde{S}_{ia}}{\omega \mathrm{Tr}(\boldsymbol{\beta})} \tag{3.7}$$

where the second equality comes from Eqs. 3.4 or 3.6. Since by definition $\sum_{ia} \hat{S}_{ia} = 1$ for each molecule and perturbation type, each $\hat{S}_{ia}$ value gives the relative contribution of that transition to the total OR. The sign of $\hat{S}_{ia}$ is positive if $\tilde{S}_{ia}$ has the same sign as $\mathrm{Tr}(\boldsymbol{\beta})$ and negative if $\tilde{S}_{ia}$ has the opposite sign.

## 3.4 Results

In this section, we compare $\hat{S}$ values computed with the various perturbed densities for the molecules in Figure 3.1. To make a comprehensive comparison, we analyze the $\hat{S}$ values of each molecule with increasing level of detail. In Section 3.4.1, we compare the cumulative contribution to the OR of different sets of $\hat{S}$ values, separated according to their magnitude. In Section 3.4.2, we use heat maps of $\hat{S}_{ia}$ to visualize how these contributions are distributed among the frontier orbital transitions. From these transitions, we plot the perturbation vectors for the largest contributors in Section 3.4.3 to determine whether these transitions are consistent across the choice of gauge and perturbed density.

Before delving into these comparisons, we briefly summarize the prior $\tilde{S}$ studies of these molecules. For molecules **1** and **2**, it has been shown that the OR can be qualitatively described by a small fraction of the total number of transitions, coming within an order of magnitude of the total OR with only 3 out of 6583 and 4 out of 3192 $\tilde{S}_{ia}$ values, respectively.[39] The $\tilde{S}_{ia}$ value for the highest occupied molecular orbital (HOMO) to lowest unoccupied molecular orbital (LUMO) transition of molecule **1** is considerably larger than the other transitions and has the same sign as the overall $\mathrm{Tr}(\boldsymbol{\beta})$. This transition involves partial charge transfer between the two double bond groups, and the favorable interaction of the magnetic and electric transition vectors is possible because of the rigid cage structure of the molecule. This large $\tilde{S}_{ia}$ element helps to explain the

large OR observed for this molecule. For molecule **2**, the four largest magnitude transitions are comparable in size to the largest transitions of molecule **1**, and also correspond to charge transfer excitations between the double bond groups. However, in this case the $\tilde{S}_{ia}$ values come in opposite signed pairs because of the molecule's symmetry, so their combined contribution is small. This near cancellation provides an explanation for why the observed OR for molecule **2** ($[\alpha]_{355}^{\text{exp}}$ = +408 deg dm$^{-1}$ (g/mL)$^{-1}$) is significantly smaller than that of molecule **1** ($[\alpha]_{355}^{\text{exp}}$ = −6310 deg dm$^{-1}$ (g/mL)$^{-1}$).

For molecules **3** and **4**, $\tilde{S}$ analysis was used to determine the origin of OR differences between conformers.[87] The conformers of **3** each had their largest contribution from the (HOMO→LUMO) transition, but these contributions were opposite in sign, as was the total OR for each conformer. This change in sign was attributed to the electric vector being nearly parallel with the magnetic vector in the axial conformer (**3a**), but nearly antiparallel in the equatorial conformer (**3b**) due to the change in orientation of the methyl. The difference in OR between the conformers of **4** mainly stemmed from differences between HOMO transitions, specifically the (HOMO→LUMO+2) transition. The electric and magnetic vectors for this transition had an angle less than 90° for conformer **4a**, but greater than 90° for conformer **4b**, resulting in a different sign for the $\tilde{S}_{ia}$ of this transition. For each set of conformers, including transitions within a factor of $10^{-3}$ of the total sum (30-40% of the transitions) captures 97% of the total OR.

Molecules **5**-**7** are helicenes, which exhibit a strong chiroptical response due to their axial chirality. $\tilde{S}$ was applied to these molecules and other functionalized helicenes to determine the effect of length and withdrawing/donating character of functional groups on the observed OR.[40] While these molecules have many $\tilde{S}$ contributions, the dominant transitions were characterized by magnetic vectors pointed parallel to the helical axis, which corresponds to electron density moving along the helix body. The magnitude of the magnetic vector for these types of transitions was found to increase with the length of the helix and with delocalization of charge density. Also of note was the number of transitions needed for a quantitative description of the OR. The helicenes have on the order of $10^5$ transitions in the aug-cc-pVDZ basis set, but a large portion of the total OR can be

recovered with only $10^{3-4}$ transitions.[40] If these significant transitions could be determined a priori, one could greatly reduce the cost of an OR calculation. Aharon and Caricato proposed truncating the MO space based on the $\tilde{S}_{ia}$ values computed using the initial guess perturbed density.[91] Using this criterion for a set of 51 organic molecules (including molecules **1**-**5**), they estimated a speedup of 2-8x depending on the size of the basis, with a mean unsigned error of < 1% for the computed OR.

### 3.4.1 Cumulative Optical Rotation Contributions

In Figure 3.2, we plot the cumulative $\hat{S}$ for each molecule. On the left hand side of each plot, only the largest contributors are included. Moving to the right, progressively smaller contributions are added until all contributions are included and $\sum \hat{S}_{ia} = 1$ for each perturbed density. Scanning across each plot, we can see that the contributions to the OR for each $\hat{S}$ definition become more similar as more transitions are included. The only notable dissimilarity is found in the first set of transitions, where the contribution for the MVG-E definition is half as large as the others. Using molecule **1** as an example, all $\hat{S}$ definitions require the same number of relevant transitions (2 out of 6583) to provide a qualitative understanding of the OR for this molecule.[87] Once $\hat{S}_{ia}$ values larger in magnitude than $10^{-2}$ are included in the sum, each $\hat{S}$ definition is no farther than 0.09 from the others. Not only are all the definitions similar at this point in the summation, but they are also within 0.07 of their total sums. Therefore, these results suggest that this MO decomposition of the OR may not be very sensitive to the choice of $\hat{S}$ definition.

Molecules **2**-**7** also show similarity between the $\hat{S}$ definitions, as depicted in Figure 3.2. There is some significant variation when only large values are included in the sum, as can be seen for molecule **2**. For the largest magnitude transitions, $|\hat{S}_{ia}| \geq 1$, the LG-M and MVG-M definitions actually have opposite signs compared to their total sums, while the MVG-E definition has the same sign as its total. Note that the four largest transitions do give the correct sign for all the definitions of $\tilde{S}$, consistent with our prior study, but here we have chosen a smaller cutoff for significant contributions and these additional transitions change the sign for the LG-M and MVG-

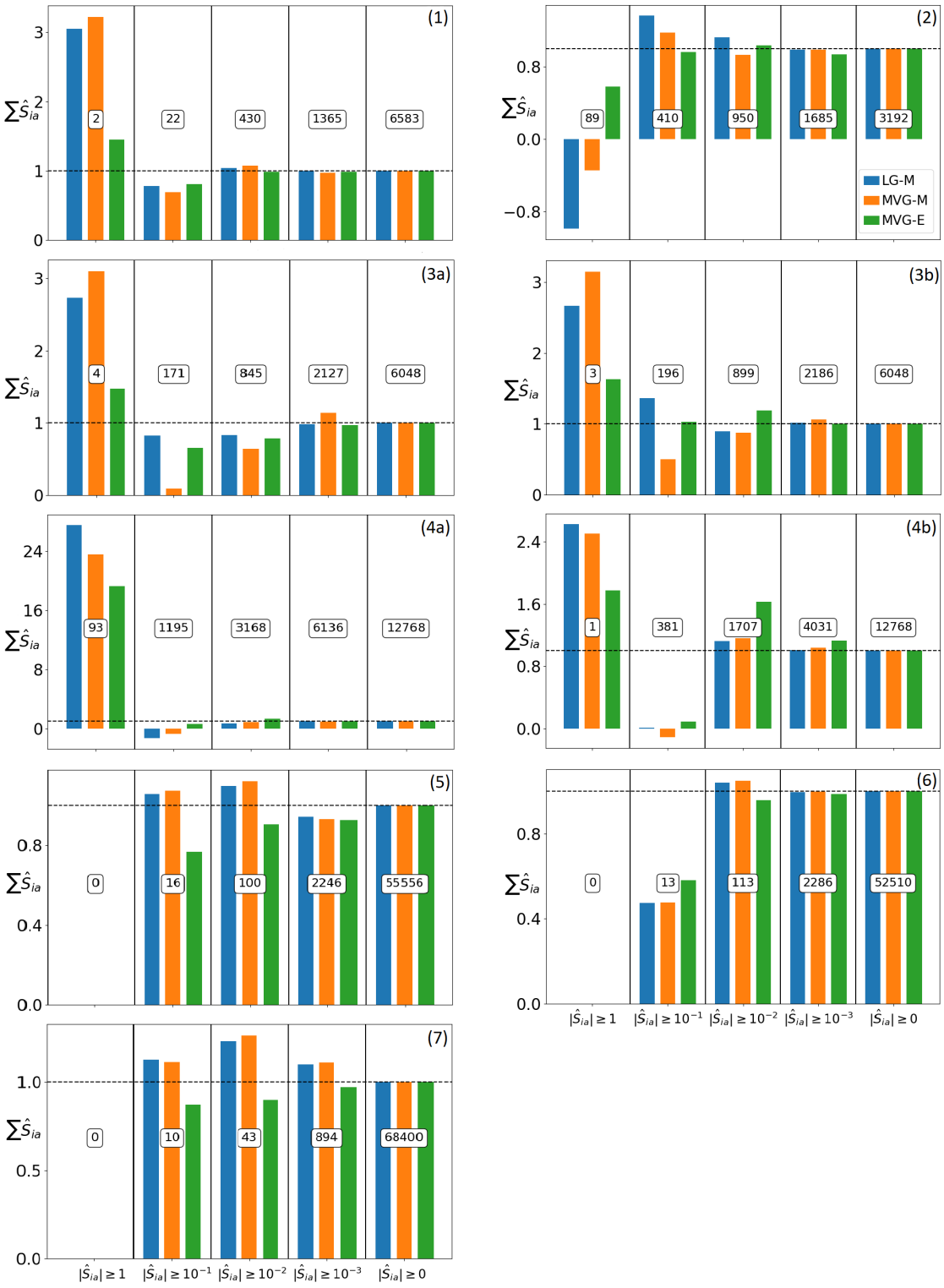


Figure 3.2: Cumulative $\hat{S}_{ia}$ for molecules **1**-**7** computed with LG-M, MVG-M, and MVG-E definitions of $\tilde{S}$. The transitions selected are those for which the LG-M $|\hat{S}_{ia}|$ is greater than the specified value. The height of each bar is the sum of these transitions, while the number of transitions is listed above the bars. The total sum is marked with a dashed line.

M definitions. For these definitions, there are 7 transitions with $1.0 > |\hat{S}_{ia}| \geq 0.95$, so the choice of cutoff can greatly influence the value and even the sign of this partial sum. While in this paper we choose consistent cutoffs in order to facilitate comparisons across molecules and $\tilde{S}$ definitions, in practice, the cutoffs used to define significant transitions should be tailored to each molecule. This is clearly exemplified in the case of molecule **2**: the $|\hat{S}_{ia}|$ values for the 4 largest transitions are around an order of magnitude larger than any of the other transitions, which suggests that they should considered separately from the other $|\hat{S}_{ia}| \geq 1$ transitions.

For the molecules tested, the different $\tilde{S}$ definitions are generally indistinguishable once contributions with $|\hat{S}_{ia}| \geq 10^{-3}$ are included. Even with just $\hat{S}_{ia}$ contributions larger than $10^{-2}$, the different definitions are qualitatively consistent. However, including just contributions with $|\hat{S}_{ia}| \geq 10^{-1}$ is sometimes insufficient for even qualitative agreement; for instance the MVG-M partial sum for **4b** is the wrong sign. As previously mentioned, these qualitative inconsistencies can be remedied using a tailored selection of transitions in the partial sum, rather than a preselected cutoff.

It is noteworthy that where there are differences between the $\hat{S}$ definitions, the LG-M and MVG-M choices tend to be very similar while MVG-E differs. This is somewhat surprising since the MVG-M and MVG-E definitions are based on the same choice of gauge and produce identical overall OR values. In the next section, we explore whether this trend extends to the level of individual transitions.

### 3.4.2 Distribution of Contributions

We have seen up to this point that by summing the same set of transitions, we can obtain a similar proportion of the total OR for each $\hat{S}$ definition. However, the cumulative summation does not show how much individual transitions contribute. To give a more detailed representation of how the OR is distributed among the transitions, we present heat maps of $\hat{S}_{ia}$ in Figures 3.3 and 3.4. As most of the sizable transitions are between frontier orbitals, the heat maps only display transitions among the 15 highest occupied and 15 lowest virtual orbitals. Full maps are reported in Figures S1-S9 of the SI.

Focusing first on **1**, the first row of Figure 3.3 plots its frontier transitions using the LG-M, MVG-M, and MVG-E definitions. The largest two $|\hat{S}_{ia}|$ values for each definition correspond to the (HOMO→LUMO) and (HOMO-1→LUMO) transitions, which are the same $|\hat{S}_{ia}| \geq 1$ transitions plotted in Figure 3.2. However, while these transitions have similar $\hat{S}_{ia}$ for LG-M and MVG-M, they are much smaller for the MVG-E definition (e.g. $\hat{S}_{29,1}$=4.65, 5.12, and 1.84 for the LG-M, MVG-M, and MVG-E definitions respectively). This pattern seems to hold across frontier transitions of **1**, with a consistent distribution of $\hat{S}_{ia}$ for each choice of perturbed density, but with smaller magnitudes with the MVG-E choice. For the rest of the test set, the distribution of $\hat{S}$ among the frontier transitions remains largely unchanged with different $\hat{S}$ definitions. While some of the molecules, such as **2** and **3a**, exhibit consistently smaller $\hat{S}$ values with MVG-E than with the other definitions, this is not the case in general. For example, $\hat{S}_{ia}$ values for **4a** are similar in magnitude for each definition and the MVG-E choice for **5** actually has some larger transitions than the other definitions, e.g. (HOMO-1→LUMO).

The similarity in the distribution of values for each definition would suggest that $\tilde{S}$ is largely invariant to the choice of gauge and perturbed density. However, the differences in magnitude seen for MVG-E could be a sign that it is producing different representations of the transitions when compared to the magnetic perturbed density definitions.

### 3.4.3 Major Optical Rotation Contributions

To address whether $\tilde{S}_{ia}$ has the same physical interpretation with different definitions, we superimpose the electric and magnetic dipole/perturbed density vectors of the largest transitions onto the molecules using the PyMOL program.[96] These vectors describe how the electronic density rearranges itself to induce OR; similar vectors for a given occupied-virtual transition indicate a similar underlying physical process. Note that we are considering the unnormalized $\tilde{S}$ in this section; this allows us to account for the slight differences in total OR when comparing between the length gauge and modified velocity gauge. The magnitude of the vectors for each transition and the angles between them are reported in Table S11 of the SI. For the MVG definitions, each vector

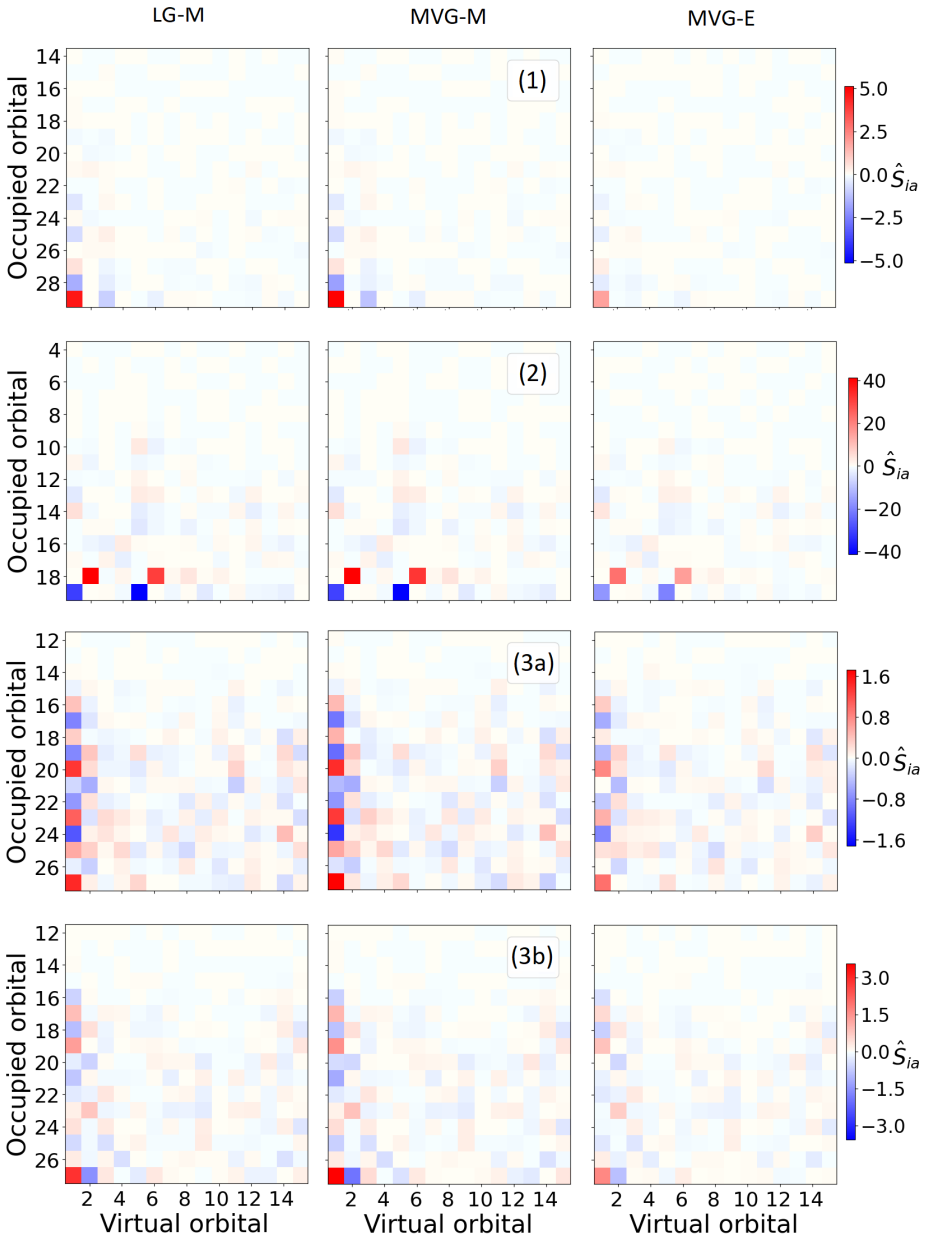


Figure 3.3: $\hat{S}_{ia}$ values for the 15 highest occupied and 15 lowest virtual MOs for molecules **1**-**3** computed with (from left to right) the LG-M, MVG-M, and MVG-E definitions of $\tilde{S}$.

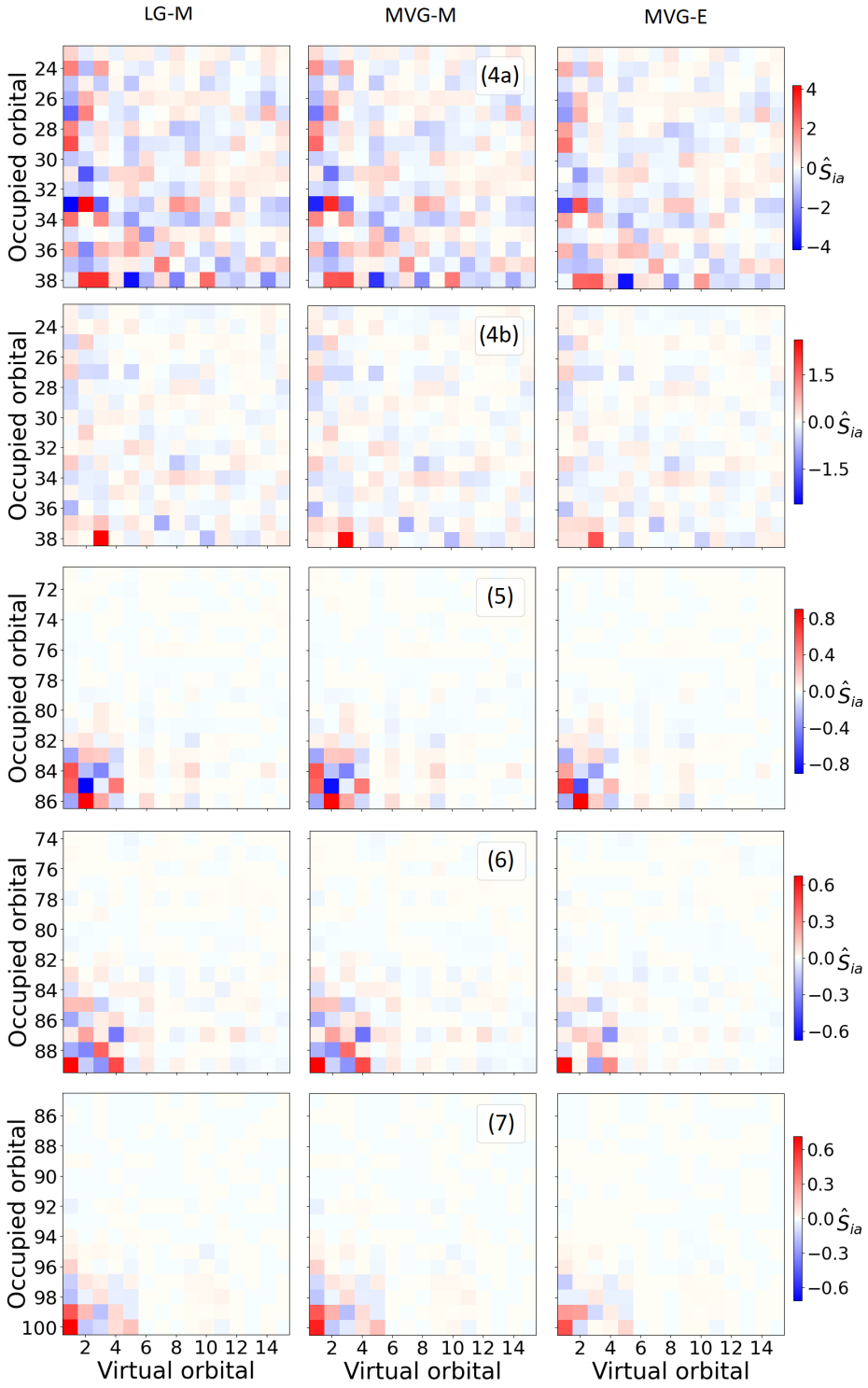


Figure 3.4: $\hat{S}_{ia}$ values for the 15 highest occupied and 15 lowest virtual MOs for molecules **4**-**7** computed with (from left to right) the LG-M, MVG-M, and MVG-E definitions of $\tilde{S}$.

is multiplied by $\omega^{-\frac{1}{2}}$, evenly splitting the factor of $\omega^{-1}$ in Eq. 3.5. This factor arises from a general response function relationship rather than a specific transformation of an operator, so the choice of how it is partitioned between the two vectors is somewhat arbitrary.[60]

Figure 3.5 depicts these vectors for each molecule, calculated with each $\tilde{S}$ definition. Focusing first on molecule **1**, it is clear from the figure that the vectors for all types of $\tilde{S}$ have the same orientation and involve the same occupied-virtual MO pair, indicating that the process is indeed the same. While the $\tilde{S}$ values for LG-M and MVG-M values are very similar in magnitude, the lengths of their vectors differ, see the left and center columns of Figure 3.5. This is due to the differences in length nearly canceling in the product (i.e. the electric vector is 1.84 times larger for LG-M than MVG-M, but the magnetic vector is 1.94 times smaller). On the other hand, the MVG-E $\tilde{S}$ value is smaller because its shorter magnetic vector length (1.24 times smaller than LG-M, 2.41 times smaller than MVG-M) is not compensated for by the electric vector length (1.93 times smaller than LG-M, 1.05 times smaller than MVG-M).

The other molecules exhibit the same qualitative behavior, where the largest $\tilde{S}_{ia}$ contributions come from the same transitions independently of the $\tilde{S}$ definition. Nonetheless, there are still some noteworthy differences between the $\tilde{S}$ definitions. Comparing LG-M and MVG-M (same perturbation, different gauge), these produce similar $\tilde{S}_{ia}$ values for a given molecule. However, the electric and magnetic vectors of these transitions differ in magnitude, with LG-M having consistently longer magnetic vectors and MVG-M having consistently longer electric vectors. For the electric vectors, LG-M and MVG-M "electric dipoles" are not exactly equivalent, as Eq. 3.2 is only an approximation in an incomplete basis. For the magnetic vectors, while they are both ostensibly the magnetic perturbed density for a 589.3 nm perturbation, LG-M incorporates additional terms due to the use of GIAOs and MVG-M subtracts out the spurious zero-frequency perturbed density, likely resulting in different magnitudes. This suggests the magnitude of the LG-M and MVG-M vectors can only be roughly compared.

Comparing MVG-M and MVG-E (same gauge, different perturbation), the $\tilde{S}_{ia}$ values differ mostly because of the different magnitude of the corresponding transition vectors, as shown in

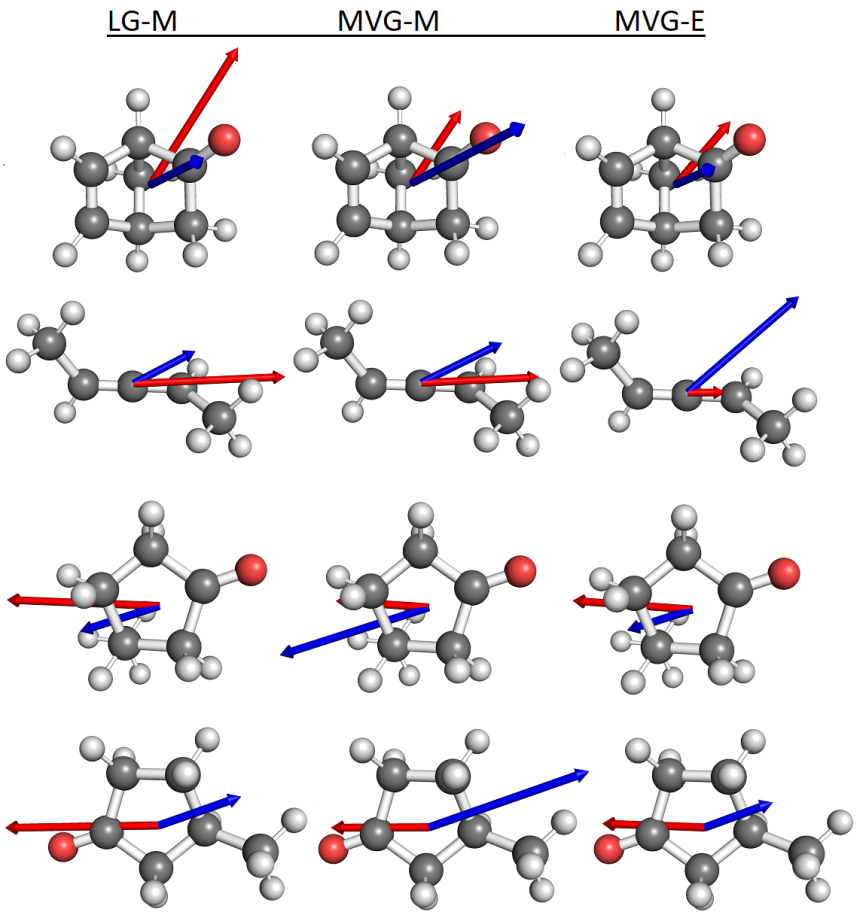


Figure 3.5: Electric (red) and magnetic (blue) dipole or perturbed density vectors for the largest transition of molecules **1**-**3**, computed with (from left to right) the LG-M, MVG-M, and MVG-E definitions of $\tilde{S}$. The transitions depicted are: **1** (HOMO→LUMO), **2** (HOMO→LUMO+4), **3a** (HOMO→LUMO), **3b** (HOMO→LUMO). For visibility, the length of the largest electric and magnetic vectors for each molecule or conformer are fixed at an arbitrary value and the other vectors are scaled relative to this length.

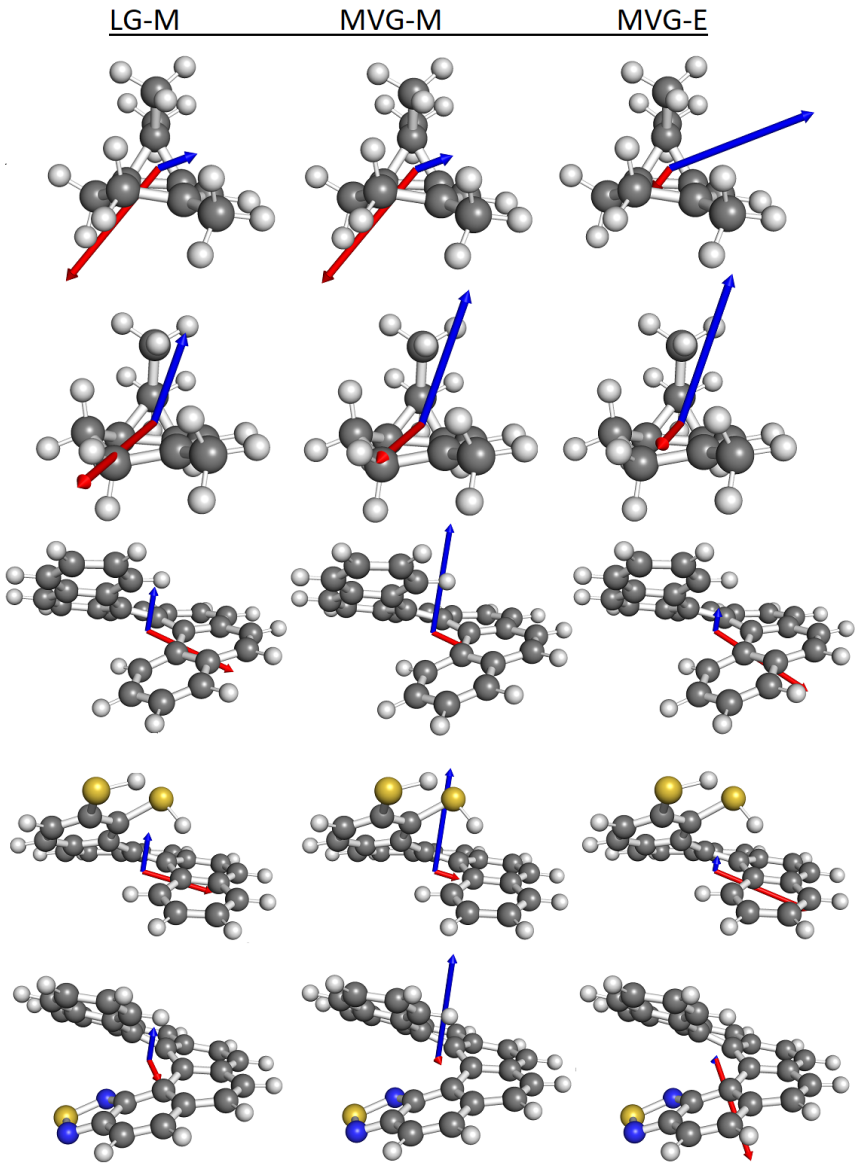


Figure 3.6: Electric (red) and magnetic (blue) dipole or perturbed density vectors for the largest transition of molecules **4**-**7**, computed with (from left to right) the LG-M, MVG-M, and MVG-E definitions of $\tilde{S}$. The transitions depicted are: **4a** (HOMO-5→LUMO), **4b** (HOMO→LUMO+2), **5** (HOMO→LUMO+1), **6** (HOMO→LUMO), and **7** (HOMO→LUMO). For visibility, the length of the largest electric and magnetic vectors for each molecule or conformer are fixed at an arbitrary value and the other vectors are scaled relative to this length.

Table S11 of the SI. This is due to how the perturbed density is computed. The main difference can be explained by looking at the form of the initial guess for the perturbed density at the beginning of the iterative solution of the LR equations:

$$\langle \phi_a | X^{\pm}_{y\alpha} | \phi_i \rangle^{(0)} = \frac{\langle \phi_a | y_\alpha | \phi_i \rangle}{\Delta\varepsilon_{ia} \mp \omega} \tag{3.8}$$

where $\Delta\varepsilon_{ia}$ is the orbital energy difference and the 0 superscript denotes the initial guess. For this initial guess, the MVG-M and MVG-E $\tilde{S}_{ia}$ values are identical, but the vectors differ in length, as the factor of $\Delta\varepsilon_{ia} \mp \omega$ is incorporated into the magnetic vector for MVG-M and the electric vector for MVG-E. Furthermore, in solving the LR-SCF equations to obtain a converged perturbed density, these initial guess terms become intermixed, changing the value of the perturbed density and consequently $\tilde{S}_{ia}$. This intermixing is responsible for the slight difference in angles between MVG-E and MVG-M $\tilde{S}_{ia}$ values, see Table S11 of the SI. However, this effect is small and the MVG-M and MVG-E vectors essentially have the same orientation for all the molecules tested, as shown in Figures 3.5-3.6. It is also apparent from Section 3.4.2 that while the MVG-M and MVG-E have different $\tilde{S}_{ia}$ values, the distribution of $\tilde{S}$ is the same (e.g. the same transitions are the largest contributors for each definition). Since the different definitions agree as to which transitions are most important (as shown by the matching distribution of $\hat{S}$) and these transitions describe qualitatively the same movement of charge (as shown by the matching vector orientation for each definition), we can say that each definition gives a consistent picture of the underlying cause of OR for a given molecule.

## 3.5 Discussion and Conclusions

We have investigated whether different definitions of the $\tilde{S}$ rotatory strength, based on different choices of gauge and perturbed density, provide the same or different pictures for the qualitative description of the optical rotation of chiral molecules. We used a sample of chiral molecules previously studied using the LG-M definition. We find that these different $\tilde{S}$ definitions produce

consistent pictures of what physical processes contribute to the optical rotation. This is true at multiple levels of detail. As shown in Section 3.4.1, partial sums of $\hat{S}$ over selected sets of transitions give similar convergence to the total OR. Going one level deeper, the distribution of $\hat{S}$ within these sets is mostly the same for each perturbation, as shown in Section 3.4.2 for the frontier orbital transitions. Finally, at the level of individual transitions, Section 3.4.3 shows that the largest transitions correspond to the same physical process even though the relative magnitude of the transition vectors changes with different $\tilde{S}$ definition. The difference in vector magnitudes is due to the choice of dipole representation (LG-M vs MVG-M) or to the choice of perturbation density (MVG-M vs MVG-E), but it has no bearing on the qualitative interpretation of the electron transition for a particular *ia* orbital pair. Therefore, we posit that $\tilde{S}$ analysis is largely independent of the choice of gauge, at least in so far as the different definitions produce similar total OR, and in the choice of perturbation.

# Chapter 4

# Exploration of Methods for Origin Invariant Decomposition of Optical Rotation

## 4.1 Introduction

Optical rotation (OR) measurements are a common method to characterize chiral systems, but it has proven challenging to establish a chemically intuitive connection between the structure of a molecule and the direction or magnitude of OR it induces. With the advent of quantum mechanical simulations of OR, significant efforts have been made to clarify this structure-property relationship. One avenue of research has focused on developing methods to decompose the OR into contributions from functional groups, bonds, or even individual atoms within the molecule.[31,33–35,38,39]

We have recently developed a scheme, referred to as $\tilde{S}$ analysis, whereby the OR is expressed as a sum of contributions $\tilde{S}_{ia}$ from occupied $i$ to virtual $a$ molecular orbital transitions.[39] We can determine what physically is driving OR and what parts of the molecule are involved by analyzing the movement of charge density described by these transitions. Using this framework, we have investigated the influence of molecular conformation and functionalization on the OR induced by rigid organic molecules.[40,87] As we showed in Chapter 3, the method is robust, providing a consistent physical interpretation of the OR provided in both modified velocity gauge (MVG) and length gauge (LG) calculations, regardless of the choice of perturbation used in solving the linear response equations.[79]

However, while both MVG and LG calculations of the total OR are (or can be made)[60,62] origin invariant, this does not necessarily ensure that individual $\tilde{S}_{ia}$ values will be origin invariant.

The total OR is in general made origin invariant via cancellation, rather than elimination, of the origin dependent terms from each $\tilde{S}_{ia}$ value. It is possible that choosing a physically meaningful origin like the center of mass may produce reasonable $\tilde{S}_{ia}$ values, though this is not guaranteed; e.g. typical LG CCSD calculations of the total OR are origin dependent and using the center of mass as the origin does not consistently lead to high accuracy in comparison with experiment.[97] In any case, the reliability of the physical interpretations provided by $\tilde{S}$ analysis would be greatly improved by removing this origin dependence.

In this chapter, we discuss two possible approaches to compute origin invariant $\tilde{S}$ values in the modified velocity gauge. The chapter is organized as follows. Section 4.2 contains a brief review of the theory behind $\tilde{S}$ analysis and then the derivation of two origin invariant formulations. The first of these procedures uses an average of the $\tilde{S}_{ia}$ values obtained from MVG electric and magnetic perturbation calculations. The second approach involves splitting the response matrix in order to form "hemi-perturbed densities" that are then contracted to form $\tilde{S}_{ia}$ values. In Section 4.3, we demonstrate both approaches by performing calculations on two small organic molecules, P-(2,3)-pentadiene and (R)-3-chloro-1-butene. We conclude with a discussion of the relative merits of these two approaches Section 4.4.

## 4.2 Theory

We begin by briefly reviewing the methodology for $\tilde{S}$ analysis. In the modified velocity gauge, the electric dipole-magnetic dipole polarizability tensor $\boldsymbol{\beta}$ can be computed through linear response theory as[28,64,88]

$$\beta_{\alpha\beta} = \frac{1}{\omega^2}\left[\underline{\mu}_\alpha^{V\dagger}\underline{D}^{\omega}\underline{m}_\beta - \underline{\mu}_\alpha^{V\dagger}\underline{D}^{0}\underline{m}_\beta\right] \tag{4.1}$$

where $\underline{D}^{\omega} = \underline{\Omega}^{-1}$ is the linear response matrix (explicitly defined in Eq. 2.23) for the perturbation frequency $\omega$, $\underline{D}^{0}$ is the zero frequency response matrix, and $\underline{\mu}_\alpha^V$ and $\underline{m}_\beta$ are respectively the $\alpha$ and $\beta$ Cartesian components of the velocity gauge electric and magnetic dipoles, represented as vectors over all orbital pairs in the given basis. For experiments in isotropic media, the observed

optical rotation is proportional to the trace of $\boldsymbol{\beta}$. We can gain insight into what underlying physical processes are contributing to this trace by splitting it into contributions from orbital pairs, referred to as $\tilde{S}$ values

$$\mathrm{Tr}(\boldsymbol{\beta}) = \frac{1}{\omega}\sum_{ia}\tilde{S}^{\omega}_{ia} - \tilde{S}^{0}_{ia} = \frac{1}{\omega}\sum_{ia}\tilde{S}^{\mathrm{MVG}}_{ia} \tag{4.2}$$

In MO basis, we can express these contributions as

$$\tilde{S}^{\mathrm{M},\omega}_{ia} = \frac{1}{\omega}\sum_{\alpha}\sum_{jb}\langle\phi_i|\mu^{V}_{\alpha}|\phi_a\rangle\left(D^{\omega}_{ia,jb}\langle\phi_b|m_{\alpha}|\phi_j\rangle\right) \tag{4.3a}$$

$$\tilde{S}^{\mathrm{E},\omega}_{ia} = \frac{1}{\omega}\sum_{\alpha}\sum_{jb}\langle\phi_i|m_{\alpha}|\phi_a\rangle\left(D^{\omega}_{ia,jb}\langle\phi_b|\mu^{V}_{\alpha}|\phi_j\rangle\right) \tag{4.3b}$$

where in 4.3a, the response matrix is contracted with the magnetic dipole (MVG-M) and in 4.3b, the response matrix is contracted with the electric dipole (MVG-E). The expressions for the static frequency $\tilde{S}^{0}_{ia}$ are the same, but with $\underline{D}^{\omega}$ replaced by $\underline{D}^{0}$ (the leading factor of $\omega$ is retained). Note Eqs. 4.3a-b are equivalent to the expressions for MVG-M and MVG-E in Eqs. 3.5a-b; we write them here in terms of the response matrix as this makes it easier to present how the origin dependence emerges (and can be removed). We have previously shown that, in general, $\tilde{S}^{\mathrm{MVG\text{-}M}}_{ia} \neq \tilde{S}^{\mathrm{MVG\text{-}E}}_{ia}$; however, the $\tilde{S}_{ia}$ values for each choice of perturbation are associated with the same physical process and the relative contribution of this process to the observed OR is roughly the same.[79]

While the total OR computed in the modified velocity gauge is origin invariant regardless of the perturbation used, this is not the case for the individual $\tilde{S}_{ia}$ values, even in the limit of an exact calculation using a complete basis set. The origin dependence of $\tilde{S}_{ia}$ for each choice of perturbation

can be expressed as

$$\tilde{S}^{\mathrm{M},\omega}_{ia}(\boldsymbol{O}') = \frac{1}{\omega}\sum_{\alpha}\sum_{jb}\langle\phi_i|\mu^V_\alpha|\phi_a\rangle D^{\omega}_{ia,jb}\left(\langle\phi_b|m_\alpha|\phi_j\rangle - \sum_{\beta\gamma}\varepsilon_{\alpha\beta\gamma}\langle\phi_b|\mu^V_\beta|\phi_j\rangle d_\gamma\right)$$
$$= \tilde{S}^{\mathrm{M},\omega}_{ia}(\boldsymbol{O}) - \frac{\omega}{2}\sum_{\alpha\beta\gamma}\varepsilon_{\alpha\beta\gamma}\alpha^{V,\omega}_{\alpha\beta,ia}d_\gamma \tag{4.4a}$$

$$\tilde{S}^{\mathrm{E},\omega}_{ia}(\boldsymbol{O}') = \frac{1}{\omega}\sum_{\alpha}\sum_{jb}\left(\langle\phi_i|m_\alpha|\phi_a\rangle - \sum_{\beta\gamma}\varepsilon_{\alpha\beta\gamma}\langle\phi_i|\mu^V_\beta|\phi_a\rangle d_\gamma\right)D^{\omega}_{ia,jb}\langle\phi_b|\mu^V_\alpha|\phi_j\rangle$$
$$= \tilde{S}^{\mathrm{E},\omega}_{ia}(\boldsymbol{O}) + \frac{\omega}{2}\sum_{\alpha\beta\gamma}\varepsilon_{\alpha\beta\gamma}\alpha^{V,\omega}_{\alpha\beta,ia}d_\gamma \tag{4.4b}$$

where $\boldsymbol{\varepsilon}$ is the Levi-Civita symbol, $\boldsymbol{\alpha^V}$ is the modified velocity gauge polarizability, and $\boldsymbol{O}' = \boldsymbol{O} + \boldsymbol{d}$ is a shifted origin, displaced by a vector $\boldsymbol{d}$. The origin dependence for each choice of perturbation comes from differences between off-diagonal elements of $\boldsymbol{\alpha}^V_{ia}$; the total MVG $\boldsymbol{\alpha}$ is symmetric, so these terms cancel out when summing over all occupied-virtual orbital pairs, but each $\boldsymbol{\alpha}^V_{ia}$ is in general not symmetric.

The form of Eqs. 4.4a and 4.4b suggests a simple approach for eliminating origin dependence from $\tilde{S}_{ia}$:

$$\tilde{S}^{\mathrm{Avg}}_{ia} = \frac{1}{2}\left(\tilde{S}^{\mathrm{MVG\text{-}M}}_{ia} + \tilde{S}^{\mathrm{MVG\text{-}E}}_{ia}\right) \tag{4.5}$$

Since the origin dependent terms for $\tilde{S}^{\mathrm{MVG\text{-}M}}_{ia}$ and $\tilde{S}^{\mathrm{MVG\text{-}E}}_{ia}$ are equal in magnitude but opposite in sign, they cancel in the average, leaving $\tilde{S}^{\mathrm{Avg}}_{ia}$ origin invariant. $\tilde{S}^{\mathrm{Avg}}_{ia}$ still satisfies Eq. 4.2, which means it can still be used as an interpretive tool to understand which occupied to virtual orbital transitions are driving OR. However, while $\tilde{S}^{\mathrm{MVG\text{-}M}}_{ia}$ and $\tilde{S}^{\mathrm{MVG\text{-}E}}_{ia}$ each have associated electric/-magnetic dipole and perturbed density vectors describing the flow of charge during the $i \rightarrow a$ transition, $\tilde{S}^{\mathrm{Avg}}_{ia}$ does not have an associated set of vectors. Simply averaging the MVG-M and MVG-E electric/magnetic vectors and taking their dot product is not equivalent to averaging the MVG-M and MVG-E dot products from Eqs. 4.3a-b, as is done in Eq. 4.5.

An alternative approach to removing the origin dependence involves mixing the electric and magnetic response prior to forming $\tilde{S}_{ia}$. The origin dependence stems from the asymmetry of $\boldsymbol{\alpha}^V_{ia}$,

so an alternative method to remove the origin dependence is to simultaneously symmetrize $\boldsymbol{\alpha}^V_{ia}$ for every occupied-virtual orbital pair. This can be achieved via Cholesky decomposition[98] of the response matrix

$$\underline{D}^{\omega} = \underline{L}^{\omega}\underline{L}^{\omega\dagger} \tag{4.6}$$

where $\underline{L}^{\omega}$ is a lower triangular square matrix with all positive real values along its diagonal. By splitting the response matrix, we can apply $\underline{L}^{\omega}$ to both the electric dipole and magnetic dipole to obtain "hemi-perturbed" densities and contract these to form

$$\tilde{S}^{\text{Hemi},\omega}_{ia} = \frac{1}{\omega}\sum_{\alpha}\sum_{jb}(L^{\omega}_{ia,jb}\langle\phi_j|\mu^V_{\alpha}|\phi_b\rangle)(L^{\omega}_{ia,jb}\langle\phi_b|m_{\alpha}|\phi_j\rangle) \tag{4.7}$$

To see that this formulation of $\tilde{S}_{ia}$ is origin invariant, we simply need to show that the corresponding electric dipole polarizability

$$\alpha^{\text{Hemi},\omega}_{\alpha\beta,ia} = \frac{1}{\omega}\sum_{jb}(L^{\omega}_{ia,jb}\langle\phi_j|\mu^V_{\alpha}|\phi_b\rangle)(L^{\omega}_{ia,jb}\langle\phi_b|\mu^V_{\beta}|\phi_j\rangle) \tag{4.8}$$

is symmetric, which is apparent since it is formed from the contraction of a vector with itself. This method also does not allow us to assign unique electric and magnetic vectors to each transition. The reason is that for MVG, this approach leads to an electric and magnetic hemi-density both for the applied perturbation frequency and the static frequency. Simply taking the difference of both perturbed and static electric/magnetic hemi-densities and contracting them does not produce the same product as Eq. 4.8. For the standard $\tilde{S}^{\text{MVG-M}}_{ia}$, only the magnetic vector is a density and for $\tilde{S}^{\text{MVG-E}}_{ia}$, only the electric vector is a density; the other vector remains the same in the static and perturbed cases. This allows us to only take the difference of the densities to produce a vector that can be used in the contractions of Eq. 4.3a-b.

Another potential obstacle to this hemi-density approach is that the Cholesky decomposition is only possible for Hermitian, positive-semidefinite matrices (i.e. matrices with only real eigenvalues greater than or equal to 0). While the response matrix is always Hermitian, it is only required to be

positive-definite in the static limit.[64,99] This method would also require extensive updates to the typical linear response implementation, as generally the response matrix is not formed explicitly and instead the perturbed density is formed through an iterative process (see section 2.2).

The goal of this chapter is to assess the applicability of these approaches to remove the origin dependence of $\tilde{S}_{ia}$. We perform Hartree-Fock/aug-cc-pVDZ calculations with a 589.3 nm perturbation wavelength on two small molecules, P-(2,3)-pentadiene (hereafter pentadiene) and (R)-3-chloro-1-butene (hereafter chlorobutene), to compare $\tilde{S}_{ia}^{\text{Avg}}$ and $\tilde{S}_{ia}^{\text{Hemi}}$ with the origin dependent $\tilde{S}_{ia}^{\text{MVG-M}}$ and $\tilde{S}_{ia}^{\text{MVG-E}}$. This comparison will help to determine if there are any differences in the physical interpretation of the OR suggested by each method. The electronic structure calculations are performed in a development version of GAUSSIAN.[95] While $\tilde{S}_{ia}^{\text{MVG-M}}$ and $\tilde{S}_{ia}^{\text{MVG-E}}$ (and consequently $\tilde{S}_{ia}^{\text{Avg}}$) can be directly computed using modified GAUSSIAN routines, $\tilde{S}_{ia}^{\text{Hemi}}$ is currently only available through an external Python script, which reads in one- and two-electron integrals from GAUSSIAN output files in order to explicitly form the response matrix.

It is not our intention in this chapter to compute experimentally accurate OR for these molecules, but rather simply to demonstrate the origin invariance of these $\tilde{S}_{ia}^{\text{Avg}}$ and $\tilde{S}_{ia}^{\text{Hemi}}$. Consequently, Hartree-Fock, rather than DFT, calculations are performed to avoid the need to also read in 4-center, 1-electron exchange-correlation integrals when computing $\tilde{S}_{ia}^{\text{Hemi}}$. While this can be done using the external Python script and further modifications to the GAUSSIAN code, the prototype code for forming and reading in these integrals, in addition to the two-electron integrals, is very time and memory intensive.

## 4.3 Results

Figures 4.1 and 4.2 plot $\tilde{S}_{ia}$ values for the 20 highest occupied and 20 lowest virtual MOs of pentadiene and chlorobutene computed with each of the definitions of $\tilde{S}$ discussed in 4.2. To demonstrate the origin independence of the $\tilde{S}_{ia}^{\text{Avg}}$ and $\tilde{S}_{ia}^{\text{Hemi}}$ approaches, we plot the results of calculations with the origin at the center of mass of the molecule and with the origin shifted by -100Å in each Cartesian direction. Starting first with the center of mass plots, we can see that

for each molecule, the distribution of contributions is very similar for each of the definitions of $\tilde{S}_{ia}$. For pentadiene, the largest negative contribution is from the (HOMO-1, LUMO+12) transition and the largest positive contribution is from the (HOMO, LUMO+14) transition for each of the definitions. Chlorobutene has a larger number of significant contributions that maintain the same sign across definitions and the largest magnitude contributions are consistently from a small band of 3 transitions from (HOMO, LUMO+7) to (HOMO, LUMO+9). That the $\tilde{S}_{ia}^{\text{MVG-M}}$ and $\tilde{S}_{ia}^{\text{MVG-E}}$ distributions closely match the origin invariant $\tilde{S}_{ia}^{\text{Avg}}$ and $\tilde{S}_{ia}^{\text{Hemi}}$ distributions suggests that the center of mass is a reasonable choice of origin, though there is no guarantee that this holds beyond this set of molecules or for calculations employing a different model chemistry.

Moving to the shifted origin plots for each molecule, while the shifted $\tilde{S}_{ia}^{\text{Avg}}$ and $\tilde{S}_{ia}^{\text{Hemi}}$ are identical to their center of mass values (within numerical accuracy), the origin dependent terms of $\tilde{S}_{ia}^{\text{MVG-M}}$ and $\tilde{S}_{ia}^{\text{MVG-E}}$ dominate the shifted origin plots. Due to the large shift in origin (and subsequently large origin dependent contribution), the difference in sign between the origin dependent terms of $\tilde{S}_{ia}^{\text{MVG-M}}$ and $\tilde{S}_{ia}^{\text{MVG-E}}$ is clearly visible in the plots. For pentadiene, the (HOMO-1, LUMO+12) and (HOMO, LUMO+14) MVG-M transitions both changed sign relative to the center of mass calculation, while also increasing in magnitude by a factor of 3, exceeding the range of the center of mass origin color map. $\tilde{S}_{ia}^{\text{MVG-M}}$ for the (HOMO-1, LUMO+10) and (HOMO, LUMO+11) transitions increased in size by more than an order of magnitude. For chlorobutene, many transitions increased in magnitude enough to exceed the range of the center of mass origin color map. The (HOMO, LUMO+7), the largest magnitude $\tilde{S}_{ia}^{\text{MVG-M}}$ of the center of mass calculations, remains the largest transition in the shifted calculations, but increases in size from $7.82 \times 10^{-3}$ to $2.78 \times 10^{-1}$.

## 4.4 Conclusions

In this chapter, we have introduced two new methods for decomposing the OR into origin invariant $\tilde{S}_{ia}$ values. The $\tilde{S}^{\text{Avg}}$ approach exploits the equal magnitude, opposite sign origin dependence of $\tilde{S}^{\text{MVG-M}}$ and $\tilde{S}^{\text{MVG-E}}$ to cancel out the origin dependence. For $\tilde{S}^{\text{Hemi}}$, we contract hemi-perturbed

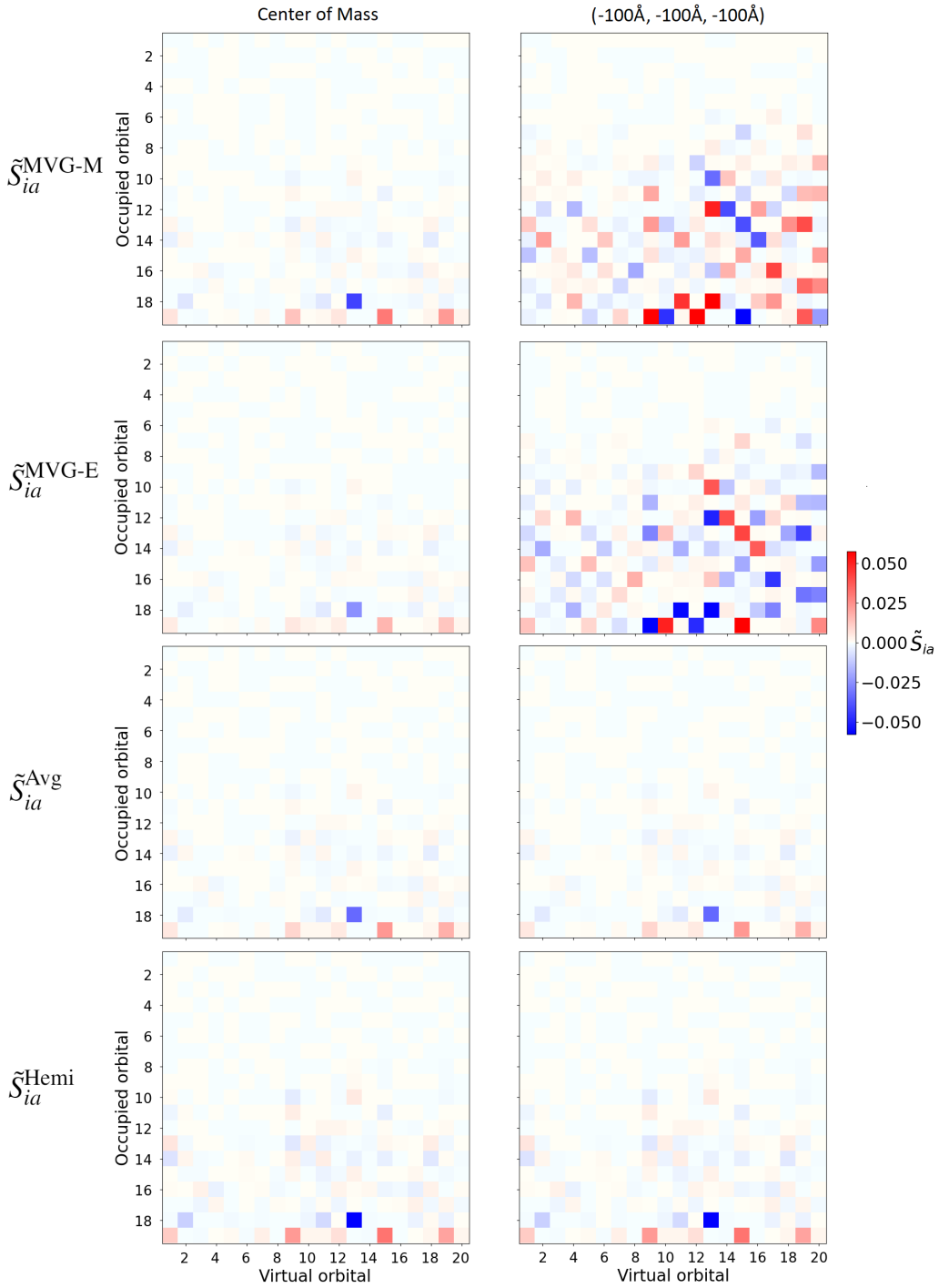


Figure 4.1: $\tilde{S}_{ia}$ values for the 20 highest occupied and 20 lowest virtual MOs of P-(2,3)-pentadiene computed with (from top to bottom) the MVG-M, MVG-E, average, and hemi-density definitions of $\tilde{S}$. Results with the origin at the center of mass and shifted by -100Å in the x, y, and z directions are plotted the in the left and right columns. The heat maps color range is from $-M$ to $M$ where $M = \max(|\tilde{S}_{ia}^{X}|), X = \{\text{MVG-M, MVG-E, Avg, Hemi}\}$ for the center of mass calculations. The shifted origin heat maps use the same color range, but some $\tilde{S}_{ia}$ exceed this range and are simply colored the same as $-M/M$.

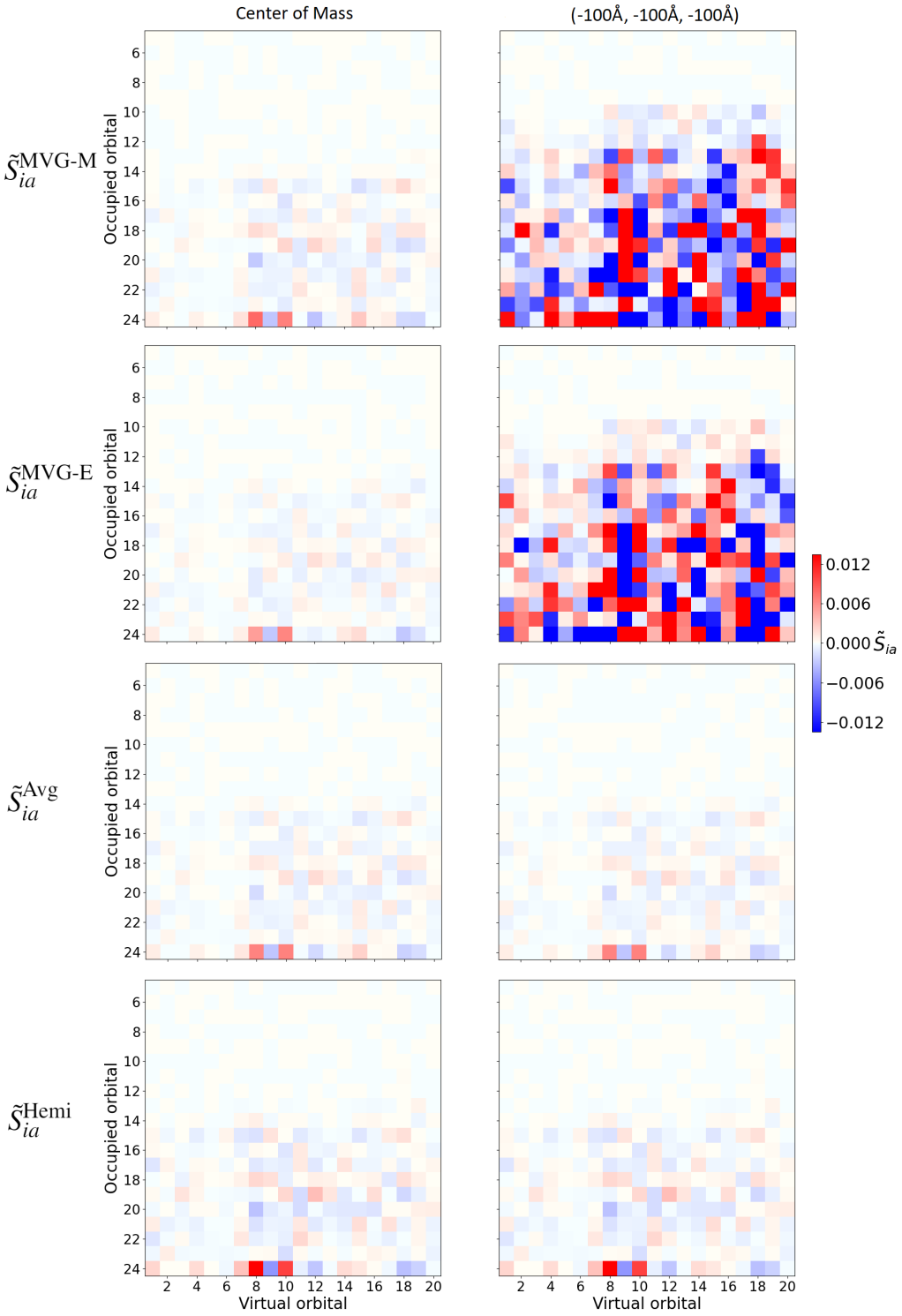


Figure 4.2: $\tilde{S}_{ia}$ values for the 20 highest occupied and 20 lowest virtual MOs of (R)-3-chloro-1-butene computed with (from top to bottom) the MVG-M, MVG-E, average, and hemi-density definitions of $\tilde{S}$. Results with the origin at the center of mass and shifted by -100Å in the x, y, and z directions are plotted the in the left and right columns. The heat maps color range is from $-M$ to $M$ where $M = \max(|\tilde{S}_{ia}^{X}|), X = \{\text{MVG-M, MVG-E, Avg, Hemi}\}$ for the center of mass calculations. The shifted origin heat maps use the same color range, but some $\tilde{S}_{ia}$ exceed this range and are simply colored the same as $-M/M$.

densities formed by splitting the linear response matrix evenly between the electric and magnetic dipoles. The MVG hemi-density electric-electric polarizability for each transition $\boldsymbol{\alpha}_{ia}^{\mathrm{Avg}}$ is inherently symmetric, which eliminates the asymmetric, origin dependent contribution to $\tilde{S}_{ia}$. From testing on two small molecules, we found that these approaches produce similar distributions of $\tilde{S}_{ia}$, regardless of the choice of coordinate origin.

Further development and research is still needed to ensure these methodology can be applied in general. While $\tilde{S}_{ia}^{\mathrm{Avg}}$ is simple to compute, its lack of a unique set of associated electric and magnetic vectors limits its use for determining the physical process associated with a particular transition. Since $\tilde{S}_{ia}^{\mathrm{Avg}}$ is simply an average of MVG-M/MVG-E, it may be possible to obtain a set of vectors from some linear combination of the $\tilde{S}_{ia}^{\text{MVG-M}}$/$\tilde{S}_{ia}^{\text{MVG-E}}$ vectors, but thus far it is only clear that an average of the MVG-M/MVG-E vectors will not suffice for this purpose. The hemi-density approach is also promising, producing a similar distribution to the average MVG $\tilde{S}_{ia}$. However, explicitly forming the response matrix as we have done here is prohibitively slow and memory intensive, so we will need to determine how to best incorporate the Cholesky decomposition into the iterative matrix-vector product routines that are typically used to solve the coupled perturbed-self consistent field (CP-SCF) equations. The Cholesky decomposition is also only possible when the response matrix is positive-semidefinite, which is not guaranteed outside the static limit.[64] In fact, we attempted calculations on (1S,4S)-norbornenone using the same model chemistry, but these failed due to negative eigenvalues in the response matrix. We could approximate the response matrix by removing these negative eigenvalues, but this would only be reasonable if they were very small in magnitude and determining the eigenvalues would also likely be prohibitively costly. We will need to determine if an alternative decomposition could be used in these cases or if a simpler, less approximate procedure can be developed to "cure" the response matrix.

Even with these caveats, these approaches offer promising avenues towards the development of a fully origin invariant $\tilde{S}$ analysis framework. At the moment, these approaches already allow for the detection of severe origin dependence in the computed $\tilde{S}_{ia}$ values, helping us to avoid drawing faulty physical interpretations from computational artifacts. Using the distribution of

origin invariant $\tilde{S}$ as a reference, we can tune the origin of MVG-M or MVG-E calculations to minimize the effect of origin dependence. Going forward, we will work to extend these methods to allow for the computation of origin invariant electric and magnetic vectors associated with each transition. We will also explore if these schemes, or some variation, can be applied when using the length gauge with GIAOs[61] or LG(OI).[62]

# Part II

# Optical Rotation of Periodic Systems

# Chapter 5

# Helical Chains of Diatomic Molecules as a Model for Solid State Optical Rotation

(This work taken with the permission of Ty Balduf and Marco Caricato from *J. Phys. Chem. C* **2019**, *123*, 4329-4430.[52] Supporting information is available online.)

## 5.1 Introduction

Controlling supramolecular chirality is an active area of research with applications in nanostructure assembly,[81,100,101] drug design,[2] and enzymatic/bio-inspired catalysis.[8,80] Optical activity measurements are commonly used to distinguish the chiral species generated in these experiments, but it remains poorly understood how the structure of a molecule and intermolecular interactions relate to the observed optical rotation (OR, referred to as *specific rotation* when it is normalized with respect to concentration and path length). Comparisons between computation and experiment could clarify this structure-property relationship, but such comparisons are hampered by limitations of existing theoretical methods. Most experimental measurements of optical activity are carried out in solution, but theoretical methods to account for solvent effects are either highly costly or unable to account for specific solvent-solute interactions, making comparison with experiment challenging. Calculations of OR in isotropic media are also highly sensitive to cancellation effects: elements of the OR tensor can vary in sign and magnitude, so the calculated specific rotation, proportional to the trace of the OR tensor,[49,51] can be greatly influenced by small errors in OR tensor components. In experiment, one only obtains the isotropic specific rotation, not the OR tensor components, so comparison can show that a calculation is incorrect, but not what is causing the error.

Some experimental work has sought to address these limitations by making measurements in different phases. Vaccaro *et al.* have used cavity ring-down polarimetry to study optical rotation in the gas phase.[102,103] Gas phase measurements and calculations do not need to take into account intermolecular interactions, but they still need to deal with spatial averaging of the observable. Kahr and coworkers have enhanced techniques for isolating the OR of crystals from the typically much larger linear birefringence signal.[104–106] These solid state measurements represent valuable benchmark data for theoretical comparison. While it is in principle possible to calculate the OR of solids and extended systems, one issue that remains is the development of a general computational procedure to do so. Perhaps due to the lack of experimental data available, neither classical nor first principles techniques for calculating the OR of solids have been thoroughly pursued. To the best of our knowledge, the works of Zhong *et al.*,[47,48] using a combination of local density approximation density functional theory (DFT) and Green's functions, and Devarajan and Glazer,[46] utilizing the classical polarizability theory of optical rotation, have been the only attempts at developing such methods. However, both approaches rely on *ad hoc* assumptions and parametrizations, thus lacking a truly general applicability.

In this chapter, we calculate the OR tensor of long helical chains of diatomic molecules (using the well established procedure for isolated molecules) as a simple model for the behavior of extended systems. These calculations test the specifications, e.g. cell size and basis set requirements, needed to accurately calculate the OR of infinite systems with the goal of using this information to develop and benchmark an efficient, general-purpose procedure to compute the optical activity of solids, in particular molecular crystals.

The chapter is organized as follows. In section 5.2, we describe our computational procedures for constructing the model helices and calculating their OR tensors. In section 5.3, we present how the OR is influenced by the length and orientation of the helices, as well as the basis set used. We conclude the paper with a discussion of these results and their implications in section 5.4.

## 5.2 Computational Procedure

In this chapter, we consider helices formed from five diatomic molecules: $H_2$, $N_2$, $F_2$, HF, and LiH. Note that the optical activity in these model systems is solely generated by the supramolecular structure of the helices, as the individual molecular units are achiral. To generate the helices, we first orient a single diatomic molecule along the y-axis, with the origin centered either on one of its atoms (CA) or its center of mass (COM). We then add more units with coordinates determined by translating a distance $R$ along the z-axis and rotating $\theta$ degrees counterclockwise in the xy plane. For each diatomic molecule, helices were generated with 15° increments up to 165°. Helices with 0° or 180° rotation have a plane of symmetry and so they do not exhibit isotropic specific rotation; for these helices, this is not due to cancellation in the trace of the OR tensor, but rather from each diagonal component individually being zero. For this reason, we do not perform any further analysis of the OR tensor for these helix conformations. The distance $R$ varies for the different molecules, with $H_2$ units separated by 2.0 Å; $N_2$, $F_2$, and HF units separated by 3.5 Å; and LiH units separated by 5.0 Å. These distances were chosen so that the molecular units would be close enough to retain significant intermolecular interactions while avoiding large electron density overlap between neighboring units. Experimental bond lengths from the NIST CCCBDB database[107] were used for each molecule: $R_{H_2} = 0.741$ Å, $R_{N_2} = 1.098$ Å, $R_{F_2} = 1.412$ Å, $R_{HF} = 0.917$ Å, $R_{LiH} = 1.595$ Å.

To compute the full, oriented OR of Eq. 2.7, denoted here as $\tilde{\mathbf{B}}$, we employ the linear response formalism described in section 2.2. All calculations were carried out using a development version of the GAUSSIAN suite of programs,[95] with the CAM-B3LYP functional[108] and aug-cc-pVDZ basis[93] in the length gauge representation, using gauge including atomic orbitals (GIAOs)[61] to ensure origin independence and 589.3 nm perturbation wavelength. Origin independence could also be achieved using the modified-velocity gauge without GIAOs, but this leads to slower convergence of the OR with basis set size.[88] This choice of functional and basis set has been shown to accurately reproduce the OR obtained from experiment[51,109] and higher levels of theory.[110]

Since these helical systems cannot be experimentally realized nor have they been studied com-

putationally, it is useful to have a model that can be related qualitatively to the calculated results. We compare these computed values with predictions based on the semi-empirical Kirkwood polarizability model,[33,34,53,111] which expresses the optical rotation in terms of the mean polarizability and anisotropy of interacting groups. This model suggests that two interacting groups with $C_{nv}(n>2)$ symmetry will contribute to the optical rotation in proportion to $\sin(2\theta)$, where $\theta$ is the dihedral angle between the groups.[53] This model can offer insight into the locality of interactions within the helices, since interactions beyond those of nearest neighbors should introduce $\sin(2k\theta)$ contributions $(k = 2,3,...)$.

## 5.3 Results

To analyze the OR tensor, we plot the Cartesian components of $\tilde{\mathbf{B}}$, the contributions of $\boldsymbol{\beta}$ and $\mathbf{A}$ to these components, and Tr($\tilde{\mathbf{B}}$). All component contributions/values are scaled by the inverse of the number of units in the helix so that they can be directly related to the specific rotation. Since the helix is not periodic in the x or y direction, the XX and YY component contributions are combined to reflect the averaged value that would be observed in a hypothetical experimental measurement. We include plots of the trace of the OR tensor for completeness and clarity of certain comparisons; we emphasize that for solid state calculations, the specific rotation in each direction is given by Eq. 2.8, rather than the trace of $\tilde{\mathbf{B}}$, as there is no isotropic averaging.

In section 5.3.1, we discuss the dependence of the OR tensor components (in particular, how quickly they converge) on the length of the helix considered. In section 5.3.2, we discuss how the dihedral angle between each molecular unit influences the OR and utilize the Kirkwood model to probe interaction distances within the helices. In section 5.3.3, we examine the basis set dependence of the results obtained for $H_2$.

### 5.3.1 Length Dependence

To model the calculation size necessary to converge the OR tensor components, we created helices of various lengths for each of the diatomic molecules. For $H_2$, we constructed all helices with up to 300 molecules for each dihedral angle and rotation axis considered, while for the other four molecules we obtained all lengths up to 100 molecules. We plot the diagonal Cartesian components of $\tilde{\mathbf{B}}$, as well as the contributions to these components from $\boldsymbol{\beta}$ and $\mathbf{A}$, with respect to the number of molecular units for representative helices with $\theta = 45°$ in Figures 5.1-5.4.

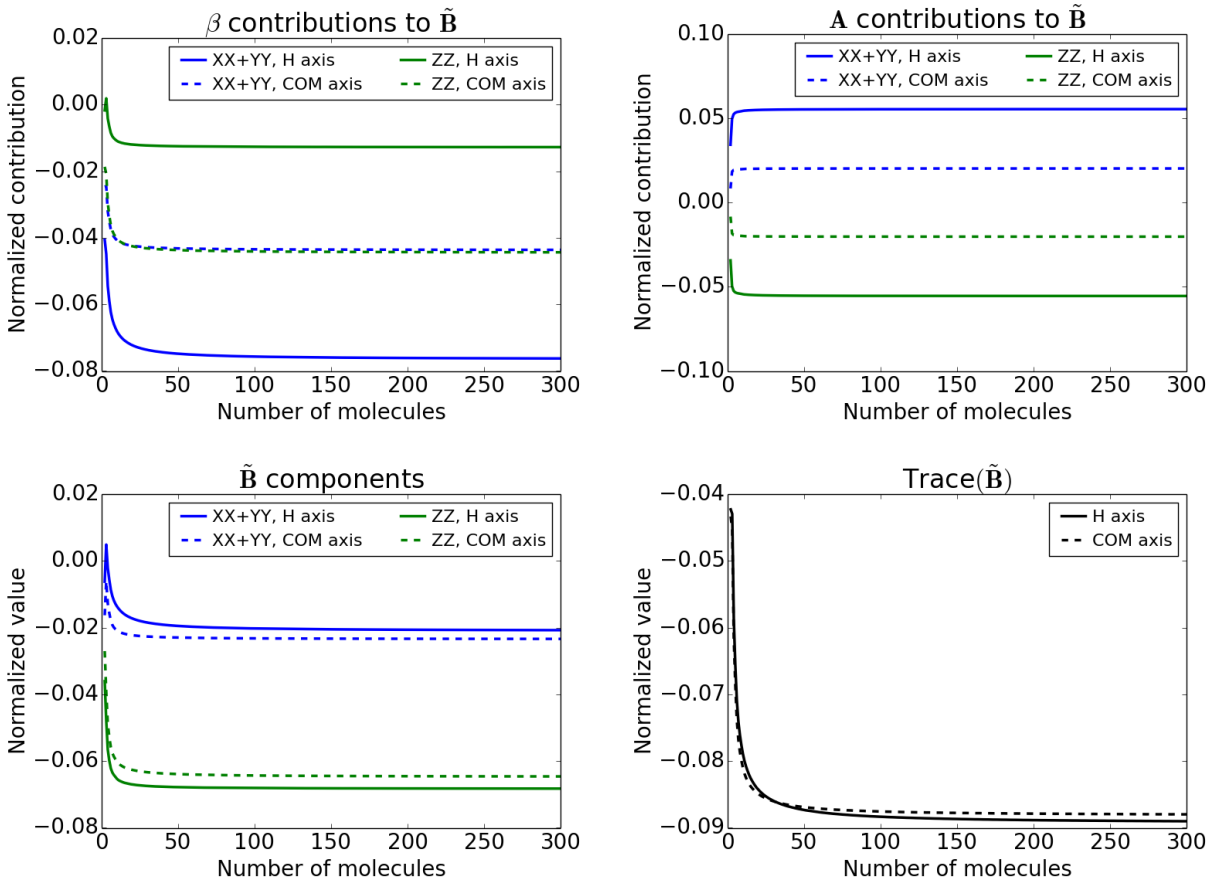


Figure 5.1: Plots of contributions to/components of $\tilde{\mathbf{B}}$ (a.u) for an $H_2$ helix with 45° rotation about the COM and H atom axes. The separation between hydrogen molecules is 2.0 Å. Each value is scaled by the inverse of the number of units in the helix.

As expected based on the sum-over-states formula, the contribution of $\mathbf{A}$ to Tr($\tilde{\mathbf{B}}$) is zero (i.e. XX+YY=-ZZ) and this is the case for all the types of diatomic helices tested. However, the $\mathbf{A}$

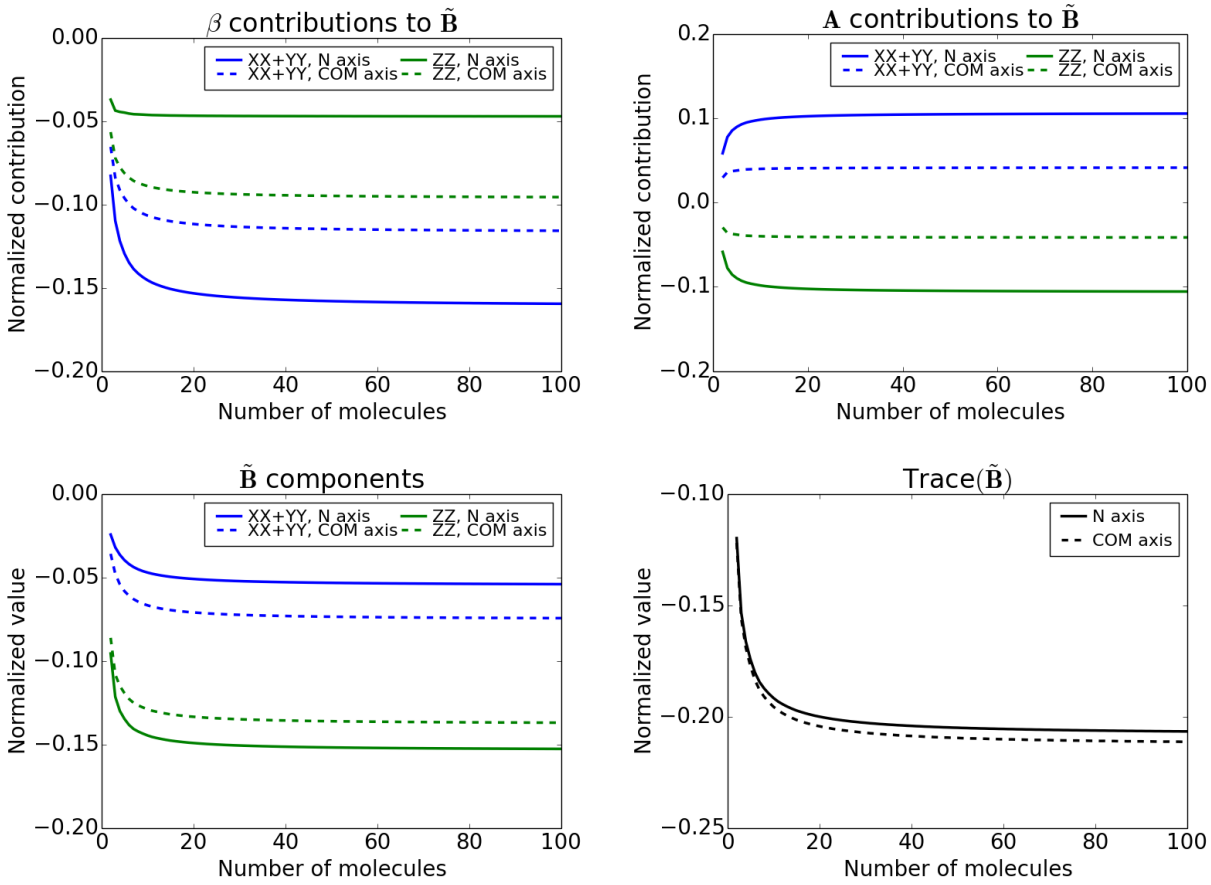


Figure 5.2: Plots of contributions to/components of $\tilde{\mathbf{B}}$ (a.u) for an $N_2$ helix with 45° rotation about the COM axis and N atom axes. The separation between nitrogen molecules is 3.5 Å. Each value is scaled by the inverse of the number of units in the helix.

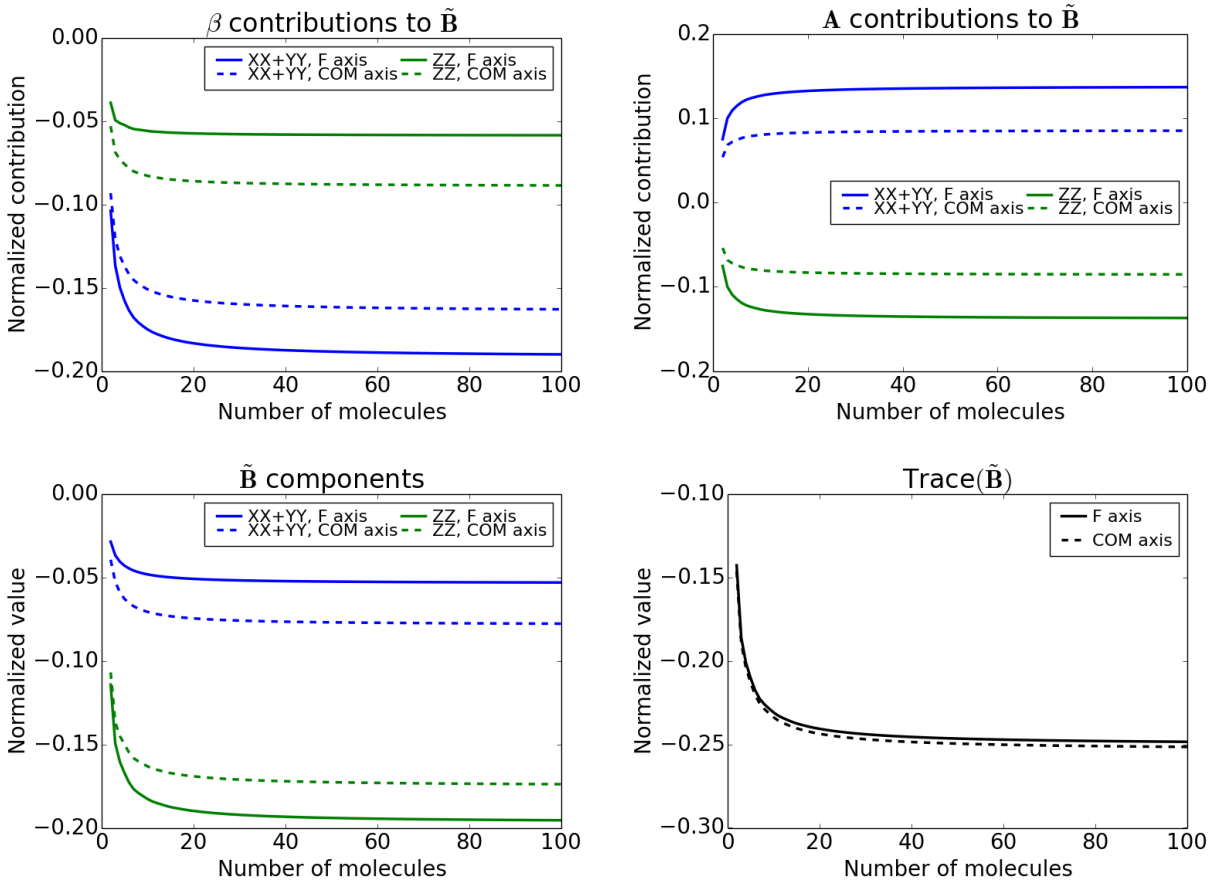


Figure 5.3: Plots of contributions to/components of $\tilde{\mathbf{B}}$ (a.u) for an $F_2$ helix with 45° rotation about the COM and F atom axes. The separation between fluorine molecules is 3.5 Å. Each value is scaled by the inverse of the number of units in the helix.

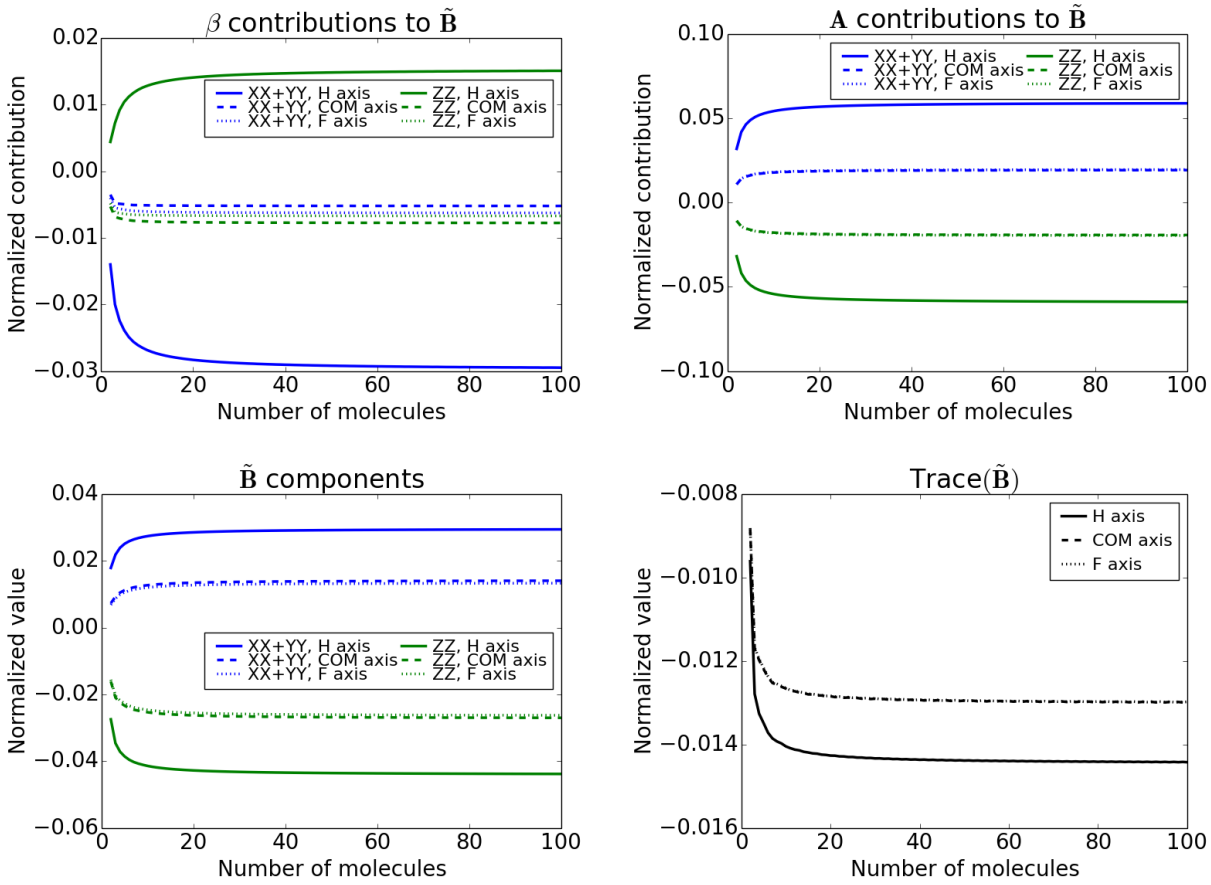


Figure 5.4: Plots of contributions to/components of $\tilde{\mathbf{B}}$ (a.u) for an HF helix with 45° rotation about the COM, H atom, and F atom axes. The separation between HF molecules is 3.5 Å. Each value is scaled by the inverse of the number of units in the helix.

tensor does influence the relative and absolute magnitude of the directional components of $\tilde{\mathbf{B}}$. For the COM helix described in Figure 5.1, had we assumed that $\boldsymbol{\beta}$ on its own represented the OR, the XX+YY and ZZ components would be predicted to have nearly equal magnitude, but with the contribution from $\mathbf{A}$ included, the ZZ component is actually about three times larger than the XX+YY. For the H axis helix, $\mathbf{A}$ contributions essentially invert the ratio between the Cartesian components. These examples illustrate that including the electric dipole-electric quadrupole contribution is crucial to obtain the correct qualitative relationship between the Cartesian components of the OR tensor.

The homonuclear molecules exhibit remarkably similar variation of the OR tensor with respect to length, as can be seen by comparing Figures 5.1-5.3. This similarity holds across the whole range of dihedral angles, which implies that the qualitative OR behavior of these homonuclear systems is largely governed by the helix geometry rather than by the nature of the substituents. For $H_2$, the Cartesian components of $\tilde{\mathbf{B}}$ are typically 95% of their converged value within $\sim 25-35$ molecules for all dihedral angles. For $N_2, F_2$, and HF, around 20 molecules are needed to reach 95% of the converged value for every dihedral angle (LiH exhibited markedly different convergence behavior than the other molecules, thus it will be treated separately below).

For small clusters ($\sim 2-7$ molecules), the OR tensor components do not agree with their converged values. The OR components of the small helices vary inconsistently with increased length, unlike the monotonic convergence seen for larger helices, making it impossible to extrapolate to the large helix limit. In fact, the components for the small helices can even have the wrong sign compared to their converged values. Interestingly, this poor agreement for small helices, and the rate of convergence in general, seem to be independent of unit cell size. The latter depends on the intermolecular relative orientation, determined by the angle $\theta$, and varies between 2(COM homonuclear)/3(CA or heteronuclear) molecular units for $\theta = 90^\circ$ and 12(COM homonuclear)/24(CA or heteronuclear) molecular units for $\theta = 15^\circ, 75^\circ, 105^\circ, 165^\circ$. One might have expected that helices with a smaller unit cell would have required fewer units to converge their OR tensors, as the structure is in some sense more complete for a given number of molecules. Furthermore, molecular

crystals that are periodic in 2 or 3 dimensions will likely require many more molecules to converge the OR tensor than these simple models. Based on the deviation from the converged OR tensor values for these small helices, we expect that calculations on just the unit cell will not be sufficient to obtain accurate results. This indicates that calculations on finite clusters are likely not practical for more realistic systems, and periodic boundary conditions (PBC) may be necessary.[68,75]

Changing the separation distance between molecules does not qualitatively alter $\boldsymbol{\beta}$ and $\mathbf{A}$, but it does affect their relative contributions to $\tilde{\mathbf{B}}$, as shown in Figure 5.5 for $H_2$. These plots are for helices with the same dihedral angle as Figure 5.1, but with the separation distance between units increased from 2.0 Å to 3.0 Å (the latter plots only extend to helices 100 molecules in length, as the tensor components were sufficiently converged by this point). The $\boldsymbol{\beta}$ contribution to each component decreases in magnitude, while the $\mathbf{A}$ contributions increase in magnitude. Since the ZZ contribution from each tensor is negative, $\tilde{B}_{ZZ}$ remains negative and essentially of the same magnitude, but the XX+YY contributions have different signs and the shift in magnitude causes this component of $\tilde{\mathbf{B}}$ to become positive. This again highlights the importance of the $\mathbf{A}$ tensor contributions: not only are the components of $\boldsymbol{\beta}$ and $\tilde{\mathbf{B}}$ qualitatively different for a given separation distance, but they also do not respond in the same manner to a change in separation distance.

In the prior analysis, we have not discussed LiH helices; unlike the other molecules, the $\tilde{\mathbf{B}}$ tensor components for many of the LiH helices do not converge with the length of the helix, as shown in Figure 5.6 by the plots of the H axis rotated helix. The slow convergence of the OR tensor is particularly noticeable in the plot of $\mathrm{Tr}(\tilde{\mathbf{B}})$. Even considering helices of length up to 250 units, $\mathrm{Tr}(\tilde{\mathbf{B}})$ continues to change significantly when additional units are added. Changing the separation distance between LiH units from 5.0 Å to 7.5 Å or 10.0 Å (reported in Figures S1-S2 of the supporting information, SI) did not make the $\mathrm{Tr}(\tilde{\mathbf{B}})$ curves smoother, although it did improve the rate of convergence for the 10.0 Å separation. That the error persists at different separation distances shows that it is not simply a consequence of the chosen geometry. Hartree-Fock calculations of the trace (see Figure S3 in the SI) seem to converge, albeit with oscillations about the converged value. To test whether the error is due to basis set incompleteness, Figure S4 in the SI

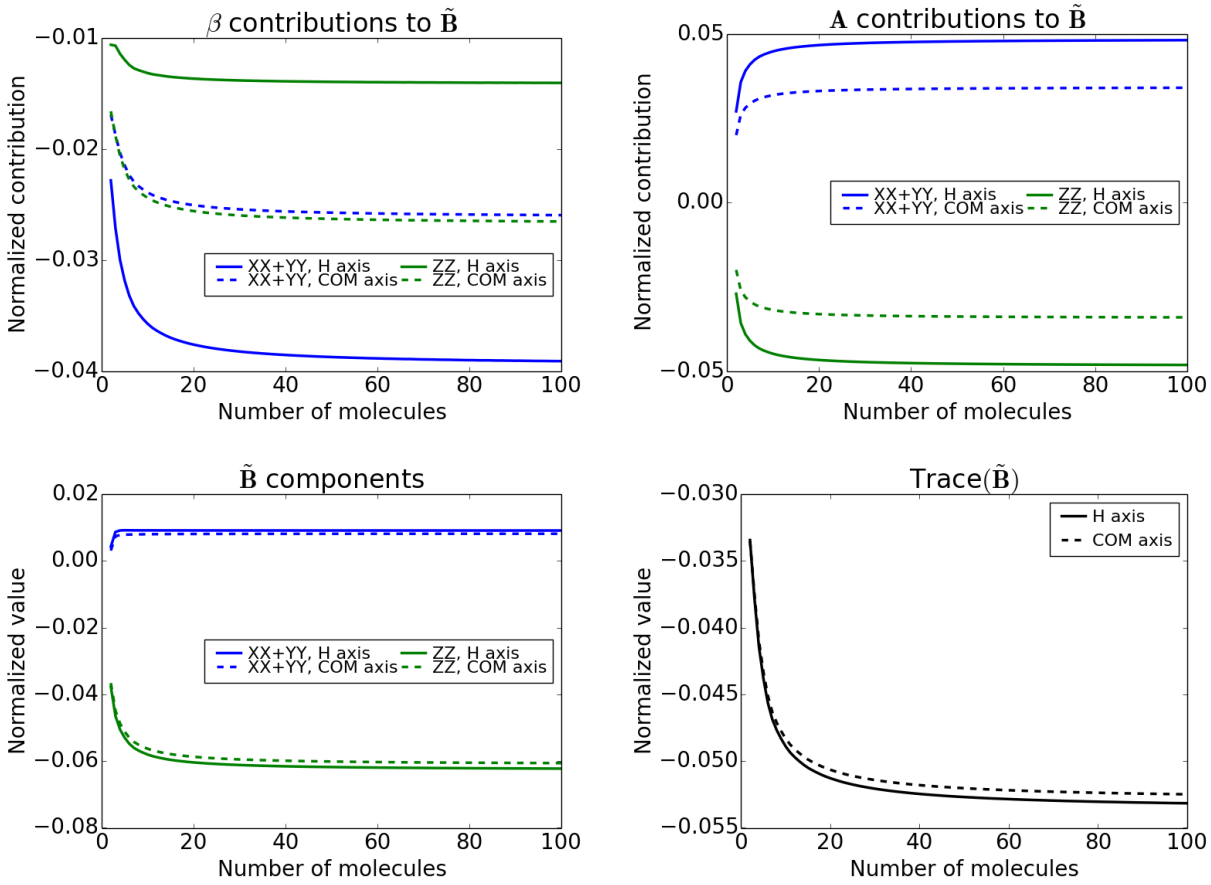


Figure 5.5: Plots of contributions to/components of $\tilde{\mathbf{B}}$ (a.u) for an $H_2$ helix with 45° rotation about the COM axis. The separation between hydrogen molecules is 3.0 Å. Each value is scaled by the inverse of the number of units in the helix.

reports Tr($\tilde{\mathbf{B}}$) calculated with the aug-cc-pVTZ basis, which shows that the OR tensor still fails to converge for some orientations and exhibits small oscillations throughout. These oscillations and jagged appearance of the Tr($\tilde{\mathbf{B}}$) curve occur in both DFT and Hartree-Fock calculations, suggesting they may be caused by edge effects from our finite model. Such edge effects may be present also in cluster models of realistic chiral crystals, further highlighting potential pitfalls of using finite models to describe solid crystals.

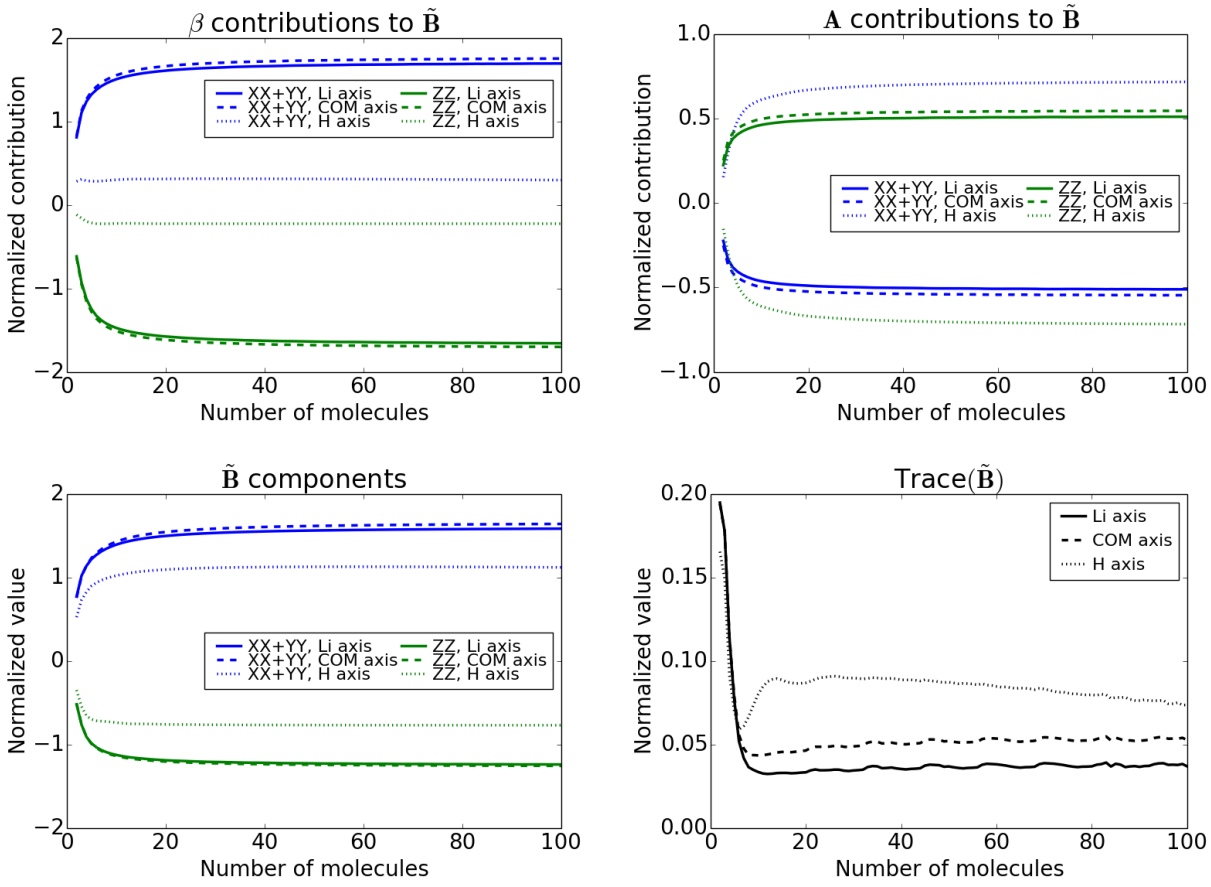


Figure 5.6: Plots of contributions to/components of $\tilde{\mathbf{B}}$ (a.u) for an LiH helix with 30° rotation about the COM, Li atom, and H atom axes. The separation between LiH molecules is 5.0 Å. Each value is scaled by the inverse of the number of units in the helix.

### 5.3.2 Angle Dependence

To study the influence of the helix geometry on the OR, we consider how the converged values of the OR tensor and its contributions vary as functions of the dihedral angle between units. In Figure

5.7, we plot the Tr($\tilde{\mathbf{B}}$) values for a series of $H_2$ helices with different dihedral angles $\theta$ for rotations around the COM and H atom axis. The plots for the individual components of the $\tilde{\mathbf{B}}$ tensor are reported for all molecules (but LiH) in Figures S5-S13 of the SI. The plots in Figure 5.7 seem to have a sinusoidal dependence on the angle that can be rationalized through the semiempirical Kirkwood polarizability model, as mentioned in section 5.2.[33,111] Since the units of the helices have $D_{\infty h}$ or $C_{\infty v}$ symmetry depending on whether the molecule is homonuclear or heteronuclear, the Kirkwood model predicts interacting molecules will contribute to the optical rotation as $\sim \sin(2\theta')$ with respect to the angle $\theta'$ between them. For a given helix, $\theta' = k\theta$, $k = 1, 2, \ldots$.

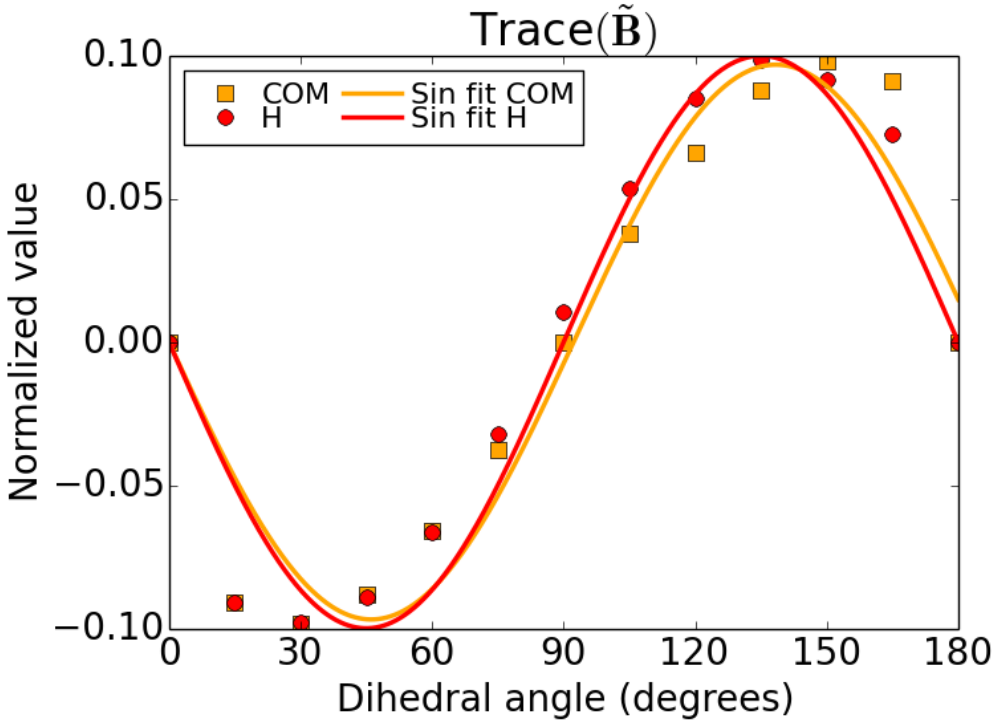


Figure 5.7: Converged Tr($\tilde{\mathbf{B}}$) (a.u.) for $H_2$ helices with $\theta°$ rotation about the H atom and COM axes. Points are the calculated values and solid lines are single sin fits to the same colored calculated values. The separation between hydrogen molecules is 2.0 Å.

To see how well this model describes the results, we fit Tr($\tilde{\mathbf{B}}$) for each type of helix (excluding LiH, as its Tr($\tilde{\mathbf{B}}$) did not converge for many different dihedral angles and rotation axes) to the functional form $C\sin(\nu\theta)$, using a nonlinear regression where $C$ and $\nu$ are varied to minimize the

sum of the mean-squared error with respect to the computed values. Figure 5.7 reports the fitted line for the $H_2$ helices (similar plots for $N_2$, $F_2$, and HF are given in Figures S14-S16 of the SI), while Tables 5.1 and 5.2 report the fitted $C$ and $\nu$ parameters for all molecules. The fits for each axis are approximately $\sin(2\theta)$, which suggests that the largest contribution to the OR comes from nearest neighbor interactions. Interestingly, the individual components of the $(\tilde{\mathbf{B}})$ tensor show a sinusoidal behavior only for helices with the COM axis, while for the helices with the axis on one of the atoms, only $\mathrm{Tr}(\tilde{\mathbf{B}})$ has a sinusoidal shape, see Figures S5-S13 in the SI.

Fitting the data to a function of the form $C\sin(\nu\theta) + D\sin(\eta\theta)$ reduces the sum of the mean-square error by a factor of 2-4. These values are also reported in Tables 5.1-5.2 for all molecules (except LiH). For the homonuclear molecules, the refitting only marginally changes the values of $C$ and $\nu$ (typically resulting in $\nu$ being closer to 2); $D$ is anywhere from 4 to 10 times smaller than $C$, and $\eta \approx 3.95$ for each molecule. These $\eta$ and $D$ values are consistent with the Kirkwood term arising from $2^{nd}$ nearest neighbor interactions, which we would expect to be the next largest contribution to the OR. For HF, while the error is reduced by the double sin fit, the $\nu$ and $\eta$ parameters obtained for COM and F axis rotation clearly do not align with the Kirkwood model prediction. Different initial guesses for the HF double sin fit parameters did not decrease the mean squared error, nor they significantly improve agreement with the Kirkwood model relative to the double sin fit in Tables 5.1 and 5.2. It is noteworthy that homonuclear $H_2$ and $F_2$ are described quite well with the Kirkwood model, but a heteronuclear combination of these elements generally does not agree with the model.

While the Kirkwood model provides a reasonable first-pass description of the OR, it is not a complete model even for homonuclear molecular units. In particular, angles close to 0° and 180° deviate from the fit, and the extrema of the fits (45° and 135°) do not match the extrema of the data (30° and 150°). While closer, the double sin fits' extrema also fail to match the data in the same regions. Since Kirkwood contributions all have the form $\sin(2k\theta)$, the model predicts that interacting groups rotated by 90° cannot contribute to the optical rotation, however, the atom axis helices for all molecules have a nonzero trace at 90°. Unlike rotation through the COM,

Table 5.1: Fitting parameters for single and double sin fits of Tr($\tilde{\mathbf{B}}$) for COM axis rotations. Parameters were determined via nonlinear regression to minimize the sum of the mean-squared errors with respect to the calculated values of Tr($\tilde{\mathbf{B}}$) for each helix.

| **COM Axis Helices** | **Single sin fit** | | **Double sin fit** | | | |
|---|---|---|---|---|---|---|
| Fitting Parameters | $C \times 10^2$ | $\nu$ | $C \times 10^2$ | $\nu$ | $D \times 10^2$ | $\eta$ |
| $H_2$ | −9.69 | 1.951 | −9.81 | 1.992 | −2.38 | 3.924 |
| $N_2$ | −21.9 | 1.980 | −22.0 | 1.998 | −2.35 | 3.943 |
| $F_2$ | −26.6 | 1.963 | −26.8 | 1.996 | −5.33 | 3.946 |
| HF | −0.932 | 1.743 | −0.935 | 1.346 | −0.601 | 2.338 |

Table 5.2: Fitting parameters for single and double sin fits of Tr($\tilde{\mathbf{B}}$) for atom axis rotations. Parameters were determined via nonlinear regression to minimize the sum of the mean-squared errors with respect to the calculated values of Tr($\tilde{\mathbf{B}}$) for each helix.

| **Atom Axis Helices** | **Single sin fit** | | **Double sin fit** | | | |
|---|---|---|---|---|---|---|
| Fitting Parameters | $C \times 10^2$ | $\nu$ | $C \times 10^2$ | $\nu$ | $D \times 10^2$ | $\eta$ |
| $H_2$ | −10.0 | 2.001 | −10.3 | 2.034 | −1.91 | 3.853 |
| $N_2$ | −17.9 | 1.946 | −18.3 | 1.979 | −2.67 | 3.706 |
| $F_2$ | −21.2 | 1.921 | −21.8 | 1.979 | −5.35 | 3.774 |
| HF(H) | −1.45 | 1.885 | −1.43 | 1.902 | −0.296 | 4.140 |
| HF(F) | −0.940 | 1.721 | −0.932 | 1.341 | −0.597 | 2.337 |

90° atom axis rotations leave the helix chiral, thus Tr($\tilde{\mathbf{B}}$) can only be zero accidentally in this case. Moreover, a description of the OR in terms of solely local ($k^{\text{th}}$ nearest neighbor for small $k$) interactions would not be consistent with the results in Figures 5.1-5.4, as this would cause the OR tensor components to reach their converged values sooner. The fact that the OR tensor takes around 25 units to converge suggests there must be some longer range effects at play that are not properly described by the Kirkwood model, and might be teased out from PBC calculations.

### 5.3.3 Basis sets

The aug-cc-pVDZ basis set is generally a good compromise between computational cost and accuracy, however, in developing a solid state OR procedure, it is important to verify that the conclusions we have drawn with this basis set are qualitatively consistent with the results of more complete basis sets. To this end, we compare aug-cc-pVDZ results for $H_2$ with calculations using aug-cc-pVTZ. In Figure 5.8, we plot the tensor contributions to/components of $\tilde{\mathbf{B}}$ using the aug-cc-pVTZ basis set, along with Tr($\tilde{\mathbf{B}}$) for the same $H_2$ helix as in Figure 5.1. Due to the cost associated with increasing the size of the basis set, results were only obtained for helices up to 100 molecules in length, which is sufficient to observe convergence of the tensor components. The tensor components obtained using the triple-$\zeta$ basis set exhibit similar qualitative behavior with respect to convergence as their double–$\zeta$ counterparts. The only noteworthy difference is that the $\tilde{\mathbf{B}}$ components shift to decrease their magnitude, leading to reduced optical response. This effect is essentially absent for the 90° helices, but increases as $\theta$ goes towards 0° or 180°.

One exception to the qualitative similarity of the OR tensors in each basis set is the 15° helix, which shows an unusual stair-step pattern in the OR tensor and its trace, see Figure 5.9. These sharp jumps seem to be an unphysical artifact of the basis set, as evidenced by the incongruity between the periodicity of the steps and any geometric feature of the helix. To probe what part of the basis set causes this behavior, we repeated the calculations by systematically removing diffuse functions from the aug-cc-pVTZ basis. We focused on diffuse functions because they have been shown to be more important for accurately computing the OR than other elements of the basis set.[112] We find

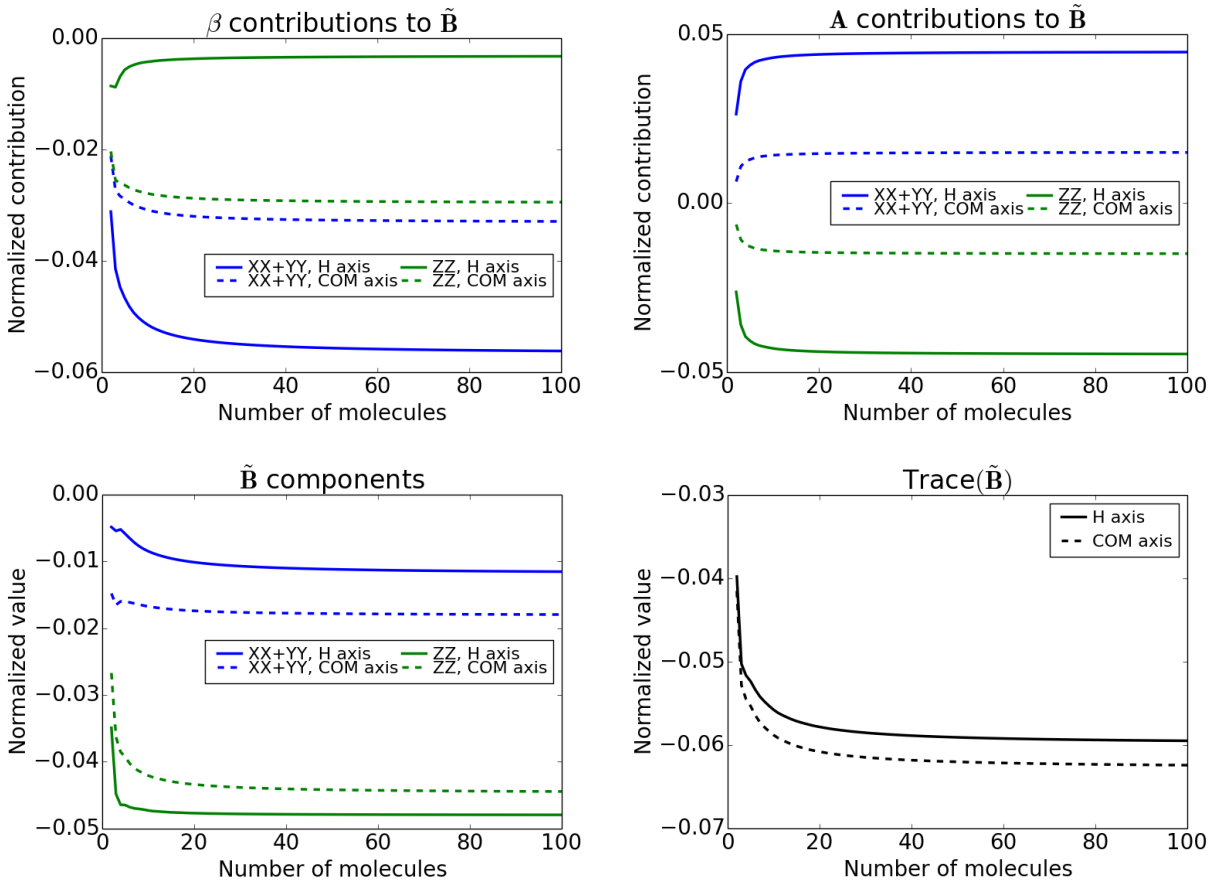


Figure 5.8: Plots of contributions to/components of $\tilde{\mathbf{B}}$ (a.u.) for an $H_2$ helix with 45° rotation about the H atom and COM axes. Calculations were carried out using aug-cc-pVTZ basis set. The separation between hydrogen molecules is 2.0 Å. Each value is scaled by the inverse of the number of units in the helix.

that removing the s-type diffuse functions from the basis eliminates the sharp jumps and changes the values of the tensor components. On the other hand, removing d-type diffuse functions has little effect on the component values, while removing the p-functions substantially alters the OR tensor, leading to poor agreement with larger basis set results (*vide infra*).

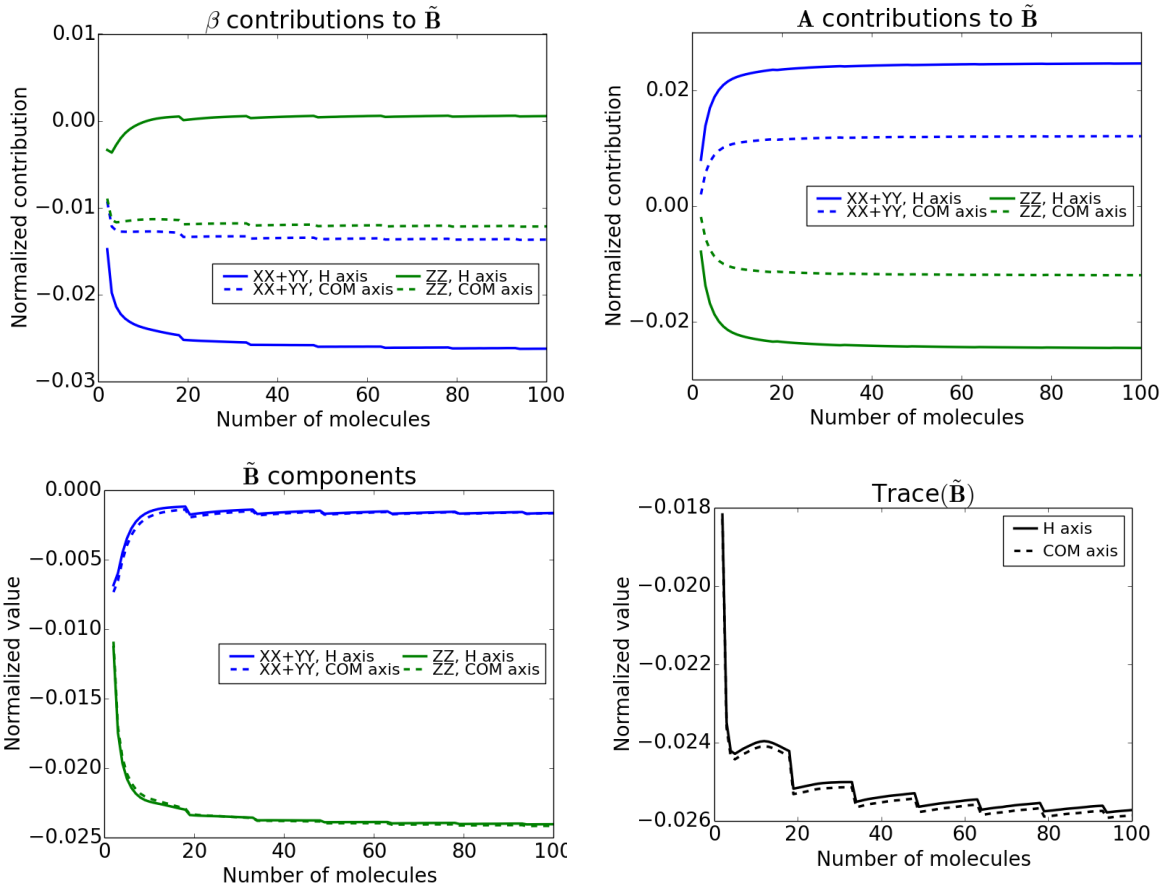


Figure 5.9: Plots of contributions to/components of $\tilde{\mathbf{B}}$ (a.u.) for an $H_2$ helix with 15° rotation about the H atom and COM axes. Calculations were carried out using aug-cc-pVTZ basis set. The separation between hydrogen molecules is 2.0 Å. Each value is scaled by the inverse of the number of units in the helix.

To see whether this unphysical effect is present for larger bases and to test the accuracy of these diffuse-removed basis sets, we recalculated the OR using aug-cc-pVQZ and aug-cc-pV5Z, along with their variants where some of the diffuse functions were removed. Figure 5.10 plots the trace of the OR tensor for a selection of these basis sets. We use the aug-cc-pV5Z results as the standard for accuracy, given that this is the largest basis set considered, although we note that

p-pV5Z (the same basis set, but with only the p-type diffuse functions) gives very similar results. There are still jumps in $\tilde{\mathbf{B}}$ for the aug-cc-pVQZ basis, however, they do not occur in the same locations as they did for the aug-cc-pVTZ basis, supporting their unphysical nature. Comparing the fully augmented basis functions to those with just p-type diffuse functions, the results show that not only does removing the s-type functions smooth out the curves of Trace($\tilde{\mathbf{B}}$), but it also brings their values closer in line with those of the aug-cc-pV5Z basis set. This is a promising result, as it suggests that the accuracy of a large basis set can be achieved using a smaller cc-pVXZ basis with fewer diffuse functions.

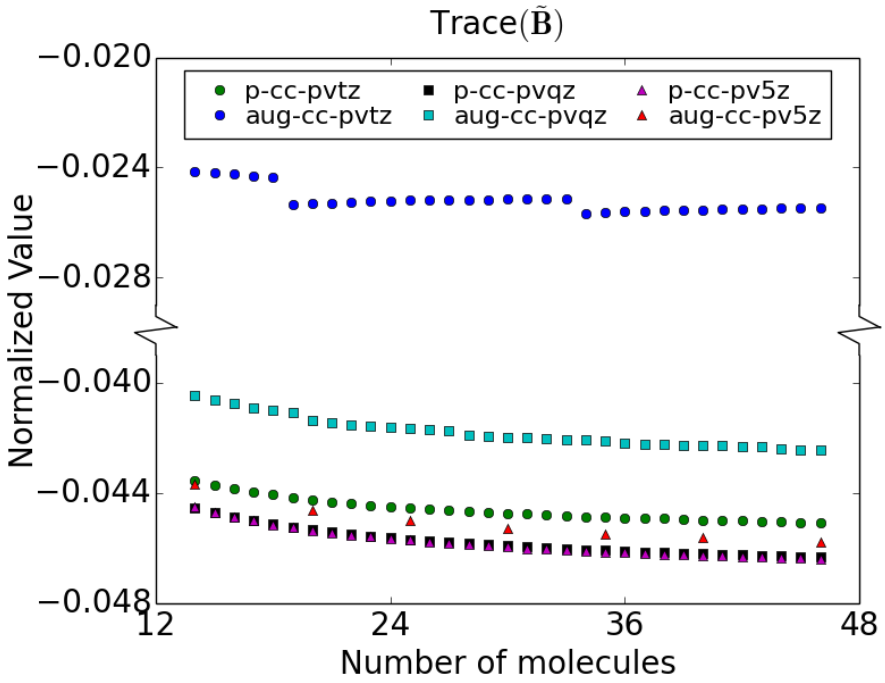


Figure 5.10: Plots of Trace($\tilde{\mathbf{B}}$) (a.u.) for an $H_2$ helix with 15° rotation about the COM axis. The separation between hydrogen molecules is 2.0 Å. Calculations were carried out with the aug-cc-pVXZ basis sets, X=T,Q,5, and with the corresponding p-cc-pVXZ basis sets containing only the p-type diffuse functions.

## 5.4 Discussion and Conclusions

This work provides benchmark data on a set of helical chains of diatomic molecules as a first step towards developing a general purpose electronic structure method to compute the OR of solids. We study the rate and extent of convergence of the OR tensor, $\tilde{\mathbf{B}}$, with respect to the length of the helix, and find that the values for unit cell sized clusters do not agree with the converged values for these helices. This lack of dependence on the unit cell size holds across different molecules and helix conformations, suggesting larger calculation cells may be needed, particularly for the more relevant 3D periodic compounds that we intend to study. The OR behavior of the helices of homonuclear molecular units with respect to the dihedral angle $\theta$ is in good agreement with the semiempirical Kirkwood polarizability model, indicating that nearest-neighbor interactions are the strongest contributions to the OR. At the same time, this simple model is not able to account for important long-range, cumulative interactions, and it does not describe heteronuclear helices very well. The difficulties of the Kirkwood model in describing even such simple systems reveal that approaches based on local polarizability may be insufficient to characterize chiral supramolecular assemblies. A comparison of the results for $H_2$ helices with different basis sets shows that the aug-cc-pVDZ results qualitatively agree with the results of the larger aug-cc-pVTZ basis set. Reasonable results are also obtained using basis sets with a reduced number of diffuse functions, allowing for more efficient computation of the OR.

The outcomes of this study suggest that the implementation of periodic boundary condition methods for the evaluation of the OR seems the best approach to study the optical activity of solids, and this work is currently under way in our group. Furthermore, the insight gained from these helix models will help us to tune various parameters of the PBC procedure to strike the right balance between accuracy and feasibility.

# Chapter 6

# Effect of Gauge on Computations of the Origin Invariant, Full Optical Rotation Tensor

(This work was adapted from *Chirality* **2021**, *33*, 303-314[70] and *J. Chem. Phys.* **2021**,*155* 024118[71] with permission from the authors. Supporting information is available online.)

## 6.1 Introduction

Although most measurements and simulations of optical rotation focus on isotropic systems, e.g. in solution and in gas phase, it is also interesting to investigate oriented systems. In fact, for the latter it is possible to measure each individual element of the tensor, which can provide more information about the properties of the system.[113,114] From the experimental point of view, such measurements can only be performed on crystals because of the limited intensity of the signal. These solid-state measurements are still difficult because they require a very smooth surface to distinguish the optical rotation signal (circular birefringence) from more intense signals (linear birefringence), and only a limited amount of experimental data is available.[113,114]

Theoretical simulations are of paramount importance for the correct assignment of the absolute configuration of chiral molecules, and many methods have been developed mostly based on density functional theory (DFT) and coupled cluster (CC) theory.[26,60,72,90,97,109,111,115–123] However, the majority of these studies focus on the isotropic OR because these methods were developed for molecules, not solid materials, and because of the predominance of experimental data in gas and solution phase. Thus, few studies report the full OR tensor.[36,51,52,54,124] In Chapter 5, we presented results for calculations of the full OR tensor at the DFT level for models of solid state

systems. There, we noted the importance of both the electric dipole-magnetic dipole and electric dipole-electric quadrupole polarizability contributions to the individual components of the full OR tensor.

While isotropic OR is simpler to compute, computing the full OR tensor can reduce numerical error, allowing for clearer comparisons with experiment. This is because the isotropic specific rotation, as shown in Eq. 2.6, is computed as the trace of the full OR tensor, which is equal to the trace of the electric dipole-magnetic dipole polarizability tensor. In other words, the electric dipole-electric quadrupole polarizability tensor contribution to the full OR tensor is traceless and does not need to be evaluated. Since the diagonal elements of the OR tensor are signed quantities, their sum is consistently one or two orders of magnitude smaller than their individual values. This means that a relatively small error in each element may lead to a significantly larger error in the trace and thus in the isotropic OR. Therefore, it is desirable to be able to evaluate the full OR tensor using highly accurate levels of theory despite the extra computational effort.

Since practical electronic QM calculations only provide an approximate solution to the Schrödinger equation, using an incomplete basis expansion for the electron density, the numerical results depend on the choice of gauge for the electric dipole and quadrupole operators.[72,89,125,126] As we noted in section 2.1, length gauge (LG) calculations using variational methods such as Hartree-Fock (HF) and DFT, the origin-dependence issue is resolved using London orbitals, also known as gauge including atomic orbitals (GIAOs).[61,90,115,127,128] However, GIAOs cannot be utilized with standard CC methods, where the reference orbitals are not reoptimized, because orbital relaxation is neglected in the linear response (LR) equations to avoid unphysical poles in the LR function due to the reference wave function.[73,129] Alternative formulations of CC methods with orbital optimization have been proposed,[130,131] but they are not commonly used because they are computationally intensive and often have convergence issues. On the other hand, the modified velocity (MVG)[60,70] recipe requires the explicit evaluation and removal of the unphysical static limit, and so far it has been the preferred approach for OR calculations at CC level.[72,97,132,133]

Caricato recently proposed the so called LG(OI) method to overcome the origin-dependence

issue of LG calculations of specific rotation without the complication of London orbitals.[62] This is based on a transformation of the electric dipole-magnetic dipole polarizability tensor using the singular value decomposition (SVD) eigenvectors of the mixed-gauge electric dipole-electric dipole polarizability tensor. The LG(OI) approach is simpler than the LG-GIAOs approach and faster than the MVG approach, but it shares with the latter the applicability to any approximate method. This allows for the first origin-invariant LG simulations of specific rotation with standard CC methods.

In this chapter, we compare the calculation of the full, origin independent Buckingham/Dunn tensor at single and double excitation level (CCSD) using both the MVG and LG(OI) approaches to determine the effect of the gauge on the computed OR. Although coupled cluster calculations of OR are only reasonable for individual molecules at this point, this is an important step towards the development of the corresponding method for solids using periodic boundary conditions, which in turn can be compared directly to experimental measurements. Here, we compare the CCSD-LG(OI) and MVG approaches on a test set of 22 organic molecules to assess whether they provide physically equivalent results. We also compute the MVG $\boldsymbol{\mathcal{B}}$ tensor from Eq. 2.7 for these molecules using two common DFT functionals for OR calculations (B3LYP and CAM-B3LYP) in order to assess how the treatment of electron correlation affects the OR. The chapter is organized as follows: a brief review of the LG(OI) approach is presented in section 6.2, details of the calculations are reported in section 6.3, the results of numerical simulations are discussed in section 6.4, and concluding remarks are summarized in section 6.5.

## 6.2 Theory

In this section, we present the theory for the evaluation of the full OR tensor with the LG(OI) approach.[62] Throughout, we reference the equations for the Buckingham/Dunn optical activity tensor $\boldsymbol{\mathcal{B}}$ given in Chapter 2. The tensors $\boldsymbol{\beta}^L$ and $\mathbf{A}^L$ in Eqs. 2.2-2.3 are both origin dependent according to:

$$\beta^L_{\alpha\beta}(\boldsymbol{O}') = \beta^L_{\alpha\beta}(\boldsymbol{O}) + \frac{1}{2}\varepsilon_{\beta\gamma\delta}\alpha^{(R,P)}_{\alpha\gamma} d_\delta \tag{6.1}$$

$$A^{L}_{\alpha,\beta\gamma}(\boldsymbol{O}') = A^{L}_{\alpha,\beta\gamma}(\boldsymbol{O}) - \frac{3}{2}\alpha^{(R,R)}_{\alpha\gamma} d_{\beta} - \frac{3}{2}\alpha^{(R,R)}_{\alpha\beta} d_{\gamma} + \alpha^{(R,R)}_{\alpha\delta} d_{\delta}\delta_{\beta\gamma} \tag{6.2}$$

where a sum over common indices is implied in the last term on the right-hand side of both equations, $\boldsymbol{O}$ is a particular choice of origin of the coordinate system and

$$\boldsymbol{O}' = \boldsymbol{O} + \boldsymbol{d} \tag{6.3}$$

is a displaced origin, $\boldsymbol{\delta}$ is the Kronecker delta tensor, and $\boldsymbol{\alpha}$ is the electric dipole-electric dipole polarizability tensor expressed with two gauge representations for the dipole operator: the superscript $(R,R)$ in Eq. 6.2 indicates the length gauge representation for both occurrences of the dipole operator (shown in Eq. 2.12), while the superscript $(R,P)$ in Eq. 6.1 indicates a mixed representation with the length gauge for one occurrence of the dipole operator and the velocity gauge for the other one (note that a factor of $\omega^{-1}$ is to be included in the mixed-gauge LR function to account for the change in representation of the electric dipole operator).[60] In an exact calculation, the individual origin dependent terms of $\boldsymbol{\beta}^{L}$ and $\mathbf{A}^{L}$ perfectly cancel out when they are combined to form $\boldsymbol{\mathcal{B}}$ in Eq. 2.7, such that the latter is origin invariant. However, this is no longer the case in approximate calculations and the individual elements of $\boldsymbol{\mathcal{B}}$ as well as its trace are origin dependent.

In the velocity gauge, $\boldsymbol{\mathcal{B}}$ is inherently origin independent because the origin dependent terms for both $\boldsymbol{\beta}^{V}$ and $\mathbf{A}^{V}$ are related to $\boldsymbol{\alpha}^{(R,R)}$. Since $\boldsymbol{\alpha}^{(R,R)}$ is symmetric even in an approximate calculation, these terms cancel when the two tensors are added together. We can use this same idea to construct an origin-invariant $\boldsymbol{\mathcal{B}}$ tensor in the length gauge for any approximate method. The approach is based on the same transformation we suggested for the specific rotation in Ref. 62, which we called LG(OI). First, we need a form of the $\mathbf{A}$ tensor that has a qualitatively similar origin dependence as the $\boldsymbol{\beta}$ vector in Eq. 6.1, i.e. dependent on the mixed-gauge electric dipole-electric dipole polarizability $\boldsymbol{\alpha}^{(R,P)}$. This is accomplished by contracting the length electric dipole operator with the velocity quadrupole operator:

$$A^{(R,P)}_{\alpha,\beta\gamma} = \frac{2}{\omega}\sum_{j\neq 0}\mathrm{Im}\frac{\langle\psi_0|\mu^L_\alpha|\psi_j\rangle\langle\psi_j|\Theta^V_{\beta\gamma}|\psi_0\rangle}{\omega_j^2-\omega^2} \tag{6.4}$$

such that the **A** tensor uses a mixed-gauge representation: length for the dipole operator and velocity for the quadrupole operator. This form of the **A** tensor preserves the expression of the $\boldsymbol{\mathcal{B}}$ tensor in Eq. 2.1, but leads to the following expression of its origin dependence:

$$\begin{aligned} A^{(R,P)}_{\alpha,\beta\gamma}(\boldsymbol{O}') =& A^{(R,P)}_{\alpha,\beta\gamma}(\boldsymbol{O}) - \frac{3}{2}\alpha^{(R,P)}_{\alpha\gamma}d_\beta \\ & - \frac{3}{2}\alpha^{(R,P)}_{\alpha\beta}d_\gamma + \alpha^{(R,P)}_{\alpha\delta}d_\delta\delta_{\beta\gamma} \end{aligned} \tag{6.5}$$

which is now compatible with that of the $\boldsymbol{\beta}$ tensor in Eq. 6.1. However, the use of the mixed-gauge form of the **A** tensor is not sufficient to lead to an origin-invariant $\boldsymbol{\mathcal{B}}$ tensor because the $\boldsymbol{\alpha}^{(R,P)}$ tensor is not symmetric. Therefore, the contributions of the $\boldsymbol{\beta}$ and $\mathbf{A}^{(R,P)}$ tensors to the $\boldsymbol{\mathcal{B}}$ tensor do not perfectly cancel out. To achieve that, we diagonalize $\boldsymbol{\alpha}^{(R,P)}$ tensor with a singular value decomposition (SVD):

$$\boldsymbol{\alpha}^{(R,P)} = \boldsymbol{U}\,\boldsymbol{\alpha}^{(R,P)}_D\boldsymbol{V}^\dagger \tag{6.6}$$

where $\boldsymbol{\alpha}^{(R,P)}_D$ is diagonal, and $\boldsymbol{U}$ and $\boldsymbol{V}^\dagger$ are unitary transformations. We then apply the inverse transformation to the $\boldsymbol{\beta}$ and $\mathbf{A}^{(R,P)}$ tensors:

$$\tilde{\beta}_{\alpha\beta} = \beta_{\alpha'\beta'}U_{\alpha'\alpha}V_{\beta'\beta} \tag{6.7}$$

$$\tilde{A}^{(R,P)}_{\alpha,\beta\gamma} = A^{(R,P)}_{\alpha',\beta'\gamma'}U_{\alpha'\alpha}V_{\beta'\beta}V_{\gamma'\gamma} \tag{6.8}$$

Using the $\tilde{\boldsymbol{\beta}}$ and $\tilde{\mathbf{A}}^{(R,P)}$ tensors in Eq. 2.1 makes the $\boldsymbol{\mathcal{B}}$ tensor fully origin invariant without recourse to GIAOs. The transformations in Eqs. 6.7-6.8 are sensitive to the phase of the eigenvectors and their order in the unitary matrices. Therefore, it is important to avoid inadvertently changing the handedness of the coordinate system with the transformation.

A legitimate question about the LG(OI) $\boldsymbol{\mathcal{B}}$ tensor is what orientation of the molecule corresponds to this tensor values. In an exact calculation, $\boldsymbol{\alpha}^{(R,P)}$ is symmetric and $\boldsymbol{U} \equiv \boldsymbol{V}$ in Eq. 6.6.

Therefore, the transformed $\mathcal{B}$ tensor corresponds to an orientation where the $\alpha^{(R,P)}$ tensor is diagonal. For approximate calculations, the closest we can get to this orientation is to use the eigenvectors that diagonalize the symmetric part of the $\alpha^{(R,P)}$ tensor, which we collect in the unitary matrix $\boldsymbol{W}$. As we approach the exact solution $\boldsymbol{U},\boldsymbol{V} \to \boldsymbol{W}$. When comparing LG(OI) results with other gauges, the orientation defined by $\boldsymbol{W}$ should be used so that the $\mathcal{B}$ tensor components of all calculations are aligned properly.

## 6.3 Computational Details

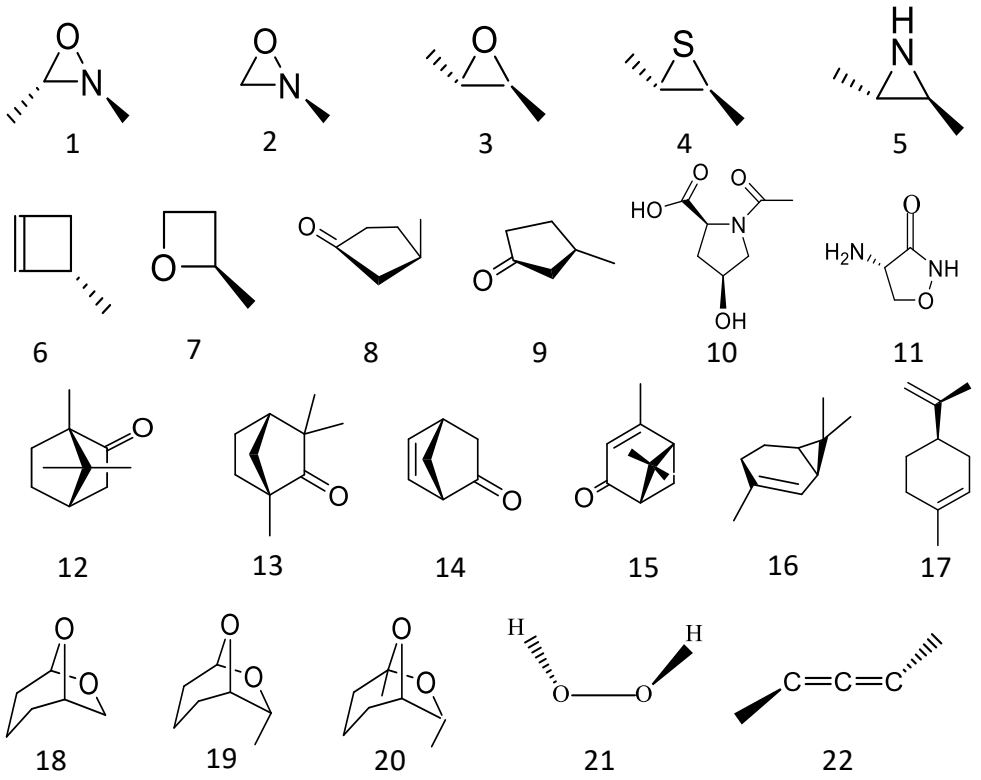


Figure 6.1: Structures of the molecules in the test set.

The molecules used in the test set are reported in Figure 6.1. The geometries were taken from Stephens et al.[112] for molecules **1** and **2**, from Srebro et al.[134] for molecules **3**-**7**, **12**, **13**, **15**, and **18**-**20**, from Aharon and Caricato[41] for molecules **8**-**11**, **16**, and **17**, and from Ref. 62 for molecules **14**, **21**, and **22**. The $\mathcal{B}$ calculations were performed as discussed in the previous section at the

sodium D line (589.3 nm) because this is a typical wavelength used in experiment and it is far from resonance for all molecules. The OR calculations were performed with three methods: CCSD, B3LYP,[94,135] and CAM-B3LYP.[108] The former functional was chosen as the most popular choice in quantum chemistry, and the latter because it showed very good performance in specific rotation calculations.[134] The aug-cc-pVDZ[93] basis set was used for the OR calculations with all methods because it represents a reasonable cost-accuracy compromise. All calculations were performed with a development version of the GAUSSIAN suite of programs.[95] The CCSD tensors were computed using standard LR theory, with frozen core orbitals for the correlation energy evaluation and frozen orbitals for the LR function.

In comparing the CCSD and DFT MVG calculations, we assume that the CCSD results can be used as reference, and errors are defined with respect to this reference. Obviously, there is no guarantee that CCSD provides the results most in agreement with hypothetical experimental measurements. However, since CC theory provides a systematically improvable hierarchy of methods and CCSD is the best we can do at the moment, we consider it the "theoretical reference". Furthermore, in order to simplify the comparison between methods, the molecules were oriented along the principal axes of their CCSD-MVG $\boldsymbol{\mathcal{B}}$ tensor, and the rotated structures are reported in Tables S1-S22 of the supporting information (SI).[70] In this way, the off-diagonal elements of the $\boldsymbol{\mathcal{B}}$ tensor are zero and the corresponding values for the DFT-MVG tensor should be rather small. This allows us to concentrate our analysis on the diagonal elements of the $\boldsymbol{\mathcal{B}}$ tensor and evaluate absolute errors for the off-diagonal elements. Further analysis of the off-diagonal elements of $\boldsymbol{\mathcal{B}}$, as well the individual components of the $\boldsymbol{\beta}$ and $\mathbf{A}$ tensors, is described elsewhere.[70] For the MVG and LG(OI) comparisons, we retained the geometries from the original references and simply rotated the MVG $\boldsymbol{\mathcal{B}}$ tensor directly using the $\boldsymbol{W}$ rotation matrix.

All of these tensors are extensive quantities, but it is better to compare intensive quantities like the specific rotation in Eq. 2.6. Therefore, the correlation plots in section 6.4 use mass-scaled tensor values:

$$T^{M}_{\alpha\beta} = \frac{T_{\alpha\beta}}{M} \times 10^{3} \tag{6.9}$$

where $T_{\alpha\beta}$ is an element of tensor **T** in a.u., $M$ is the molar mass of the molecule in amu, and $10^3$ is a scaling factor used for clarity. The $T_{\alpha\beta}$ and $M$ values, together with the isotropic specific rotation of all molecules, are reported in Tables S23-S32 of the SI.[70] The correlation plots also include linear fits of the data with a zero-intercept constraint (an achiral molecule has zero trace with all methods). The fits are evaluated with and without molecule **14**, (1S,4S)-(–)-norbornenone, because this is a notoriously difficult molecule for optical rotation calculations.[31,136,137]

## 6.4 Results

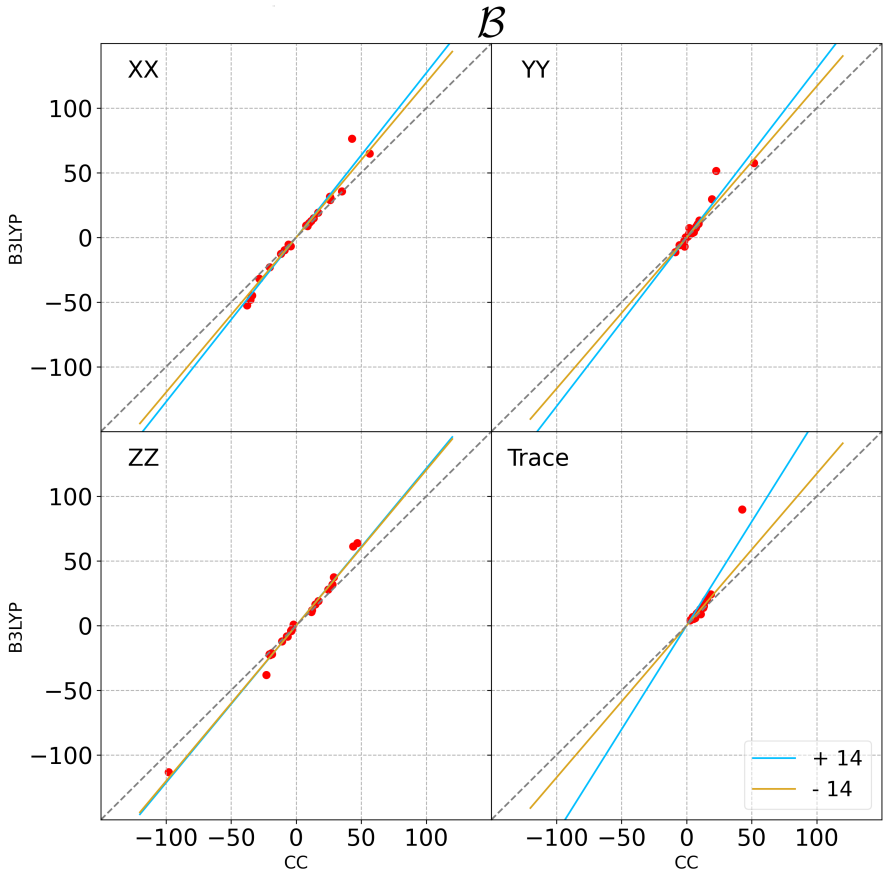


Figure 6.2: B3LYP vs CCSD correlation plots for the mass-normalized diagonal elements and trace (a.u. $\times 10^3$) of the MVG $\mathcal{B}$ tensor, see Eq. 6.9. The linear fit was performed with (cyan) and without (yellow) molecule **14**.

We start by comparing MVG CCSD and DFT calculations on the test set in Figure 6.1 in order to determine the effect of electron correlation on OR. For this comparison, the molecules are

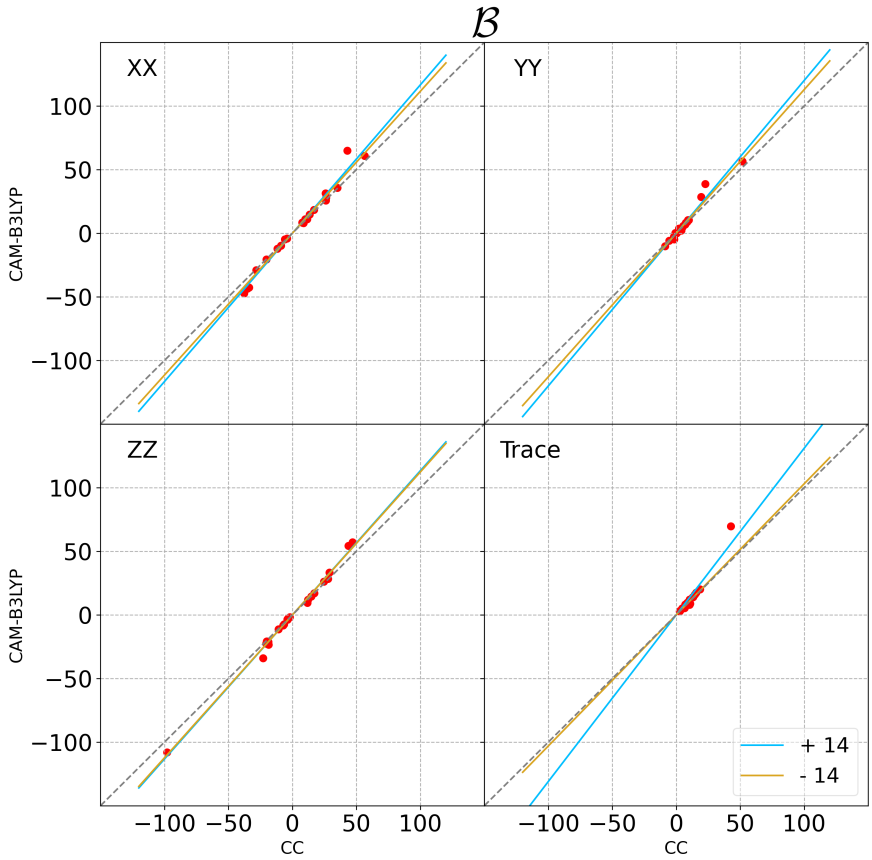


Figure 6.3: CAM-B3LYP vs CCSD correlation plots for the mass-normalized elements and trace (a.u.$\times 10^3$) of the $\boldsymbol{\mathcal{B}}$ tensor computed at the CCSD/aug-cc-pVDZ level for the molecules in Figure 6.1. The linear fit was performed with (cyan) and without (yellow) molecule **14**.

oriented such that the CCSD $\mathcal{B}$ tensor is diagonal so that we can directly assess the correlation of the diagonal elements, as shown in Figures 6.2 and 6.3. In these plots, the tensor elements and the trace are divided by the molecular mass to produce an intensive quantity, which allows us to compare molecules of various sizes on equal footing. Linear fits on the whole test set were noticeably affected by molecule **14** (i.e., the notorious (S)-(–)-norbornenone), so we also present fits that exclude this molecule.

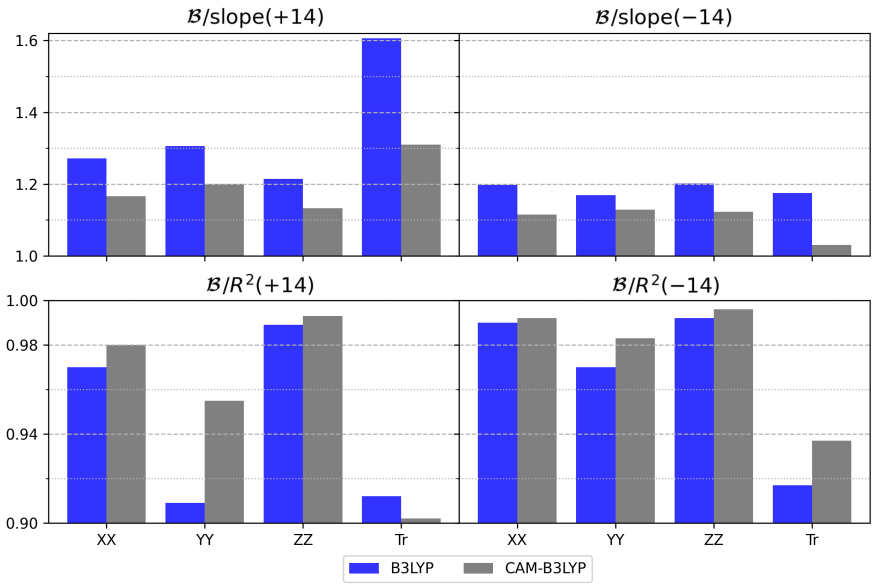


Figure 6.4: Slope and $R^2$ correlation parameters of the linear fits for the diagonal elements of the $\mathcal{B}$ tensor including (left-hand side plots) and excluding (right-hand side plots) molecule **14**.

Starting with the B3LYP data in Figures 6.2 and 6.4, while this functional consistently overestimates the magnitude of the $\mathcal{B}$ components and trace compared to CCSD (slopes greater than 1.2), the correlation between the two methods is strong ($R^2 \geq 0.90$ for all components and the trace). As previously mentioned, molecule **14** has an outsized influence on the fit. Removing this point improves the correlation, particularly for the YY component, but still results in a roughly 20% overestimation of CCSD results. The individual tensor elements ($T^M_{\alpha\beta}$ in Eq. 6.9) are spread out over a large range of positive and negative values, but the spread of their sum (the Trace plot) is significantly smaller, which confirms the notion that in general the isotropic OR is considerably smaller than what could be measured for an oriented system. The reduced magnitude of the trace

makes it susceptible to significant error propagation stemming from (relatively) small errors in the diagonal components. This error propagation is clearly visible when comparing the fits with and without molecule **14**; removing this outlier reduces the average overestimation of the trace from 60% to 20% while the overestimation of the individual tensor elements is reduced from roughly 30% down to 20%.

The results for CAM-B3LYP are qualitatively the same as those for B3LYP, but the agreement with the CCSD results is better for the former functional, with a slope smaller than 1.2 for all components and the trace. Excluding molecule **14**, the overestimation of the trace is reduced considerably, with a slope approaching 1. The $R^2$ of the trace remains lower than that of the components whether or not molecule **14** is included, again reflecting the significant effect of error propagation.

The LG(OI) and MVG $\mathcal{B}$ tensor elements and trace for the entire set of 22 molecules are compared in the correlation plots in Figure 6.5. As with the DFT/CCSD comparison, we performed linear fits on the data both with and without molecule **14**, and the slope and $R^2$ parameters for these fits are plotted in Figure 6.6. Since the molecules are oriented according to the LG(OI) rotation matrix **W** for this comparison, rather than the principal axes of the MVG $\mathcal{B}$ tensor, we also consider the correlation of off-diagonal elements of the tensor. It is clear from these figures that there is a strong linear correlation between the MVG and LG(OI) $\mathcal{B}$ tensor components, with the $R^2$ of each fit approaching 1.0. This strong correlation persists regardless of whether molecule **14** is included in the fit; however, this molecule has an outsized influence on the predicted slope. While the fit with molecule **14** suggests that MVG $\mathcal{B}$ components consistently underestimate the corresponding LG(OI) tensor elements by 10-20%, the fit excluding **14** predicts a consistent overestimation relative to LG(OI) for all components. In this case, the overestimation is smaller in magnitude ($< 5\%$) except for the yy ($\sim 10\%$) and yz (15%) elements. The MVG trace is consistently smaller than that with LG(OI), but again removing molecule **14** from the fit reduces this underestimation to only about 5%. The $R^2$ parameters are all $> 0.95$ and all but the yy component increase when molecule **14** is excluded from the fitting. The MVG yy elements for molecules **21** and **22** signif-

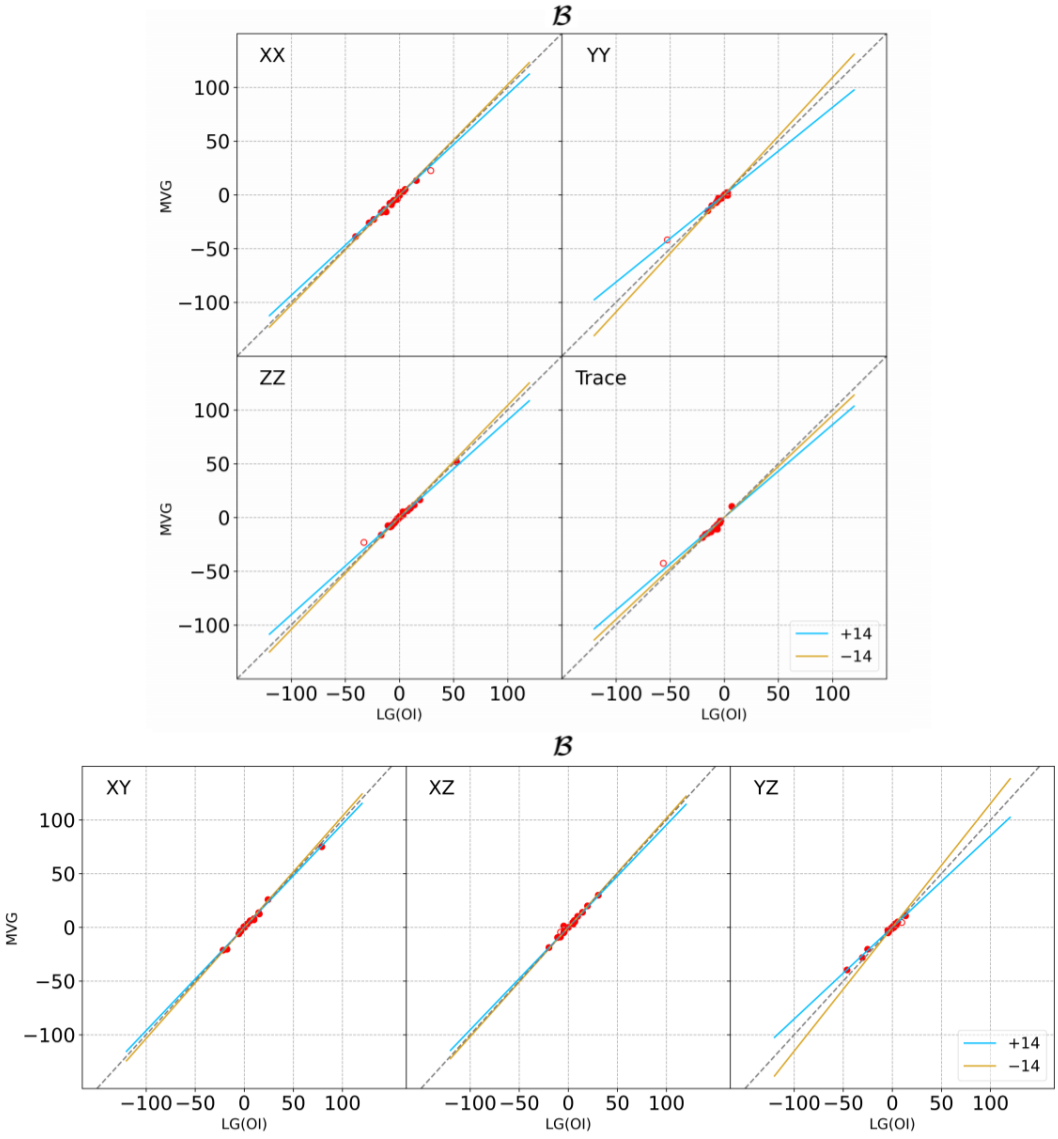


Figure 6.5: MVG vs LG(OI) correlation plots for the mass-normalized elements and trace (a.u.$\times 10^3$) of the $\mathcal{B}$ tensor computed at the CCSD/aug-cc-pVDZ level for the molecules in Figure 6.1. The linear fits were performed with (cyan) and without (yellow) molecule **14**. Molecule **14** is denoted by the unfilled marker in each of the plots.

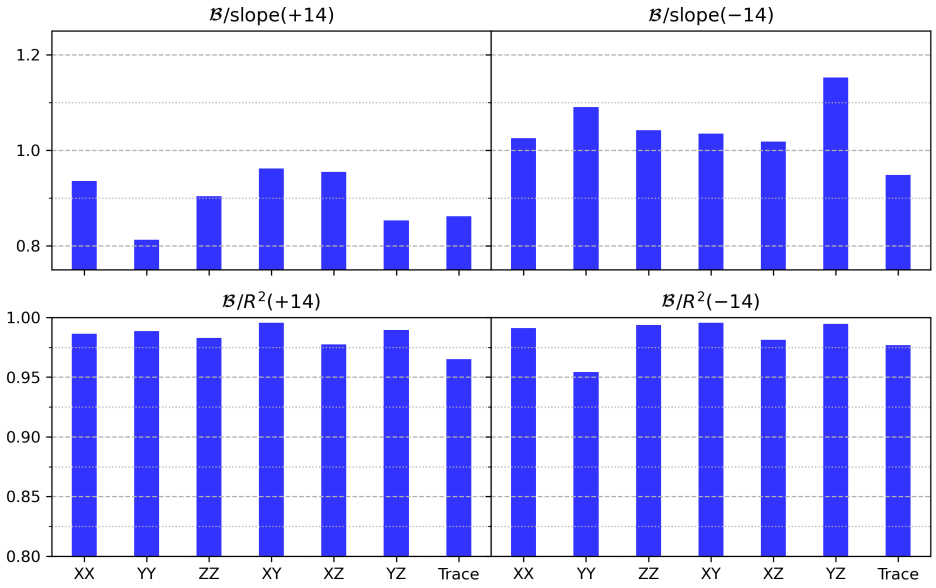


Figure 6.6: Slope and $R^2$ correlation parameters of the MVG vs LG(OI) linear fits of $\mathcal{B}$ components including (left) and excluding (right) molecule **14**.

icantly underestimate their LG(OI) counterparts, with the component for **22** even changing sign. This underestimation leads to a slightly larger $R^2$ in the fit with 14, since the slope is < 1.0 in this case. The large effect of including/excluding molecule **14** suggests that this level of theory and basis set are somewhat far from convergence for this compound, as MVG and LG(OI) should be equivalent for an exact, complete basis calculation.

Despite the general accuracy of the fits, there are still individual cases for which the MVG and LG(OI) CCSD $\mathcal{B}$ tensor elements differ qualitatively: molecules **7** (xx), **8** (yz), **14** (xy), **15** (xz), and **21** (yy) all have one MVG $\mathcal{B}$ element that differs in sign from the LG(OI) $\mathcal{B}$ tensor. These sign changes are not specific to any particular Cartesian component and tend to occur in cases where the absolute value of the LG(OI) $\mathcal{B}$ elements are small (< 0.10 a.u.). These small values lead to relatively small squared errors, so the $R^2$ parameters of the fits remain close to 1.0. Nevertheless, it is not obvious whether these deviations stem from incompleteness of the basis or insufficient treatment of correlation.

## 6.5 Discussion and Conclusions

In this work, we assess the influence of gauge and level of theory on computations of the full OR tensor $\mathcal{B}$. We found that MVG DFT calculations of the OR tensor $\mathcal{B}$ were well correlated with MVG CCSD calculations, though they consistently overestimate individual components of the tensor, as well its trace. Depending on the chosen functional, the extent of overestimation can vary greatly; excluding molecule **14**, traces computed with CAM-B3LYP results are in remarkable agreement with CCSD, while even without this problematic molecule, B3LYP overestimates the trace by an average of 30%. The outsized influence of molecule **14** on fits of the trace demonstrates the impact that error propagation from the (typically much larger) tensor components can have on this property.

Figures 6.5 and 6.6 show a clear linear correlation between the mass-scaled MVG and LG(OI) CCSD tensor elements for a wider range of molecules. However, there remain unexplained discrepancies between the two gauges. Large component differences, like those seen for the yy element of molecule **14** in Figure 6.5, can have a pronounced effect on the obtained linear fits for these components, even when the sign is consistent. It is unclear whether the deviations of these components are solely due to the small basis set used for these calculations (aug-cc-pVDZ) or if the level of theory plays a role.

A recent study by our group analyzed the effect of increasing basis set size on the computed specific rotation and found that MVG and LG(OI) CCSD did not converge to the same complete basis limit.[138] This discrepancy would suggest that the treatment of electron correlation in CCSD is insufficient, as the OR tensor should be invariant to the choice of gauge for an exact calculation. Including triple excitation contributions using CCSDT is cost-prohibitive even for small molecules and may still be insufficiently converged. Consequently, some care must taken with the choice of gauge in practical, approximate calculations. LG(OI) OR may be the more cost-effective choice, as it seems to converge more quickly to the complete basis set limit than MVG.[138]

In summary, we assessed various approaches to express the Buckingham/Dunn tensor in an origin-invariant formulation. We found that the $\mathcal{B}$ tensor computed with DFT methods was gen-

erally well correlated with that of CCSD, but that it typically overestimated the components by 20-30% and could deviate substantially for challenging cases like molecule **14**. $\mathcal{B}$ computed with the CCSD LG(OI) and MVG approaches were highly correlated, with only a few noticeable discrepancies in individual tensor components. While we have only analyzed molecular systems here, insight into different methods of computing the full $\mathcal{B}$ tensor will aid in the evaluation of optical rotation in oriented systems such as chiral crystals.

# Chapter 7

# Full Optical Rotation Tensor for Periodic Systems

## 7.1 Introduction

Measurements of optical rotation (OR) have long been used to characterize chiral systems, but it was only relatively recently that the first Hartree-Fock level simulations were performed to compute this property.[26] Great strides have been made towards more accurate, efficient, and interpretable OR simulations since these initial calculations: extensions to post-Hartree-Fock levels of theory,[62,111,112] development of OR-optimized basis sets,[91,132,139,140] improved understanding of the role of solvation,[41–44,141] and approaches for decomposing the OR into more chemically intuitive contributions.[31,34,35,38,39] In spite of these efforts, OR remains a very challenging property to compute accurately and a clear structure-property relationship remains unknown.

Many of the current difficulties related to computing OR stem from comparisons made with solution phase experiments. Accurately replicating the experimental OR in solution is not as simple as just determining the OR for the minimum energy structure; multiple configurations of the molecule can exist in solution, requiring a conformational search and subsequent Boltzmann averaging of the OR from each conformer. Not only does this require an extensive, costly search to ensure that all relevant conformers are found, but the energies of each conformer must also be determined to a high accuracy, as any errors in the exponential energy-weighting could greatly distort the conformer populations and thus the average OR. Compounding these issues is the need to model the surrounding solvent, which can influence the conformer populations and contribute to the OR itself through the formation of a chiral shell around a given conformer. Implicit electrostatic models of solvation account for some of these effects, but neglect specific solvent-solute

interactions and the induced chirality of the solvation shell. Incorporating some explicit solvent molecules has mixed results, as the decision of how many solvent molecules to include and where to place them is typically made in an *ad hoc* fashion. Including a full solvation shell would be less arbitrary, but would require MD simulations to determine reasonable solvation shell configurations and the levels of theory that can be used for these system sizes may not be sufficiently accurate.[42–44] Larger extents of solvation currently preclude the use of QM calculations altogether.

Comparisons with crystalline solids, particularly molecular crystals, would eliminate or greatly simplify a number of these issues. The rigidity and regularity of a crystal would not only limit the number of conformations, but would also simplify the "solvent" environment. The extended structure around a molecule can be determined from the crystal structure and can be readily simulated through the use of periodic boundary conditions (PBC). Additionally, solid state experiments allow the entire OR tensor to be resolved, unlike solution or gas phase where only the isotropic OR can be measured. However, general, first-principles methods to compute optical activity of periodic systems had not been developed until very recently.[46–48,142–145] Scuseria et al. laid the groundwork by deriving an approach to compute the PBC electric dipole polarizability tensor using an AO basis.[68] Rerát and Kirtman recently implemented the PBC electric dipole-magnetic dipole polarizability, which together with contributions from the electric dipole-electric quadrupole polarizability forms the full OR tensor, into the CRYSTAL electronic structure platform.[50]

In this chapter, we derive equations for the PBC electric dipole-magnetic dipole polarizability, as well as the electric dipole-electric quadrupole polarizability, and describe our work implementing these tensors into Gaussian in order to build the full periodic OR tensor. In section 7.2, we derive expressions for the electric dipole-magnetic dipole and electric dipole-electric quadrupole polarizabilities that satisfy periodic boundary conditions (PBC). In section 7.3, we discuss some details of how these OR polarizabilities are implemented in GAUSSIAN. Section 7.4 presents preliminary calculations on simple model systems using our current implementation of the PBC OR tensor. We conclude the chapter with a discussion of these results and their implications in section 7.5.

## 7.2 Theory

Optical activity is governed by the rank-2 Buckingham/Dunn **B** tensor, which in turn has contributions from $\boldsymbol{\beta}$ and **A**, respectively the frequency dependent electric dipole-magnetic dipole and electric dipole-electric quadrupole polarizability tensors as detailed in Chapter 2. Extending these tensors to periodic systems reduces to the problem of defining $\boldsymbol{\mu}$, $\boldsymbol{m}$, and $\boldsymbol{\Theta}$ operators that are translation invariant (i.e. consistent with periodic boundary conditions) that can be evaluated over an AO basis. For the electric dipole, this has already been derived and we reproduce the key steps of the derivation in section 2.3.[68,77] We can proceed in a similar fashion to derive the PBC magnetic dipole and electric quadrupole operators.

### 7.2.1 Periodic Magnetic Dipole

In real space, the matrix elements of the magnetic dipole for a periodic system can be written as

$$\boldsymbol{m}^{\mathbf{0}\boldsymbol{l}}_{\mu\nu} = -\langle \mu_{\mathbf{0}} | \boldsymbol{r} \times \boldsymbol{p} | \nu_{\boldsymbol{l}} \rangle = -\langle \mu_{\mathbf{0}} | (\boldsymbol{r} \times \boldsymbol{p}) e^{-i\boldsymbol{D}\boldsymbol{l}\cdot\boldsymbol{p}} | \nu_{\mathbf{0}} \rangle \tag{7.1}$$

where, as in section 2.3, $\mu, \nu$ are atomic orbitals indices, $\boldsymbol{D}$ is a matrix with (in general) three translation vectors as columns, $\boldsymbol{l}$ is a vector of indexing the cells the along each translation vector, with $\mathbf{0}$ the center cell, and $e^{-i\boldsymbol{D}\boldsymbol{l}\cdot\boldsymbol{p}}$ is the translation operator. Due to the presence of the translation operator, the magnetic dipole matrix is not Hermitian, i.e. $\boldsymbol{m}^{\mathbf{0}\boldsymbol{l}}_{\mu\nu} \neq \boldsymbol{m}^{\mathbf{0}-\boldsymbol{l}}_{\nu\mu}$:

$$\begin{aligned}
\boldsymbol{m}^{\mathbf{0}\boldsymbol{l}}_{\mu\nu} &= -\langle \mu_{\mathbf{0}} | (\boldsymbol{r} \times \boldsymbol{p}) e^{-i\boldsymbol{D}\boldsymbol{l}\cdot\boldsymbol{p}} | \nu_{\mathbf{0}} \rangle \\
&= -\langle \mu_{-\boldsymbol{l}} | \boldsymbol{r} \times \boldsymbol{p} | \nu_{\mathbf{0}} \rangle - \langle \mu_{\mathbf{0}} | [\boldsymbol{r}, e^{-i\boldsymbol{D}\boldsymbol{l}\cdot\boldsymbol{p}}] \times \boldsymbol{p} | \nu_{\mathbf{0}} \rangle \\
&= -\langle \mu_{-\boldsymbol{l}} | \boldsymbol{r} \times \boldsymbol{p} | \nu_{\mathbf{0}} \rangle - \boldsymbol{D}\boldsymbol{l} \times \langle \mu_{-\boldsymbol{l}} | \boldsymbol{p} | \nu_{\mathbf{0}} \rangle \\
&= -\boldsymbol{m}^{\mathbf{0}-\boldsymbol{l}}_{\nu\mu} + \boldsymbol{D}\boldsymbol{l} \times \boldsymbol{p}^{\mathbf{0}-\boldsymbol{l}}_{\nu\mu}
\end{aligned} \tag{7.2}$$

The second step is the result of $\boldsymbol{r} e^{-i\boldsymbol{D}\boldsymbol{l}\cdot\boldsymbol{p}} = e^{-i\boldsymbol{D}\boldsymbol{l}\cdot\boldsymbol{p}} \boldsymbol{r} + [\boldsymbol{r}, e^{-i\boldsymbol{D}\boldsymbol{l}\cdot\boldsymbol{p}}]$ where $[\cdot,\cdot]$ is the commutator. The third step evaluates the commutator in the second term using the relationship $[\boldsymbol{r}, f(\boldsymbol{p})] = i f'(\boldsymbol{p})$

which holds for any function of momentum $f(\boldsymbol{p})$ expressed in atomic units.[78] The last step uses that the magnetic dipole and momentum matrix elements are pure imaginary, so taking their complex conjugates is equivalent to negating them.

Expressing the magnetic dipole matrix in a Hermitian form will help us to collect terms that arise during the derivation and is convenient for storing/manipulating the matrix in our implementation. Defining

$$\tilde{\boldsymbol{m}}^{\boldsymbol{0l}}_{\mu\nu} = \frac{1}{2}\left(\boldsymbol{m}^{\boldsymbol{0l}}_{\mu\nu} - \boldsymbol{m}^{\boldsymbol{0-l}}_{\nu\mu}\right) = \boldsymbol{m}^{\boldsymbol{0l}}_{\mu\nu} + \frac{1}{2}\boldsymbol{Dl}\times\boldsymbol{p}^{\boldsymbol{0l}}_{\mu\nu} = (\tilde{\boldsymbol{m}}^{\boldsymbol{0-l}}_{\nu\mu})^{*} \tag{7.3}$$

produces a Hermitian matrix by splitting the origin dependent momentum term between the given matrix element and the corresponding adjoint matrix element. This is equivalent to the procedure used to make the electric dipole Hermitian in Eq. 2.42.

We can now proceed to transforming the magnetic dipole operator to a translation invariant form. As with the PBC electric dipole derivation shown in section 2.3, this involves substituting the length operator $\boldsymbol{r}$ with a translation invariant form as shown in Eq. 2.35.[68,75,77] Applying this transformation to the magnetic dipole operator yields

$$\begin{aligned}
-\boldsymbol{r}\times\boldsymbol{p} \simeq& (-ie^{i\boldsymbol{k}\cdot\boldsymbol{r}}\nabla_{\boldsymbol{k}}e^{-i\boldsymbol{k}\cdot\boldsymbol{r}})\times\boldsymbol{p} \\
=& e^{i\boldsymbol{k}\cdot\boldsymbol{r}}(-i\nabla_{\boldsymbol{k}}\times\boldsymbol{p})e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} + e^{i\boldsymbol{k}\cdot\boldsymbol{r}}(-i\nabla_{\boldsymbol{k}}\times[e^{-i\boldsymbol{k}\cdot\boldsymbol{r}},\boldsymbol{p}]) \\
=& e^{i\boldsymbol{k}\cdot\boldsymbol{r}}(-i\nabla_{\boldsymbol{k}}\times\boldsymbol{p})e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} + e^{i\boldsymbol{k}\cdot\boldsymbol{r}}(-i\nabla_{\boldsymbol{k}}\times\boldsymbol{k})e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \\
=& -ie^{i\boldsymbol{k}\cdot\boldsymbol{r}}\left[(\nabla_{\boldsymbol{k}}\times\boldsymbol{p}) + (\nabla_{\boldsymbol{k}}\times\boldsymbol{k})\right]e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}
\end{aligned} \tag{7.4}$$

The second step uses the identity $\exp(i\boldsymbol{k}\cdot\boldsymbol{r})\boldsymbol{p} = \boldsymbol{p}\exp(i\boldsymbol{k}\cdot\boldsymbol{r}) + [\exp(i\boldsymbol{k}\cdot\boldsymbol{r}),\boldsymbol{p}]$. The commutator is then evaluated in the third step using $[f(\boldsymbol{r}),\boldsymbol{p},] = if'(\boldsymbol{r})$.[78]

Using the transformed operator, we can now derive magnetic dipole matrix elements between Bloch functions

$$\begin{aligned}
\hat{\boldsymbol{m}}_{p\boldsymbol{k},q\boldsymbol{k}} &= -i\langle p\boldsymbol{k}|e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\left((\nabla_{\boldsymbol{k}}\times\boldsymbol{p}) + (\nabla_{\boldsymbol{k}}\times\boldsymbol{k})\right)e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}|q\boldsymbol{k}\rangle \\
&= -i\langle u_{p\boldsymbol{k}}|\nabla_{\boldsymbol{k}}\times\boldsymbol{p}|u_{q\boldsymbol{k}}\rangle - i\langle u_{p\boldsymbol{k}}|\nabla_{\boldsymbol{k}}\times\boldsymbol{k}|u_{q\boldsymbol{k}}\rangle
\end{aligned} \tag{7.5}$$

where the Bloch functions $|q\boldsymbol{k}\rangle$ and their cell periodic part $u_{q\boldsymbol{k}}(\boldsymbol{r})$ are defined in Eqs. 2.30 and 2.37

respectively. We write the crystal orbital PBC magnetic dipole matrix as $\underline{\hat{\boldsymbol{m}}}$ to distinguish it with from the AO matrix $\underline{\boldsymbol{m}}$ and the Hermitian AO matrix $\underline{\tilde{\boldsymbol{m}}}$. We now must express these reciprocal-space, crystal orbital matrix elements in terms of real-space, atomic orbital matrix elements so they can be evaluated.

We can start to evaluate the first term on the right hand side Eq. 7.5 by applying the operators sequentially to the cell periodic ket

$$\begin{aligned} -i\boldsymbol{p}|u_{q\boldsymbol{k}}\rangle &= \left[ -\sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}} e^{i\boldsymbol{k}\cdot(\boldsymbol{D}\boldsymbol{l}-\boldsymbol{r})} \sum_{\nu=1}^{\Omega} C_{q\nu}(\boldsymbol{k}) \frac{\partial \chi_\nu^{\boldsymbol{l}}(\boldsymbol{r})}{\partial \boldsymbol{r}} + i\boldsymbol{k} \sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}} e^{i\boldsymbol{k}\cdot(\boldsymbol{D}\boldsymbol{l}-\boldsymbol{r})} \sum_{\nu=1}^{\Omega} C_{q\nu}(\boldsymbol{k}) \chi_\nu^{\boldsymbol{l}}(\boldsymbol{r}) \right] \\ &= \boldsymbol{\Xi} \end{aligned} \tag{7.6}$$

where $\boldsymbol{\Xi}$ is simply a collection of the terms produced by applying the momentum operator to the ket. Continuing to apply $\nabla_{\boldsymbol{k}}\times$ yields

$$\nabla_{\boldsymbol{k}} \times \boldsymbol{\Xi} = \Bigg[ i\boldsymbol{r} \times \sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}} e^{i\boldsymbol{k}\cdot(\boldsymbol{D}\boldsymbol{l}-\boldsymbol{r})} \sum_{\nu=1}^{\Omega} C_{q\nu}(\boldsymbol{k}) \frac{\partial \chi_\nu^{\boldsymbol{l}}(\boldsymbol{r})}{\partial \boldsymbol{r}} \tag{7.7a}$$

$$- i \sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}} \boldsymbol{D}\boldsymbol{l} \times e^{i\boldsymbol{k}\cdot(\boldsymbol{D}\boldsymbol{l}-\boldsymbol{r})} \sum_{\nu=1}^{\Omega} C_{q\nu}(\boldsymbol{k}) \frac{\partial \chi_\nu^{\boldsymbol{l}}(\boldsymbol{r})}{\partial \boldsymbol{r}} \tag{7.7b}$$

$$- \sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}} e^{i\boldsymbol{k}\cdot(\boldsymbol{D}\boldsymbol{l}-\boldsymbol{r})} \sum_{\nu=1}^{\Omega} \frac{\partial C_{q\nu}}{\partial \boldsymbol{k}} \times \frac{\partial \chi_\nu^{\boldsymbol{l}}(\boldsymbol{r})}{\partial \boldsymbol{r}} \tag{7.7c}$$

$$+ \boldsymbol{r} \times \boldsymbol{k} \sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}} e^{i\boldsymbol{k}\cdot(\boldsymbol{D}\boldsymbol{l}-\boldsymbol{r})} \sum_{\nu=1}^{\Omega} C_{q\nu}(\boldsymbol{k}) \chi_\nu^{\boldsymbol{l}}(\boldsymbol{r}) \tag{7.7d}$$

$$- \sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}} (\boldsymbol{D}\boldsymbol{l}) \times \boldsymbol{k} e^{i\boldsymbol{k}\cdot(\boldsymbol{D}\boldsymbol{l}-\boldsymbol{r})} \sum_{\nu=1}^{\Omega} C_{q\nu}(\boldsymbol{k}) \chi_\nu^{\boldsymbol{l}}(\boldsymbol{r}) \tag{7.7e}$$

$$+ i\boldsymbol{k} \times \sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}} e^{i\boldsymbol{k}\cdot(\boldsymbol{D}\boldsymbol{l}-\boldsymbol{r})} \sum_{\nu=1}^{\Omega} \frac{\partial C_{q\nu}(\boldsymbol{k})}{\partial \boldsymbol{k}} \chi_\nu^{\boldsymbol{l}}(\boldsymbol{r}) \Bigg] \tag{7.7f}$$

We can evaluate the ket of the second term of Eq 7.5 in the same way

$$-i\nabla_{\boldsymbol{k}} \times \boldsymbol{k}|u_{q\boldsymbol{k}}\rangle = \Bigg[ -\boldsymbol{r} \times \boldsymbol{k} \sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}} e^{i\boldsymbol{k}\cdot(\boldsymbol{D}\boldsymbol{l}-\boldsymbol{r})} \sum_{\nu=1}^{\Omega} C_{q\nu}(\boldsymbol{k}) \chi_{\nu}^{\boldsymbol{l}}(\boldsymbol{r}) \tag{7.8d}$$

$$+ \sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}} (\boldsymbol{D}\boldsymbol{l}) \times \boldsymbol{k} e^{i\boldsymbol{k}\cdot(\boldsymbol{D}\boldsymbol{l}-\boldsymbol{r})} \sum_{\nu=1}^{\Omega} C_{q\nu}(\boldsymbol{k}) \chi_{\nu}^{\boldsymbol{l}}(\boldsymbol{r}) \tag{7.8e}$$

$$- i\boldsymbol{k} \sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}} e^{i\boldsymbol{k}\cdot(\boldsymbol{D}\boldsymbol{l}-\boldsymbol{r})} \sum_{\nu=1}^{\Omega} \frac{\partial C_{q\nu}(\boldsymbol{k})}{\partial \boldsymbol{k}} \chi_{\nu}^{\boldsymbol{l}}(\boldsymbol{r}) \Bigg] \tag{7.8f}$$

Notice, Eqs. 7.7d-f and 7.8d-f are identical except for their sign, so these contributions will cancel out. Acting the cell periodic bra from the right on the remaining terms in Eq. 7.7, we obtain an expression for the periodic magnetic dipole matrix

$$\hat{\underline{\boldsymbol{m}}} = \Bigg[ \underline{C}^{\dagger} \underline{\boldsymbol{m}} \underline{C} \tag{7.9a}$$

$$- i\underline{C}^{\dagger} (\nabla_{\boldsymbol{k}} \times \underline{\boldsymbol{p}}) \underline{C} \tag{7.9b}$$

$$- i\underline{C}^{\dagger} \underline{\boldsymbol{p}} \underline{C} \times \underline{\boldsymbol{U}} \Bigg] \tag{7.9c}$$

where 7.9a-c are the terms that result from applying the bra to corresponding term of 7.7a-c. $\underline{\boldsymbol{U}}$ is a vector of matrices that transforms the MO coefficients $C$ to their gradient with respect to $\boldsymbol{k}$, as defined in 2.3. We can make one further simplification by noting that

$$-i\nabla_{\boldsymbol{k}} \times \boldsymbol{p}_{\mu\nu} = -i\nabla_{\boldsymbol{k}} \times \sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}} e^{i\boldsymbol{k}\cdot\boldsymbol{D}\boldsymbol{l}} \boldsymbol{p}_{\mu\nu}^{\boldsymbol{0}\boldsymbol{l}} = \sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}} e^{i\boldsymbol{k}\cdot\boldsymbol{D}\boldsymbol{l}} (\boldsymbol{D}\boldsymbol{l}) \times \boldsymbol{p}_{\mu\nu}^{\boldsymbol{0}\boldsymbol{l}} \tag{7.10}$$

which we can utilize to transform the real-space magnetic dipole matrix $\underline{\boldsymbol{m}}$ to the Hermitian matrix $\underline{\tilde{\boldsymbol{m}}}$ in Eq. 7.3. The resulting final expression for the periodic magnetic dipole is

$$\hat{\underline{\boldsymbol{m}}} = \underline{C}^{\dagger} \underline{\tilde{\boldsymbol{m}}} \underline{C} - \frac{i}{2} \underline{C}^{\dagger} (\nabla_{\boldsymbol{k}} \times \underline{\boldsymbol{p}}) \underline{C} - i\underline{C}^{\dagger} \underline{\boldsymbol{p}} \underline{C} \times \underline{\boldsymbol{U}} \tag{7.11}$$

Recently, Rerát and Kirtman independently published a derivation of the PBC magnetic dipole, along with some proof of concept calculations of the $\boldsymbol{\beta}$ tensor.[50] While their derivation takes

different steps to the final expression and uses slightly different notation, the resulting expressions for the PBC magnetic dipole are equivalent. As Rerát and Kirtman note, ensuring that $\boldsymbol{\beta}$ has all real elements is equivalent to the constraint that $\underline{\hat{\boldsymbol{m}}}$ is Hermitian. However, the second term in Eq. 7.11 is anti-Hermitian and the third term in neither Hermitian nor anti-Hermitian. To ensure that the OR tensor is real, we explicitly symmetrize the magnetic dipole matrix. In practice, we only need to symmetrize the third term of Eq. 7.11, as the first term is already Hermitian and the second term, being anti-Hermitian, is canceled out by the symmetrization.

### 7.2.2 Periodic Electric Quadrupole

The derivation of electric quadrupole for a periodic system proceeds in much the same way as for the electric and magnetic dipoles. The following will focus on the velocity form of the electric quadrupole operator, as this is the form used in both the modified velocity gauge (MVG) and origin-independent length gauge [LG(OI)]. In the velocity gauge, the real space matrix elements of the electric quadrupole are written as

$$\boldsymbol{Q}^{\mathbf{0}\boldsymbol{l}}_{\mu\nu} = \langle \mu_{\mathbf{0}} | \frac{1}{2} (\boldsymbol{r} \otimes \boldsymbol{p} + \boldsymbol{p} \otimes \boldsymbol{r}) | \nu_{\boldsymbol{l}} \rangle = \langle \mu_{\mathbf{0}} | \boldsymbol{r} \odot \boldsymbol{p} | \nu_{\boldsymbol{l}} \rangle = \langle \mu_{\mathbf{0}} | \boldsymbol{r} \odot \boldsymbol{p} e^{-i\boldsymbol{D}\boldsymbol{l}\cdot\boldsymbol{p}} | \nu_{\mathbf{0}} \rangle \tag{7.12}$$

where $\otimes$ is the outer product and $\odot$ is the symmetric outer product. Again, the presence of the translation operator leads to a non-Hermitian matrix, which we can correct by reallocating the deviation from Hermiticity for a given matrix element evenly between itself and its corresponding adjoint element. We first determine the form of this deviation by moving the translation operator from the ket to the bra

$$\begin{aligned}
\boldsymbol{Q}^{\mathbf{0}\boldsymbol{l}}_{\mu\nu} &= \langle \mu_{\mathbf{0}} | \boldsymbol{r} \odot \boldsymbol{p} e^{-i\boldsymbol{D}\boldsymbol{l}\cdot\boldsymbol{p}} | \nu_{\mathbf{0}} \rangle \\
&= \langle \mu_{-\boldsymbol{l}} | \boldsymbol{r} \odot \boldsymbol{p} | \nu_{\mathbf{0}} \rangle + \langle \mu_{\mathbf{0}} | [\boldsymbol{r}, e^{-i\boldsymbol{D}\boldsymbol{l}\cdot\boldsymbol{p}}] \odot \boldsymbol{p} | \nu_{\mathbf{0}} \rangle \\
&= \langle \mu_{-\boldsymbol{l}} | \boldsymbol{r} \odot \boldsymbol{p} | \nu_{\mathbf{0}} \rangle + \boldsymbol{D}\boldsymbol{l} \odot \langle \mu_{-\boldsymbol{l}} | \boldsymbol{p} | \nu_{\mathbf{0}} \rangle \\
&= -\boldsymbol{Q}^{\mathbf{0}-\boldsymbol{l}}_{\nu\mu} - \boldsymbol{D}\boldsymbol{l} \odot \boldsymbol{p}^{\mathbf{0}-\boldsymbol{l}}_{\nu\mu}
\end{aligned} \tag{7.13}$$

Knowing the form of the deviation, we can define a Hermitian quadrupole in the same manner as for the magnetic and electric dipoles

$$\tilde{\boldsymbol{Q}}^{\boldsymbol{0l}}_{\mu\nu} = \frac{1}{2}\left(\boldsymbol{Q}^{\boldsymbol{0l}}_{\mu\nu} - \boldsymbol{Q}^{\boldsymbol{0-l}}_{\nu\mu}\right) = \boldsymbol{Q}^{\boldsymbol{0l}}_{\mu\nu} - \frac{1}{2}\boldsymbol{Dl} \odot \boldsymbol{p}^{\boldsymbol{0l}}_{\mu\nu} = (\tilde{\boldsymbol{Q}}^{\boldsymbol{0-l}}_{\nu\mu})^* \tag{7.14}$$

We can now proceed to apply the transformation of Eq. 2.35 to the quadrupole operator in order to make it translation invariant.

$$\begin{aligned}
\boldsymbol{r} \odot \boldsymbol{p} \simeq & \left(ie^{i\boldsymbol{k}\cdot\boldsymbol{r}} \nabla_{\boldsymbol{k}} e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\right) \odot \boldsymbol{p} \\
= & e^{i\boldsymbol{k}\cdot\boldsymbol{r}} (i\nabla_{\boldsymbol{k}} \odot \boldsymbol{p}) e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} + e^{i\boldsymbol{k}\cdot\boldsymbol{r}} (i\nabla_{\boldsymbol{k}} \odot [e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}, \boldsymbol{p}]) \\
= & e^{i\boldsymbol{k}\cdot\boldsymbol{r}} (-i\nabla_{\boldsymbol{k}} \times \boldsymbol{p}) e^{i\boldsymbol{k}\cdot\boldsymbol{r}} + e^{i\boldsymbol{k}\cdot\boldsymbol{r}} (i\nabla_{\boldsymbol{k}} \odot \boldsymbol{k}) e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \\
= & ie^{i\boldsymbol{k}\cdot\boldsymbol{r}} \Big( (\nabla_{\boldsymbol{k}} \odot \boldsymbol{p}) + (\nabla_{\boldsymbol{k}} \odot \boldsymbol{k}) \Big) e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}
\end{aligned} \tag{7.15}$$

The expression for the PBC quadrupole operator in Eq. 7.4 is very similar in form to that of the PBC magnetic dipole in Eq. 7.4, with the only difference being a change of sign and the substitution $\times \to \odot$. Consequently, applying the PBC quadrupole operator to the cell periodic ket leads to similar contributions. Applying the first term of the transformed operator to the cell periodic ket, we obtain

$$\begin{aligned}
i\boldsymbol{p}|u_{q\boldsymbol{k}}\rangle = & \left[ \sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}} e^{i\boldsymbol{k}\cdot(\boldsymbol{Dl}-\boldsymbol{r})} \sum_{\nu=1}^{\Omega} C_{q\nu}(\boldsymbol{k}) \frac{\partial \chi^{\boldsymbol{l}}_{\nu}(\boldsymbol{r})}{\partial \boldsymbol{r}} - i\boldsymbol{k} \sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}} e^{i\boldsymbol{k}\cdot(\boldsymbol{Dl}-\boldsymbol{r})} \sum_{\nu=1}^{\Omega} C_{q\nu}(\boldsymbol{k}) \chi^{\boldsymbol{l}}_{\nu}(\boldsymbol{r}) \right] \\
= & \boldsymbol{\Gamma}
\end{aligned} \tag{7.16}$$

where $\boldsymbol{\Gamma}$ just collects the intermediate terms produced by applying the momentum operator. Con-

tinuing to apply $\nabla_{\boldsymbol{k}}\odot$ yields

$$\nabla_{\boldsymbol{k}}\odot\boldsymbol{\Gamma}=\Bigg[-i\boldsymbol{r}\odot\sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}}e^{i\boldsymbol{k}\cdot(\boldsymbol{D}\boldsymbol{l}-\boldsymbol{r})}\sum_{\nu=1}^{\Omega}C_{q\nu}(\boldsymbol{k})\frac{\partial\chi_{\nu}^{\boldsymbol{l}}(\boldsymbol{r})}{\partial\boldsymbol{r}} \tag{7.17a}$$

$$+i\sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}}\boldsymbol{D}\boldsymbol{l}\odot e^{i\boldsymbol{k}\cdot(\boldsymbol{D}\boldsymbol{l}-\boldsymbol{r})}\sum_{\nu=1}^{\Omega}C_{q\nu}(\boldsymbol{k})\frac{\partial\chi_{\nu}^{\boldsymbol{l}}(\boldsymbol{r})}{\partial\boldsymbol{r}} \tag{7.17b}$$

$$+\sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}}e^{i\boldsymbol{k}\cdot(\boldsymbol{D}\boldsymbol{l}-\boldsymbol{r})}\sum_{\nu=1}^{\Omega}\frac{\partial C_{q\nu}}{\partial\boldsymbol{k}}\odot\frac{\partial\chi_{\nu}^{\boldsymbol{l}}(\boldsymbol{r})}{\partial\boldsymbol{r}} \tag{7.17c}$$

$$-\boldsymbol{r}\odot\boldsymbol{k}\sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}}e^{i\boldsymbol{k}\cdot(\boldsymbol{D}\boldsymbol{l}-\boldsymbol{r})}\sum_{\nu=1}^{\Omega}C_{q\nu}(\boldsymbol{k})\chi_{\nu}^{\boldsymbol{l}}(\boldsymbol{r}) \tag{7.17d}$$

$$+\sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}}(\boldsymbol{D}\boldsymbol{l})\odot\boldsymbol{k}e^{i\boldsymbol{k}\cdot(\boldsymbol{D}\boldsymbol{l}-\boldsymbol{r})}\sum_{\nu=1}^{\Omega}C_{q\nu}(\boldsymbol{k})\chi_{\nu}^{\boldsymbol{l}}(\boldsymbol{r}) \tag{7.17e}$$

$$-i\boldsymbol{k}\odot\sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}}e^{i\boldsymbol{k}\cdot(\boldsymbol{D}\boldsymbol{l}-\boldsymbol{r})}\sum_{\nu=1}^{\Omega}\frac{\partial C_{q\nu}(\boldsymbol{k})}{\partial\boldsymbol{k}}\chi_{\nu}^{\boldsymbol{l}}(\boldsymbol{r}) \tag{7.17f}$$

$$-i\boldsymbol{\delta}\odot\sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}}e^{i\boldsymbol{k}\cdot(\boldsymbol{D}\boldsymbol{l}-\boldsymbol{r})}\sum_{\nu=1}^{\Omega}C_{q\nu}(\boldsymbol{k})\chi_{\nu}^{\boldsymbol{l}}(\boldsymbol{r})\Bigg] \tag{7.17g}$$

where $\delta$ is the Kronecker delta. The "extra" term 7.17g (compare to Eq. 7.7) appears since the outer product allows $\nabla_{\boldsymbol{k}}$ to act on the $\boldsymbol{k}$ in the second term of Eq. 7.16.

Applying to the ket of the second term proceeds similarly

$$i\nabla_{\boldsymbol{k}}\times\boldsymbol{k}=\Bigg[\boldsymbol{r}\times\boldsymbol{k}\sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}}e^{i\boldsymbol{k}\cdot(\boldsymbol{D}\boldsymbol{l}-\boldsymbol{r})}\sum_{\nu=1}^{\Omega}C_{q\nu}(\boldsymbol{k})\chi_{\nu}^{\boldsymbol{l}}(\boldsymbol{r}) \tag{7.18d}$$

$$-\sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}}(\boldsymbol{D}\boldsymbol{l})\times\boldsymbol{k}e^{i\boldsymbol{k}\cdot(\boldsymbol{D}\boldsymbol{l}-\boldsymbol{r})}\sum_{\nu=1}^{\Omega}C_{q\nu}(\boldsymbol{k})\chi_{\nu}^{\boldsymbol{l}}(\boldsymbol{r}) \tag{7.18e}$$

$$+i\boldsymbol{k}\sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}}e^{i\boldsymbol{k}\cdot(\boldsymbol{D}\boldsymbol{l}-\boldsymbol{r})}\sum_{\nu=1}^{\Omega}\frac{\partial C_{q\nu}(\boldsymbol{k})}{\partial\boldsymbol{k}}\chi_{\nu}^{\boldsymbol{l}}(\boldsymbol{r}) \tag{7.18f}$$

$$+i\boldsymbol{\delta}\sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}}e^{i\boldsymbol{k}\cdot(\boldsymbol{D}\boldsymbol{l}-\boldsymbol{r})}\sum_{\nu=1}^{\Omega}C_{q\nu}(\boldsymbol{k})\chi_{\nu}^{\boldsymbol{l}}(\boldsymbol{r})\Bigg] \tag{7.18g}$$

where these terms are equal and opposite in sign to Eq. 7.17d-g and thus cancel out of the final

expression. After applying the bra, the remaining terms simplify to

$$\hat{\underline{\boldsymbol{Q}}} = \Bigg[ \underline{C}^{\dagger} \underline{\boldsymbol{Q}} \underline{C} \tag{7.19a}$$

$$+ i\underline{C}^{\dagger} (\nabla_{\boldsymbol{k}} \odot \underline{\boldsymbol{p}}) \underline{C} \tag{7.19b}$$

$$+ i\underline{C}^{\dagger} \underline{\boldsymbol{p}} \underline{C} \odot \underline{\boldsymbol{U}} \Bigg] \tag{7.19c}$$

It is convenient to re-express the AO basis quadrupole $\boldsymbol{Q}$ as the Hermitian matrix $\tilde{\boldsymbol{Q}}$. Noting that

$$i\nabla_{\boldsymbol{k}} \odot \boldsymbol{p}_{\mu\nu} = i\nabla_{\boldsymbol{k}} \odot \sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}} e^{i\boldsymbol{k}\cdot\boldsymbol{D}\boldsymbol{l}} \boldsymbol{p}_{\mu\nu}^{\boldsymbol{0}\boldsymbol{l}} = -\sum_{\boldsymbol{l}=-\boldsymbol{L}}^{\boldsymbol{L}} e^{i\boldsymbol{k}\cdot\boldsymbol{D}\boldsymbol{l}} (\boldsymbol{D}\boldsymbol{l}) \odot \boldsymbol{p}_{\mu\nu}^{\boldsymbol{0}\boldsymbol{l}} \tag{7.20}$$

the PBC electric quadrupole can be rewritten as

$$\hat{\underline{\boldsymbol{Q}}} = \underline{C}^{\dagger} \tilde{\underline{\boldsymbol{Q}}} \underline{C} + \frac{i}{2} \underline{C}^{\dagger} (\nabla_{\boldsymbol{k}} \odot \underline{\boldsymbol{p}}) \underline{C} + i\underline{C}^{\dagger} \underline{\boldsymbol{p}} \underline{C} \odot \underline{\boldsymbol{U}} \tag{7.21}$$

to incorporate $\tilde{\boldsymbol{Q}}$. To the best of our knowledge, this is the first derivation of the PBC electric quadrupole. Notice the resemblance in form with Eq. 7.11, where the main difference is the use of a symmetric direct product rather than a cross product, along with a change of sign in each term. As was the case with the Eq. 7.11, we need to ensure that $\hat{\boldsymbol{Q}}$ is Hermitian and we again achieve this via explicit symmetrization, with the second term in Eq. 7.21 canceling out due to being anti-Hermitian.

## 7.3 Implementation Details

With expressions derived for the PBC magnetic dipole and electric quadrupole, there remain several challenges in order to incorporate them into an electronic structure program (in this case, a development version of Gaussian[95]) in order to compute the full OR tensor. The main obstacles are updating the existing coupled perturbed self-consistent field (CP-SCF) routines to accept multipoles defined in reciprocal, rather than real space and modifying the integral engine to produce

Hermitian matrices for the supercell. While the OR tensor is formally defined by the sum over states expressions of Eqs. 2.2 and 2.3, in practice it is much more efficient to evaluate these tensors using linear response CP-SCF, as described in section 2.2.

Scuseria et al. provided the initial framework for these changes to the integral engine and CP-SCF procedure during their implementation of the PBC electric dipole and the corresponding dipole-dipole polarizability.[68,75] However, this existing code was not actively maintained and could only be applied to the length gauge electric dipole, so extensive updates were made in order to make it compatible with more recent versions of Gaussian and to extend its functionality to the PBC magnetic dipole and velocity gauge electric quadrupole. We also implemented the PBC velocity gauge electric dipole, but since this operator is translation invariant, it does not require any correction terms.

With these PBC operators at our disposal, we can implement MVG and LG(OI) $\boldsymbol{\beta}$ and $\mathbf{A}$ tensors. For each choice of gauge, $\boldsymbol{\beta}$ can be computed using either an electric or magnetic dipole perturbation when solving the linear response CP-SCF equations. In principle, we could also compute $\mathbf{A}$ with an electric dipole or electric quadrupole perturbation, but since using an electric dipole perturbation allows us to compute $\boldsymbol{\beta}$ and $\mathbf{A}$ simultaneously, we have not attempted to implement the quadrupole perturbation.

As part of the testing/validation process for our implementation, we separately compute the contributions to $\boldsymbol{\beta}$ and $\mathbf{A}$ from each term of Eqs. 7.11 and 7.21, respectively. The contribution from the first term of each of these equations corresponds to an uncorrected polarizability computed on the supercell, while the latter two terms are corrections that incorporate the effect of periodically extending the system. By separating out these contributions, it is easier to diagnose what is causing any errors in the computed OR tensor. Once the implementation is fully functioning, this separation can also provide a quantitative measure of the relative importance of the PBC corrections to the overall OR.

## 7.4 Results

To test our implementation, we calculate the $\boldsymbol{\beta}$ and $\mathbf{A}$ tensors for a 1D chain of $H_2O_2$ molecules repeated along the x-axis, the same simple model used by Rerát and Kirtman.[50] Since this is only meant as a proof of concept, we performed the calculation in the modified velocity gauge using the BLYP functional[135,146] and the modest 3-21G basis set.[66] As an initial test run, we performed a "pseudo-molecular" calculation where we extended the translation vector of the $H_2O_2$ chain to $10^5$Å so it could not interact with its periodic images. We found that $\boldsymbol{\beta}$ and $\mathbf{A}$ computed for this pseudo-molecule matched the results of a molecular calculation of $H_2O_2$ using the same level of theory and basis set.

To make an assessment of our implementation for a true periodic system, we computed the OR tensors at the same level of theory for two large, but finite chains with 121 and 119 $H_2O_2$ units. Taking the difference of the OR tensors from these calculations and dividing by two gives a very good approximation of the OR tensor per unit cell for the infinitely extended periodic system. We confirmed that the OR tensor values per unit cell were sufficiently converged by carrying out the same procedure with chains of 151 and 149 $H_2O_2$ units, for which we saw no deviation of the OR tensor components.

The results of these calculations for the $\boldsymbol{\beta}$ tensor are reported in Table 7.1. We can see that while the supercell and PBC tensor components mostly agree, the yy, zz, and xz show significant deviation. Checking the decomposed PBC $\boldsymbol{\beta}$ tensor, we can see that the components with errors are those with nonzero correction terms. The xx, yx, and zx components should trivially have zero correction: since the system is periodic along the x-axis, there are no corrections to the PBC magnetic dipole along the x-axis, since all correction terms involve a cross product with the translation vector. However, it is noteworthy that the xy, yz, and zy components also have zero correction, despite the PBC magnetic dipole having corrections along the y and z axes. While this may just be due to a fortuitous choice for an example system, it does demonstrate that our implementation of the uncorrected $\boldsymbol{\beta}$ accurately calculates the total $\boldsymbol{\beta}$ when the corrections are zero, whether incidentally or due to symmetry.

The deviation of yy, zz, and xz, however, suggest that our implementation of the correction terms is faulty. While the uncorrected $\boldsymbol{\beta}$ components are relatively close to the values from the supercell calculations, adding in the correction does not improve the agreement. The error in these components is not a consistent over/underestimation that could be resolved by rescaling or shifting the correction; while the correction to the PBC zz component is in the right direction, but overshoots the true value, the xz component correction shifts it even farther from the true value and the yy component correction is negligible. Further testing will need to be done to isolate the error in the implementation of the PBC magnetic dipole correction terms.

Table 7.1: Modified velocity gauge $\boldsymbol{\beta}$ tensor elements (a.u.) per unit cell for an $H_2O_2$ chain computed at the BLYP/3-21G level. The first column reports a finite supercell calculation, while the second column shows the results of a PBC calculation. The third and fourth columns split the PBC tensor into uncorrected and correction terms.

| | $\boldsymbol{\beta}$(super) | $\boldsymbol{\beta}$(PBC total) | $\boldsymbol{\beta}$(PBC uncorrected) | $\boldsymbol{\beta}$(PBC corrections) |
|---|---|---|---|---|
| xx | 1.571 | 1.571 | 1.571 | 0.000 |
| yy | 4.054 | 9.278 | 3.481 | 5.797 |
| zz | -5.716 | -10.856 | -4.760 | -6.086 |
| xy | 1.114 | 1.114 | 1.114 | 0.000 |
| yx | 0.000 | 0.000 | 0.000 | 0.000 |
| xz | -5.163 | -21.972 | -6.696 | -15.276 |
| zx | -2.438 | -2.438 | -2.438 | 0.000 |
| yz | -4.541 | -4.541 | -4.541 | 0.000 |
| zy | 3.671 | 3.671 | 3.671 | 0.000 |

Since the implementation of the $\boldsymbol{\beta}$ tensor is still incomplete, we forego discussion of the **A** tensor until we can make an adequate comparison based on its contribution to the full OR tensor.

## 7.5 Conclusion

In this chapter, we have derived expressions for the PBC magnetic dipole and PBC velocity gauge electric quadrupole operators. While an alternative derivation of the PBC magnetic dipole has been recently reported elsewhere and applied to compute the periodic electric dipole-magnetic dipole polarizability tensor $\beta$, to our knowledge, this work is the first to report an expression for the PBC

electric quadrupole. With these operators, along with the previously derived PBC length gauge electric dipole and the inherently translation invariant velocity gauge electric dipole, it should be possible to compute both the $\boldsymbol{\beta}$ and $\mathbf{A}$ contributions to the full optical rotation tensor in both the modified velocity gauge and origin independent length gauge.

# Chapter 8

# Concluding Remarks

Optical rotation has proven to be a challenging property to predict; chemical intuition linking the structure of a molecule and its OR is limited and this intrinsic OR can often be masked by optical activity induced by the surrounding environment. In this thesis, we have tried to address this challenge on two fronts. We have improved upon the $\tilde{S}$ analysis framework, which offers an interpretation of the OR in terms of occupied-virtual orbital pair transitions. Our efforts have helped to make this model applicable for a wider range of molecular systems and computational protocols, allowing for more robust predictions of the physical processes underlying OR. We have also worked to develop methods for computing OR of periodic systems in order to provide greater insight into the effect of the environment.

In Chapter 3, in our effort to better characterize intrinsic OR, we compared the results of $\tilde{S}$ analysis in the modified velocity gauge to prior results computed with the length gauge to determine if the physical interpretations produced in each gauge were consistent. We found that different choices of gauge and perturbation provided a consistent physical description at various levels of detail. The distribution of $\tilde{S}_{ia}$ contributions, including the largest magnitude contribution, was identical for any choice of gauge or perturbation. Summing the contributions from largest to smallest in magnitude, each gauge approached the total OR at the same rate. Most importantly, at the level of individual transitions, the $\tilde{S}_{ia}$ vectors in different gauges provided the same physical description of how charge was flowing within the molecule to induce OR.

We also examined approaches for removing the arbitrary dependence of $\tilde{S}_{ia}$ values on the chosen coordinate origin in Chapter 4. We derived two different origin invariant formulations of MVG $\tilde{S}$ analysis: $\tilde{S}^{\text{Avg}}$, which averages the result of $\tilde{S}^{\text{MVG-M}}$ and $\tilde{S}^{\text{MVG-E}}$, and $\tilde{S}^{\text{Hemi}}$, which contracts

together Hemi-perturbed electric and magnetic densities formed via a Cholesky decomposition of the response matrix. The origin invariance of both approaches was confirmed by comparing test molecule calculations performed at the center of mass and at a displaced origin. Comparisons with $\tilde{S}^{\mathrm{Avg}}$ or $\tilde{S}^{\mathrm{Hemi}}$ provide a simple test for detecting significant effects from the choice of origin in MVG-M/MVG-M calculations. While still requiring further development to ensure general applicability, these approaches offer a promising step towards an entirely origin invariant formulation of $\tilde{S}$ analysis.

To study the effect of environmental interactions on the OR, in Chapter 5 we constructed a test set of helical chains of diatomic molecules. These helical chains serve as a simple model of periodic solids for which the OR can be computed using existing molecular methods. Using this model we found that, even for simple 1D systems with fairly small unit cell sizes, convergence of the OR to the macroscopic limit was slow with increasing system size. The semiempirical Kirkwood polarizability model was found to provide a reasonable description of the OR for homonuclear helices, but could not achieve even qualitative agreement for a heteronuclear helix. We demonstrated that the electric dipole-electric quadrupole tensor, which is not needed to model solution phase OR, has a significant contribution to the solid OR tensor. Calculations on these models offered clear evidence that a quantum mechanical approach implementing periodic boundary conditions (PBC) would be needed to accurately compute the OR for solid systems. The results gathered for these helical models offers a convenient test-bed to validate the results of our PBC implementation.

In Chapter 6, recognizing the need for highly accurate computational procedures to compute solid state OR, we assessed the effect of gauge when computing the full OR tensor with newly available CCSD methods and compared these with the results from two commonly used DFT functionals. We found that while the tensor components and traces computed with MVG DFT and CCSD were highly correlated, DFT consistently overestimated the CCSD results. The choice of functional had an effect on the extent of overestimation, as B3LYP overestimated by 20-30%, while the range-separated CAM-B3LYP functional was closer to 5% overestimation. MVG and LG(OI) CCSD calculations were in very close agreement. However, for comparisons between DFT and

CCSD and comparisons between different gauges at the CCSD level, even individual outliers have a substantial effect on the correlation for a given component or the trace.

In Chapter 7, we have developed an approach to compute the full OR tensor for periodic systems by deriving translation invariant expressions for the magnetic dipole and electric quadrupole. The magnetic dipole derivation matches that of the recently published work by Rerát and Kirtman.[50] To the best of our knowledge, ours is the first published derivation of a translation invariant electric quadrupole. We are in the process of implementing these multipoles into GAUSSIAN. Once the implementation is complete, we will be able to compute the full PBC OR tensor with LG(OI) or the MVG using either an electric or a magnetic perturbation.

Efforts to characterize chiral mixtures, identify new chiral species, and design molecular/-supramolecular structures with specially tuned optical properties would all benefit from an improved understanding of how OR relates to molecular structure. Integrating this knowledge into computational models could greatly expedite, or possibly even circumvent, the often laborious and cost-intensive process of experimentally assigning absolute configurations. Our work on developing $\tilde{S}$ analysis and methods for solid state optical rotation should aid in future efforts to computationally disentangle structural, intrinsic OR from that induced via the influence of the surrounding environment.

There are a number of pathways along which our development work could be extended in order to provide more accurate, intuitive, and general methods for assessing chiroptical response. An immediate goal of our research group is to combining the $\tilde{S}$ framework with our PBC OR implementation. $\tilde{S}$ analysis of solids, in particular molecular crystals, would be an invaluable tool for interrogating to what extent the OR arise from point centered molecular chirality versus bulk environmental chirality. Our use of atomic centered basis functions, rather than the plane waves more commonly used to study solid state systems, will provide fine-grained detail into both the local and bulk interactions leading to OR.

We would also like to extend both $\tilde{S}$ analysis and PBC OR to other post Hartree-Fock methods, specifically coupled cluster. DFT methods offer a cost-effective way to probe chiroptical response,

but if a functional provides an insufficient treatment of electron correlation for a given molecule, there is not a systematic path to improvement. CCSD calculations already provide reasonable agreement with experiment and this will only improve as higher excitations levels become feasible through algorithm development and improvement of computing resources.

An interesting application of $\tilde{S}$ analysis that we have yet to explore is its use in molecular design. $\tilde{S}$ provides a physical picture of how charge moving within a molecule enhances or detracts from its chiroptical response. Using this physical interpretation, it should be possible to predict how structural changes and functionalization within a family of molecules would influence their OR. If accurate predictions of the effect of substitution can be made, $\tilde{S}_{ia}$ analysis can be used to develop design criteria for tuning the OR of a template molecule.

While optical rotation has been the focus of our work thus far, many of these developments are likely applicable to circular dichroism (CD) with slight modification. Both vibrational and electronic CD have increasingly become standard tools for studying chiral biological systems.[20,147–150] While OR arises from the imaginary part of the electric dipole-magnetic dipole polarizability, CD comes from the real part, so the $\tilde{S}$ framework that decomposes this tensor in orbital pair contributions should be equally applicable to CD. Recently, a procedure has been published for computing vibrational CD for periodic systems using *ab initio* molecular dynamics.[151] Their approach uses symmetry properties of the time correlation function to resolve the issue of translation dependence of the magnetic dipole, but it would be interesting to see if the VCD could also be determined using static electronic structure calculations with the newly derived translation-invariant magnetic dipole.

Looking toward the future, a number of challenges still remain to develop a comprehensive, chemically intuitive understanding of the physical processes underlying optical rotation and chiroptical response more generally. While gas and solid phase simulations will hopefully provide insight into how to disentangle intrinsic and environmentally induced OR, accurate simulations of solution phase systems would be highly desirable, but seem out of reach with current computational resources and methodologies. Coupled cluster methods of sufficient accuracy are too

expensive to explicitly model enough of the solvent environment; DFT methods allow for much larger simulations, but are unable to consistently compute even the intrinsic OR and lack a means of systematic improvement. Since OR stems from a nonadditive combination of a molecule and its environment, it is not easily amenable to multiscale methods where different parts of the system are treated at different levels of theory.

Machine learning has recently emerged as a possible route to overcoming this obstacle.[152,153] Using a sufficiently accurate quantum mechanical method to simulate the OR of an isolated molecule, it may be possible to construct a machine learning model for the solvent shift to match experiment. This is similar to the Δ-machine learning approach that has been applied to compute experimentally accurate thermochemical properties.[154] While an entirely black-box approach would do little to provide insight in OR, recent work has focused on making chemical machine learning models more interpretable.[155,156] Such a model may even be able to incorporate information from existing OR frameworks, using the outputs of an OR decomposition method like $\tilde{S}$ as the inputs for its predictions.

In general, chiroptical research will need to focus on cost-reduction of OR simulations. There are now a number of usable computational models for interpreting OR, but these are only useful insofar as the underlying calculations accurately models experiment and sufficiently accurate methods seem to be currently out of reach for even modest system sizes. By speeding up accurate OR calculations, whether through new hardware, algorithmic improvements, or machine learning supplementation, we can more confidently apply the results of OR interpretive models and delve into what physical processes are actually inducing optical rotation.

# References


[1] Lien Ai Nguyen, Hua He, and Chuong Pham-Huy. Chiral Drugs: An Overview. *Int J Biomed Sci*, 2(2):85–100, June 2006.

[2] W.H. Brooks, W.C. Guida, and K.G. Daniel. The significance of chirality in drug design and development. *Curr Top Med Chem*, 11(7):760–770, 2011.

[3] Silas W. Smith. Chiral Toxicology: It's the Same Thing...Only Different. *Toxicol. Sci.*, 110(1):4–30, July 2009.

[4] Gérald Lelais and David W C MacMillan. Modern Strategies in Organic Catalysis: The Advent and Development of Iminium Activation. *Aldrichimica Acta*, 39(3):11, 2006.

[5] Barry M. Trost. Asymmetric Catalysis: An Enabling Science. *ChemInform*, 35(36), September 2004.

[6] Nikhil Srivastava, Lingamurthy Macha, and Hyun-Joon Ha. Total synthesis and stereochemical revision of biemamides b and d. *Org. Lett.*, 21(22):8992–8996, November 2019.

[7] Andreas Pfaltz and William J. Drury. Design of chiral ligands for asymmetric catalysis: From C2-symmetric P,P- and N,N-ligands to sterically and electronically nonsymmetrical P,N-ligands. *PNAS*, 101(16):5723–5726, April 2004.

[8] Lei Hu, Yansong Ren, and Olof Ramström. Chirality control in enzyme-catalyzed dynamic kinetic resolution of 1,3-Oxathiolanes. *J. Org. Chem.*, 80(16):8478–8481, 2015.

[9] Eiji Yashima, Naoki Ousaka, Daisuke Taura, Kouhei Shimomura, Tomoyuki Ikai, and Katsuhiro Maeda. Supramolecular Helical Systems: Helical Assemblies of Small Molecules,

Foldamers, and Polymers with Chiral Amplification and Their Functions. *Chem. Rev.*, 116(22):13752–13990, November 2016.

[10] Marcin L. Ślęczkowski, Mathijs F. J. Mabesoone, Piotr Ślęczkowski, Anja R. A. Palmans, and E. W. Meijer. Competition between chiral solvents and chiral monomers in the helical bias of supramolecular polymers. *Nat. Chem.*, 13(2):200–207, February 2021.

[11] Tom W. Anderson, Jeremy K. M. Sanders, and G. Dan Pantoş. The sergeants-and-soldiers effect: Chiral amplification in naphthalenediimide nanotubes. *Org. Biomol. Chem.*, 8(19):4274, 2010.

[12] Yuuya Nagata, Ryohei Takeda, and Michinori Suginome. Asymmetric Catalysis in Chiral Solvents: Chirality Transfer with Amplification of Homochirality through a Helical Macromolecular Scaffold. *ACS Cent. Sci.*, 5(7):1235–1240, July 2019.

[13] Timothy J. Ward and Karen D. Ward. Chiral Separations: A Review of Current Topics and Trends. *Anal. Chem.*, 84(2):626–635, January 2012.

[14] Louis Pasteur. Mémoire sur la relation qui peut exister entre la forme cristalline et la composition chimique, et sur la cause de la polarisation rotatoire. *Comptes Rendus Hebd. Séances L'Académie Sci.*, 26:535–538, 1848.

[15] Ghislaine Vantomme and Jeanne Crassous. Pasteur and chirality: A story of how serendipity favors the prepared minds. *Chirality*, 33(10):597–601, 2021.

[16] H. D. Flack and G. Bernardinelli. The use of X-ray crystallography to determine absolute configuration. *Chirality*, 20(5):681–690, 2008.

[17] Amber L. Thompson and David John Watkin. X-ray crystallography and chirality: Understanding the limitations. *Tetrahedron: Asymmetry*, 20(6-8):712–717, May 2009.

[18] Hisako Sato. A new horizon for vibrational circular dichroism spectroscopy: A challenge for supramolecular chirality. *Phys. Chem. Chem. Phys.*, 22(15):7671–7679, 2020.

[19] Haoke Zhang, Xiaoyan Zheng, Ryan T. K. Kwok, Jia Wang, Nelson L. C. Leung, Lin Shi, Jing Zhi Sun, Zhiyong Tang, Jacky W. Y. Lam, Anjun Qin, and Ben Zhong Tang. In situ monitoring of molecular aggregation using circular dichroism. *Nat Commun*, 9(1):4961, December 2018.

[20] András Micsonai, Frank Wien, Linda Kernya, Young-Ho Lee, Yuji Goto, Matthieu Réfrégiers, and József Kardos. Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. *Proc Natl Acad Sci USA*, 112(24):E3095–E3103, June 2015.

[21] Oleksii Shemchuk, Fabrizia Grepioni, Tom Leyssens, and Dario Braga. Chiral Resolution via Cocrystallization with Inorganic Salts. *Isr. J. Chem.*, 61(9-10):563–572, 2021.

[22] Michael Guillot, Joséphine de Meester, Sarah Huynen, Laurent Collard, Koen Robeyns, Olivier Riant, and Tom Leyssens. Cocrystallization-Induced Spontaneous Deracemization: A General Thermodynamic Approach to Deracemization. *Angew. Chem. Int. Ed.*, 59(28):11303–11306, 2020.

[23] François Arago. *Sur Une Modification Remarquable Qu'éprouvent Les Rayons Lumineux Dans Leur Passage à Travers Certains Corps Diaphanes, et Sur Quelques Autres Nouveaux Phénomènes d'optiques: Lu Le 11 Août 1811*. Firmin-Didot, 1812.

[24] Jean-Baptiste Biot. *Memoire Sur Un Nouveau Genre d'oscillation Que Les Molecules de La Lumiere Eprouvent En Traversant Certains Cristaux*. Chez Firmin Didot, 1814.

[25] Bart Kahr and Oriol Arteaga. Arago's Best Paper. *ChemPhysChem*, 13(1):79–88, 2012.

[26] P. L. Polavarapu. Ab initio molecular optical rotations and absolute configurations. *Mol. Phys.*, 91(3):551–554, June 1997.

[27] Magdalena Pecul and Kenneth Ruud. The Ab Initio Calculation of Optical Rotation and Electronic Circular Dichroism. In *Advances in Quantum Chemistry*, volume 50, pages 185–212. Elsevier, 2005.

[28] T. Daniel Crawford, Mary C. Tam, and Micah L. Abrams. The current state of ab initio calculations of optical rotation and electronic circular dichroism spectra. *J. Phys. Chem. A*, 111(48):12057–12068, 2007.

[29] Patrick Norman, Kenneth Ruud, and Trygve Helgaker. Density-functional theory calculations of optical rotatory dispersion in the nonresonant and resonant frequency regions. *J. Chem. Phys.*, 120(11):5027–5035, 2004.

[30] Priyanka Lahiri, Kenneth B. Wiberg, and Patrick H. Vaccaro. A tale of two carenes: Intrinsic optical activity and large-amplitude nuclear displacement. *J. Phys. Chem. A*, 116(38):9516–9533, September 2012.

[31] Kenneth B. Wiberg, Marco Caricato, Yi-Gui Wang, and Patrick H. Vaccaro. Towards the accurate and efficient calculation of optical rotatory dispersion using augmented minimal basis sets. *Chirality*, 25(10):606–616, 2013.

[32] Shokouh Haghdani, Bård Helge Hoff, Henrik Koch, and Per-Olof Åstrand. Solvent Effects on Optical Rotation: On the Balance between Hydrogen Bonding and Shifts in Dihedral Angles. *J. Phys. Chem. A*, 121(25):4765–4777, June 2017.

[33] John G. Kirkwood. On the theory of optical rotatory power. *J. Chem. Phys.*, 5(6):479, 1937.

[34] P. L. Polavarapu, D. K. Chakraborty, and K. Ruud. Molecular optical rotation: An evaluation of semiempirical models. *Chem. Phys. Lett.*, 319(5):595–600, 2000.

[35] Rama K. Kondru, Peter Wipf, and David N. Beratan. Atomic contributions to the optical rotation angle as a quantitative probe of molecular chirality. *Science*, 282(5397):2247–2250, December 1998.

[36] Veronica L. Murphy and Bart Kahr. Hückel theory and optical activity. *J. Am. Chem. Soc.*, 137(15):5177–5183, 2015.

[37] Marc de Wergifosse and Stefan Grimme. A Unified Strategy for the Chemically Intuitive Interpretation of Molecular Optical Response Properties. *J. Chem. Theory Comput.*, 16(12):7709–7720, December 2020.

[38] Barry Moore, Monika Srebro, and Jochen Autschbach. Analysis of optical activity in terms of bonds and lone-pairs: The exceptionally large optical rotation of norbornenone. *J. Chem. Theory Comput.*, 8(11):4336–4346, 2012.

[39] Marco Caricato. Conformational effects on specific rotation: A theoretical study based on the $\tilde{S}_k$ method. *J. Phys. Chem. A*, 119(30):8303–8310, 2015.

[40] Tal Aharon and Marco Caricato. Configuration space analysis of the specific rotation of helicenes. *J. Phys. Chem. A*, 123(20):4406–4418, 2019.

[41] Tal Aharon, Paul Lemler, Patrick H. Vaccaro, and Marco Caricato. Comparison of measured and predicted specific optical rotation in gas and solution phases: A test for the polarizable continuum model of solvation. *Chirality*, 30(4):383–395, 2018.

[42] Parag Mukhopadhyay, Gérard Zuber, Michael-Rock Goldsmith, Peter Wipf, and David N. Beratan. Solvent Effect on Optical Rotation: A Case Study of Methyloxirane in Water. *ChemPhysChem*, 7(12):2483–2486, 2006.

[43] Parag Mukhopadhyay, Gérard Zuber, Peter Wipf, and David N. Beratan. Contribution of a Solute's Chiral Solvent Imprint to Optical Rotation. *Angew. Chem. Int. Ed.*, 46(34):6450–6452, 2007.

[44] Ruhee D'Cunha and T. Daniel Crawford. Modeling Complex Solvent Effects on the Optical Rotation of Chiral Molecules: A Combined Molecular Dynamics and Density Functional Theory Study. *J. Phys. Chem. A*, 125(15):3095–3108, April 2021.

[45] Alexander T. Martin, Shane M. Nichols, Veronica L. Murphy, and Bart Kahr. Chiroptical anisotropy of crystals and molecules. *Chem. Commun.*, 57(66):8107–8120, 2021.

[46] V. Devarajan and A. M. Glazer. Theory and Computation of Optical Rotatory Power in Inorganic Crystals. *Acta Crystallogr.*, A42:560–569, 1986.

[47] Hua Zhong, Zachary H. Levine, Douglas C. Allan, and John W. Wilkins. Optical activity of selenium: A nearly first-principles calculation. *Phys. Rev. Lett.*, 69(2):379–382, 1992.

[48] Hua Zhong, Zachary H. Levine, Douglas C. Allan, and John W. Wilkins. Band-theoretic calculations of the optical-activity tensor of $\alpha$-quartz and trigonal Se. *Phys. Rev. B*, 48(3):1384–1403, 1993.

[49] A. D. Buckingham and M. B. Dunn. Optical activity of oriented molecules. *J. Chem. Soc. Inorg. Phys. Theor. Chem.*, (0):1988–1991, 1971.

[50] Michel Rérat and Bernard Kirtman. First-Principles Calculation of the Optical Rotatory Power of Periodic Systems: Application on $\alpha$-Quartz, Tartaric Acid Crystal, and Chiral (n,m)-Carbon Nanotubes. *J. Chem. Theory Comput.*, 17(7):4063–4076, July 2021.

[51] Jochen Autschbach. Time-dependent density functional theory for calculating origin-independent optical rotation and rotatory strength tensors. *ChemPhysChem*, 12(17):3224–3235, 2011.

[52] Ty Balduf and Marco Caricato. Helical Chains of Diatomic Molecules as a Model for Solid-State Optical Rotation. *J. Phys. Chem. C*, 123(7):4329–4340, 2019.

[53] L. D Barron. *Molecular Light Scattering and Optical Activity*. Cambridge University Press, Cambridge, UK; New York, 2004.

[54] Mykhaylo Krykunov and Jochen Autschbach. Calculation of origin-independent optical rotation tensor components in approximate time-dependent density functional theory. *J. Chem. Phys.*, 125:034102, 2006.

[55] David W. Ball. *Field Guide to Spectroscopy*. SPIE Field Guides. SPIE Press, Bellingham, Wash, 2006.

[56] David J. Griffiths. *Introduction to Electrodynamics*. Cambridge University Press, Cambridge, United Kingdom ; New York, NY, fourth edition, 2018.

[57] Jeanne L. McHale. *Molecular Spectroscopy*. CRC Press, Taylor & Francis Group, Boca Raton, second edition, 2017.

[58] Introduction to response theory. In Jerzy Leszczynski, Anna Kaczmarek-Kedziera, Manthos G. Papadopoulous, Heribert Reis, Andrzej J. Sadlej, and Manoj K. Shukla, editors, *Handbook of Computational Chemistry*, Springer Reference, pages 148–151. Springer Netherlands, first edition, 2012.

[59] Kuo-Ho Yang and Kirk T. McDonald. Formal Expressions for the Electromagnetic Potentials in Any Gauge. 2019.

[60] Thomas Bondo Pedersen, Henrik Koch, Linus Boman, and Alfredo M.J Sánchez de Merás. Origin invariant calculation of optical rotation without recourse to London orbitals. *Chemical Physics Letters*, 393(4-6):319–326, August 2004.

[61] Robert Ditchfield. Self-consistent perturbation theory of diamagnetism I. A gauge-invariant LCAO method for N.M.R. Chemical shifts. *Mol. Phys.*, 27(4):789–807, 1974.

[62] Marco Caricato. Origin invariant optical rotation in the length dipole gauge without London atomic orbitals. *J. Chem. Phys.*, 153(15):151101, October 2020.

[63] Kenneth B. Wiberg, Yi Gui Wang, Shaun M. Wilson, Patrick H. Vaccaro, and James R. Cheeseman. Sum-over-states calculation of the specific rotations of some substituted oxiranes, chloropropionitrile, ethane, and norbornenone. *J. Phys. Chem. A*, 110(51):13995–14002, 2006.

[64] Roy McWeeny. *Methods of Molecular Quantum Mechanics*. Academic Press, San Diego, second edition, 1978.

[65] R. Moccia. Time-dependent variational principle. *Int. J. Quantum Chem.*, 7(4):779–783, July 1973.

[66] J. A. Pople, R. Krishnan, H. B. Schlegel, and J. S. Binkley. Derivative Studies in Hartree-Fock and Moller-Plesset Theories. *Int. J. Quantum Chem.*, 13:225–241, 1979.

[67] Michael Frisch, Martin Head-Gordon, and John Pople. Direct analytic SCF second derivatives and electric field properties. *Chem. Phys.*, 141(2-3):189–196, 1990.

[68] Artur F. Izmaylov, Edward N. Brothers, and Gustavo E. Scuseria. Linear-scaling calculation of static and dynamic polarizabilities in Hartree-Fock and density functional theory for periodic systems. *J. Chem. Phys.*, 125(22):224105, 2006.

[69] Péter Pulay. Convergence acceleration of iterative sequences. the case of scf iteration. *Chemical Physics Letters*, 73(2):393–398, July 1980.

[70] Kaihua Zhang, Ty Balduf, and Marco Caricato. Full optical rotation tensor at coupled cluster with single and double excitations level in the modified velocity gauge. *Chirality*, 33(6):303–314, June 2021.

[71] Marco Caricato and Ty Balduf. Origin invariant full optical rotation tensor in the length dipole gauge without London atomic orbitals. *J. Chem. Phys.*, 155(2):024118, July 2021.

[72] T. Daniel Crawford. Ab initio calculation of molecular chiroptical properties. *Theor. Chem. Acc.*, 115(4):227–245, 2006.

[73] Henrik Koch and Poul Jørgensen. Coupled cluster response functions. *The Journal of Chemical Physics*, 93(5):3333–3344, September 1990.

[74] Christian Barbier, Joseph Delhalle, and Jean-Marie Andre. Ab initio calculations of the static electric polarizability of infinite polymer chains. *Journal of Molecular Structure: THEOCHEM*, 188(3-4):299–312, August 1989.

[75] Konstantin N. Kudin and Gustavo E. Scuseria. An efficient finite field approach for calculating static electric polarizabilities of periodic systems. *J. Chem. Phys.*, 113(18):7779–7785, 2000.

[76] Konstantin N. Kudin and Gustavo E. Scuseria. Linear-scaling density-functional theory with Gaussian orbitals and periodic boundary conditions: Efficient evaluation of energy and forces via the fast multipole method. *Phys. Rev. B - Condens. Matter Mater. Phys.*, 61(24):16440–16453, 2000.

[77] Benoit Champagne, Joseph G. Fripiat, and Jean-Marie Andre. From uncoupled to coupled Hartree – Fock polarizabilities of infinite polymeric chains. Pariser – Parr – Pople applications to the polyacetylene chains. *J. Chem. Phys.*, 96:8330, 1992.

[78] Mark K. Transtrum and Jean-François S. Van Huele. Commutation relations for functions of operators. *Journal of Mathematical Physics*, 46(6):063510, June 2005.

[79] Ty Balduf and Marco Caricato. Gauge Dependence of the $\tilde{S}$ Molecular Orbital Space Decomposition of Optical Rotation. *J. Phys. Chem. A*, 125(23):4976–4985, June 2021.

[80] Shohei Takata, Yuta Endo, Mohammad Shahid Ullah, and Shinichi Itsuno. Synthesis of cinchona alkaloid sulfonamide polymers as sustainable catalysts for the enantioselective desymmetrization of cyclic anhydrides. *RSC Adv.*, 2(31):10107–10111, 2016.

[81] Alexander J. Mastroianni, Shelley A. Claridge, and A. Paul Alivisatos. Pyramidal and chiral groupings of gold nanocrystals assembled using DNA scaffolds. *J. Am. Chem. Soc.*, 131(24):8455–8459, 2009.

[82] Bilu Liu, Fanqi Wu, Hui Gui, Ming Zheng, and Chongwu Zhou. Chirality-controlled synthesis and applications of single-wall carbon nanotubes. *ACS Nano*, 11(1):31–53, January 2017.

[83] Khai Q. Le, Shun Hashiyada, Masaharu Kondo, and Hiromi Okamoto. Circularly polarized photoluminescence from achiral dye molecules induced by plasmonic two-dimensional chiral nanostructures. *J. Phys. Chem. C*, 122(43):24924–24932, November 2018.

[84] Yinhui Kan, Sebastian K. H. Andersen, Fei Ding, Shailesh Kumar, Changying Zhao, and Sergey I. Bozhevolnyi. Metasurface-Enabled generation of circularly polarized single photons. *Adv. Mater.*, 32(16):1907832, April 2020.

[85] Patrick H. Vaccaro. *Comprehensive Chiroptical Spectroscopy, Instrumentation, Methodologies, and Theoretical Simulations.* John Wiley & Sons;, New York, first edition, 2012.

[86] Sotiris Droulias and Lykourgos Bougas. Absolute chiral sensing in dielectric metasurfaces using signal reversals. *Nano Lett.*, 20(8):5960–5966, August 2020.

[87] Marco Caricato. Orbital analysis of molecular optical activity based on configuration rotatory strength. *J. Chem. Theory Comput.*, 11(4):1349–1353, 2015.

[88] Jochen Autschbach, Lucia Nitsch-Velasquez, and Mark Rudolph. Time-dependent density functional response theory for electronic chiroptical properties of chiral molecules. *Top. Curr. Chem.*, 298:1–98, 2011.

[89] Stefano Pelloni and Paolo Lazzeretti. On the determination of the diagonal components of the optical activity tensor in chiral molecules. *The Journal of Chemical Physics*, 140(7):074105, February 2014.

[90] Mykhaylo Krykunov and Jochen Autschbach. Calculation of optical rotation with time-periodic magnetic-field-dependent basis functions in approximate time-dependent density-functional theory. *J. Chem. Phys.*, 123(11):114103, 2005.

[91] Tal Aharon and Marco Caricato. Compact Basis Sets for Optical Rotation Calculations. *J. Chem. Theory Comput.*, 16(7):4408–4415, 2020.

[92] Priyanka Lahiri, Kenneth B. Wiberg, and Patrick H. Vaccaro. Intrinsic optical activity and conformational flexibility: The role of size-dependent ring morphology in model cycloketones. *J. Phys. Chem. A*, 117(47):12382–12400, November 2013.

[93] Thom H. Dunning. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.*, 90(2):1007, 1989.

[94] Axel D. Becke. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.*, 98(7):5648–5652, April 1993.

[95] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox. Gaussian development version revision J.02, 2018.

[96] Schrödinger, LLC. The PyMOL molecular graphics system, version 1.8. November 2015.

[97] T. Daniel Crawford and Philip J. Stephens. Comparison of Time-Dependent Density-Functional Theory and Coupled Cluster Theory for the Calculation of the Optical Rotations of Chiral Molecules. *J. Phys. Chem. A*, 112(6):1339–1345, February 2008.

[98] Francesco Aquilante, Linus Boman, Jonas Boström, Henrik Koch, Roland Lindh, Alfredo Sánchez de Merás, and Thomas Bondo Pedersen. Cholesky Decomposition Techniques in Electronic Structure Theory. In Robert Zalesny, Manthos G. Papadopoulos, Paul G. Mezey, and Jerzy Leszczynski, editors, *Linear-Scaling Techniques in Computational Chemistry and Physics*, volume 13, pages 301–343. Springer Netherlands, Dordrecht, 2011.

[99] J. Čížek and J. Paldus. Stability Conditions for the Solutions of the Hartree—Fock Equations for Atomic and Molecular Systems. Application to the Pi-Electron Model of Cyclic Polyenes. *The Journal of Chemical Physics*, 47(10):3976–3985, November 1967.

[100] Wei Chen, Ai Bian, Ashish Agarwal, Liqiang Liu, Hebai Shen, Libing Wang, Chuanlai Xu, and Nicholas A. Kotov. Nanoparticle superstructures made by polymerase chain reaction: Collective interactions of nanoparticles and a new principle for chiral materials. *Nano Lett.*, 9(5):2153–2159, 2009.

[101] Chuan Zhang, Weimin Wu, Xiang Li, Cheng Tian, Hang Qian, Guansong Wang, Wen Jiang, and Chengde Mao. Controlling the chirality of DNA nanocages. *Angew. Chem. - Int. Ed.*, 51(32):7999–8002, 2012.

[102] Thomas Müller, Kenneth B. Wiberg, and Patrick H. Vaccaro. Cavity ring-down polarimetry (CRDP): A new scheme for probing circular birefringence and circular dichroism in the gas phase. *J. Phys. Chem. A*, 104(25):5959–5968, 2000.

[103] Thomas Müller, Kenneth B. Wiberg, Patrick H. Vaccaro, James R. Cheeseman, and Michael J. Frisch. Cavity ring-down polarimetry (CRDP): Theoretical and experimental characterization. *J. Opt. Soc. Am. B*, 19(1):125–141, 2002.

[104] Kacey Claborn, Javier Herreros Cedres, Christine Isborn, Alexey Zozulya, Edgar Weckert, Werner Kaminsky, and Bart Kahr. Optical rotation of achiral pentaerythritol. *J. Am. Chem. Soc.*, 128(46):14746–14747, 2006.

[105] Bart Kahr, John Freudenthal, and Erica Gunn. Crystals in light. *Acc. Chem. Res.*, 43(5):684–692, 2010.

[106] Alexander T. Martin, Shane M. Nichols, Sichao Li, Melissa Tan, and Bart Kahr. Double cone of eigendirections in optically active ethylenediammonium selenate crystals. *J. Appl. Crystallogr.*, 50:1117–1124, 2017.

[107] RD Johnson. Computational Chemistry Comparison and Benchmark Database, NIST Standard Reference Database 101, 2002.

[108] Takeshi Yanai, David P. Tew, and Nicholas C. Handy. A new hybrid exchange-correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chem. Phys. Lett.*, 393(21):51–57, 2004.

[109] P. J. Stephens, D. M. McCann, J. R. Cheeseman, and M. J. Frisch. Determination of absolute configurations of chiral molecules using ab initio time-dependent density functional theory calculations of optical rotation: How reliable are absolute configurations obtained for molecules with small rotations? *Chirality*, 17(S1):S52–S64, 2005.

[110] Francisco E. Jorge, Amanda Z. De Oliveira, and Thiago P. Silva. CAM-B3LYP optical rotations at different wavelengths: Comparison with CCSD results. *Int. J. Quantum Chem.*, 116(1):21–26, 2016.

[111] Kenneth Ruud and Trygve Helgaker. Optical rotation studied by density-functional and coupled-cluster methods. *Chem. Phys. Lett.*, 352(5-6):533–539, 2002.

[112] P. J. Stephens, F. J. Devlin, J. R. Cheeseman, and M. J. Frisch. Calculation of Optical Rotation Using Density Functional Theory. *J. Phys. Chem. A*, 105(22):5356–5371, 2001.

[113] Werner Kaminsky, Kacey Claborn, and Bart Kahr. Polarimetric imaging of crystals. *Chem. Soc. Rev.*, 33(8):514, 2004.

[114] John H. Freudenthal, Eva Hollis, and Bart Kahr. Imaging chiroptical artifacts. *Chirality*, 21(1E):E20–E27, 2009.

[115] James R. Cheeseman, Michael J. Frisch, Frank J. Devlin, and Philip J. Stephens. Hartree-Fock and Density Functional Theory ab Initio Calculation of Optical Rotation Using GIAOs: Basis Set Dependence. *J. Phys. Chem. A*, 104(5):1039–1046, February 2000.

[116] Stefan Grimme. Calculation of frequency dependent optical rotation using density functional response theory. *Chemical Physics Letters*, 339(5-6):380–388, May 2001.

[117] Prasad L. Polavarapu. Optical rotation: Recent advances in determining the absolute configuration. *Chirality*, 14(10):768–781, 2002.

[118] Jochen Autschbach, Serguei Patchkovskii, Tom Ziegler, Stan J. A. van Gisbergen, and Evert Jan Baerends. Chiroptical properties from time-dependent density functional theory. II. Optical rotations of small to medium sized organic molecules. *The Journal of Chemical Physics*, 117(2):581–592, July 2002.

[119] Stefan Grimme, Arnold Bahlmann, and Günter Haufe. Ab initio calculations for the optical rotations of conformationally flexible molecules: A case study on six-, seven-, and eight-membered fluorinated cycloalkanol esters. *Chirality*, 14(10):793–797, 2002.

[120] Kenneth Ruud, Philip J. Stephens, Frank J. Devlin, Peter R. Taylor, James R. Cheeseman, and Michael J. Frisch. Coupled-cluster calculations of optical rotation. *Chemical Physics Letters*, 373(5-6):606–614, May 2003.

[121] Mary C. Tam, Nicholas J. Russ, and T. Daniel Crawford. Coupled cluster calculations of optical rotatory dispersion of (*S*)-methyloxirane. *The Journal of Chemical Physics*, 121(8):3550–3557, August 2004.

[122] T. Daniel Crawford, Lesley S. Owens, Mary C. Tam, Peter R. Schreiner, and Henrik Koch. Ab Initio Calculation of Optical Rotation in (*P*)-(+)-[4]Triangulane. *J. Am. Chem. Soc.*, 127(5):1368–1369, February 2005.

[123] Jochen Autschbach. Computing chiroptical properties with first-principles theoretical methods: Background and illustrative examples. *Chirality*, 21(1E):E116–E152, 2009.

[124] Veronica L. Murphy, Adam Reyes, and Bart Kahr. Aromaticity and Optical Activity. *J. Am. Chem. Soc.*, 138(1):25–27, January 2016.

[125] Paolo Lazzeretti. Invariance of molecular response properties under a coordinate translation. *Int. J. Quantum Chem.*, 114(20):1364–1392, October 2014.

[126] Maria C. Caputo, Stefano Pelloni, and Paolo Lazzeretti. Theoretical prediction of the optical rotation of chiral molecules in ordered media: A computational study of (Ra)-1,3-dimethylallene, (2R)-2-methyloxirane, and (2R)-N-methyloxaziridine. *Int. J. Quantum Chem.*, 115(14):900–906, 2015.

[127] F. London. Théorie quantique des courants interatomiques dans les combinaisons aromatiques. *J. Phys. Radium*, 8(10):397–409, 1937.

[128] Trygve Helgaker, Kenneth Ruud, Keld L. Bak, Poul Jørgensen, and Jeppe Olsen. Vibrational Raman optical activity calculations using London atomic orbitals. *Faraday Discuss.*, 1994.

[129] Ove Christiansen, Poul Jørgensen, and Christof Hättig. Response functions from fourier component variational perturbation theory applied to a time-averaged quasienergy. *Int. J. Quantum Chem.*, 68(1):1–52, 1998.

[130] Thomas Bondo Pedersen, Berta Fernández, and Henrik Koch. Gauge invariant coupled cluster response theory using optimized nonorthogonal orbitals. *The Journal of Chemical Physics*, 114(16):6983–6993, April 2001.

[131] Grant D. Lindh, Taylor J. Mach, and T. Daniel Crawford. The optimized orbital coupled cluster doubles method and optical rotation. *Chemical Physics*, 401:125–129, June 2012.

[132] Taylor J. Mach and T. Daniel Crawford. Basis Set Dependence of Coupled Cluster Optical Rotation Computations. *J. Phys. Chem. A*, 115(35):10045–10051, September 2011.

[133] Shokouh Haghdani, Per-Olof Åstrand, and Henrik Koch. Optical Rotation from Coupled Cluster and Density Functional Theory: The Role of Basis Set Convergence. *J. Chem. Theory Comput.*, 12(2):535–548, February 2016.

[134] Monika Srebro, Niranjan Govind, Wibe A. de Jong, and Jochen Autschbach. Optical Rotation Calculated with Time-Dependent Density Functional Theory: The OR45 Benchmark. *J. Phys. Chem. A*, 115(40):10930–10949, October 2011.

[135] Chengteh Lee, Weitao Yang, and Robert G. Parr. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B*, 37(2):785–789, January 1988.

[136] Marco Caricato, Patrick H. Vaccaro, T. Daniel Crawford, Kenneth. B. Wiberg, and Priyanka Lahiri. Insights on the Origin of the Unusually Large Specific Rotation of (1*S*,4*S*)-Norbornenone. *J. Phys. Chem. A*, 118(26):4863–4871, July 2014.

[137] Priyanka Lahiri, Kenneth B. Wiberg, Patrick H. Vaccaro, Marco Caricato, and T. Daniel Crawford. Large Solvation Effect in the Optical Rotatory Dispersion of Norbornenone. *Angew. Chem. Int. Ed.*, 53(5):1386–1389, 2014.

[138] Taylor Parsons, Ty Balduf, James R. Cheeseman, and Marco Caricato. Basis Set Dependence of Optical Rotation Calculations with Different Choices of Gauge. *J. Phys. Chem. A*, 126(11):1861–1870, 2022.

[139] Angelika Baranowska-Laczkowska and Krzysztof Z. Laczkowski. The ORP basis set designed for optical rotation calculations. *J. Comput. Chem.*, 34(23):2006–2013, 2013.

[140] Kenneth B. Wiberg, Marco Caricato, Yi-Gui Wang, and Patrick H. Vaccaro. Towards the Accurate and Efficient Calculation of Optical Rotatory Dispersion Using Augmented Minimal Basis Sets. *Chirality*, 25(10):606–616, 2013.

[141] Benedetta Mennucci, Chiara Cappelli, Roberto Cammi, and Jacopo Tomasi. Modeling solvent effects on chiroptical properties. *Chirality*, 23(9):717–729, 2011.

[142] William Moffitt. Optical Rotatory Dispersion of Helical Polymers. *The Journal of Chemical Physics*, 25(3):467–478, September 1956.

[143] David A. Rabenold. Helical polymer circular dichroism: Transition matrix theory for polypeptides. *The Journal of Chemical Physics*, 74(2):941–947, January 1981.

[144] Alan I. Levin and Ignacio Tinoco. Classical theory of circular dichroism in infinite helical polymers. *The Journal of Chemical Physics*, 66(8):3491–3497, April 1977.

[145] William Rhodes. Generalized Susceptibility Theory. III. Spectral Properties of Extended Molecular Systems: Optical Rotation of Helical Polymers. *The Journal of Chemical Physics*, 53(9):3650–3661, November 1970.

[146] A. D. Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A*, 38(6):3098–3100, 1988.

[147] Laurence A. Nafie. Vibrational Circular Dichroism: A New Tool for the Solution-State Determination of the Structure and Absolute Configuration of Chiral Natural Product Molecules. *Natural Product Communications*, 3(3):1934578X0800300322, March 2008.

[148] David M. Rogers, Sarah B. Jasim, Naomi T. Dyer, François Auvray, Matthieu Réfrégiers, and Jonathan D. Hirst. Electronic Circular Dichroism Spectroscopy of Proteins. *Chem*, 5(11):2751–2774, November 2019.

[149] Michaela Vorlíčková, Iva Kejnovská, Klára Bednářová, Daniel Renčiuk, and Jaroslav Kypr. Circular Dichroism Spectroscopy of DNA: From Duplexes to Quadruplexes. *Chirality*, 24(9):691–698, 2012.

[150] Norma J Greenfield. Using circular dichroism spectra to estimate protein secondary structure. *Nat Protoc*, 1(6):2876–2890, December 2006.

[151] Sascha Jähnigen, Anne Zehnacker, and Rodolphe Vuilleumier. Computation of Solid-State Vibrational Circular Dichroism in the Periodic Gauge. *J. Phys. Chem. Lett.*, 12(30):7213–7220, August 2021.

[152] Mengyao Chen, Ting Wu, Kaixia Xiao, Tanfeng Zhao, Yanmei Zhou, Qingyou Zhang, and Joao Aires-de-Sousa. Machine learning to predict the specific optical rotations of chiral fluorinated molecules. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 223:117289, December 2019.

[153] Rafael Mamede, Bruno Simões de-Almeida, Mengyao Chen, Qingyou Zhang, and Joao Aires-de-Sousa. Machine Learning Classification of One-Chiral-Center Organic Molecules According to Optical Rotation. *J. Chem. Inf. Model.*, 61(1):67–75, January 2021.

[154] Raghunathan Ramakrishnan, Pavlo O. Dral, Matthias Rupp, and O. Anatole von Lilienfeld. Big Data Meets Quantum Chemistry Approximations: The Δ-Machine Learning Approach. *J. Chem. Theory Comput.*, 11(5):2087–2096, May 2015.

[155] Noushin Omidvar, Hemanth S. Pillai, Shih-Han Wang, Tianyou Mou, Siwen Wang, Andy Athawale, Luke E. K. Achenie, and Hongliang Xin. Interpretable Machine Learning of Chemical Bonding at Solid Surfaces. *J. Phys. Chem. Lett.*, 12(46):11476–11487, November 2021.

[156] Ritesh Kumar and Abhishek K. Singh. Chemical hardness-driven interpretable machine learning approach for rapid search of photocatalysts. *npj Comput Mater*, 7(1):197, December 2021.