

DECOMPOSITION OF INTERMOLECULAR  
INTERACTIONS IN *AB INITIO* SPECTROSCOPY

by

**Eric John Berquist**

B.S. Chemistry, Northeastern University, 2012

Submitted to the Graduate Faculty of  
the Kenneth P. Dietrich School of Arts and Sciences in partial  
fulfillment

of the requirements for the degree of

**Doctor of Philosophy**

University of Pittsburgh

2018

UNIVERSITY OF PITTSBURGH  
KENNETH P. DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Eric John Berquist

It was defended on

March 23rd, 2018

and approved by

Daniel S. Lambrecht, Assistant Professor, Department of Chemistry

Kenneth D. Jordan, Professor, Department of Chemistry

Sean Garrett-Roe, Assistant Professor, Department of Chemistry

David Yaron, Professor, Department of Chemistry, Carnegie Mellon University

Dissertation Director: Daniel S. Lambrecht, Assistant Professor, Department of Chemistry

Copyright © by Eric John Berquist  
2018

# DECOMPOSITION OF INTERMOLECULAR INTERACTIONS IN *AB INITIO* SPECTROSCOPY

Eric John Berquist, PhD

University of Pittsburgh, 2018

Spectroscopy, the response of matter to electromagnetic radiation of different wavelengths, is a powerful experimental tool for interrogating a molecule's structure and dynamics as it interacts with its environment. However, relating a spectroscopic signature to a molecular picture relies on sophisticated computational approaches in order to identify structures, intermolecular interactions, and their correlation with spectroscopic response. This thesis focuses on the question of how to correlate a molecule's structure and interactions with its environment via the *ab initio* calculation of spectroscopic parameters.

To build a molecular picture of CO<sub>2</sub> dynamics in ionic liquids (ILs), I performed quantum chemical calculations on small gas-phase CO<sub>2</sub>-IL clusters, qualitatively reproducing the experimental ordering for CO<sub>2</sub>'s asymmetric vibrational stretch ( $\nu_3$ ) peak position as a function of the anion. To uncover the physical origin of the shift, the language of decomposition analysis based on absolutely localized molecular orbitals (ALMO-EDA) was translated from energies to vibrational frequencies. Geometric distortion of CO<sub>2</sub>, as a result of charge transfer (CT) from the anion into the CO<sub>2</sub>, is the driving force for differentiating the CO<sub>2</sub>  $\nu_3$  shift in different IL anions.

After validating these simple models, I further decomposed the CT contribution into equilibrium structure and potential energy surface curvature mechanisms, finding that CT is a significant contributor in both the geometry optimization and frequency calculation steps. Comparing ALMO-EDA and symmetry-adapted perturbation theory (SAPT) showed that while dispersion dominates the binding energy, DFT-based ALMO-EDA showed excellent



correlation with wavefunction-based SAPT, which enabled the construction of a spectroscopic map based on chemically-intuitive descriptors at lower cost.

This work presents the first application of ALMO-EDA to construct complex spectroscopic maps, however ALMO-EDA is not generally applicable to arbitrary spectroscopies. I reformulated the canonical linear response equations for use with ALMOs to provide a direct connection between EDA terms and their corresponding contribution to spectra. Test calculations indicate that allowing CT is equally important in both the underlying ground-state wavefunction and the response calculations and should not be confused with basis set superposition error.

## TABLE OF CONTENTS

<b>PREFACE</b> . . . . .	xvii
<b>1.0 INTRODUCTION AND THEORETICAL BACKGROUND</b> . . . . .	1
1.1 TERMS AND FUNDAMENTAL DEFINITIONS . . . . .	3
1.2 MACROSCOPIC AND MICROSCOPIC RESPONSE CONNECTIONS . . . . .	5
1.3 STATIC (TIME-INDEPENDENT) RESPONSE PROPERTIES . . . . .	5
1.3.1 Phenomenological approach . . . . .	5
1.3.2 Series expansion . . . . .	11
1.3.3 Derivative evaluation . . . . .	13
1.3.4 Finite difference for numerical derivatives . . . . .	14
1.3.5 Analytic derivative theory . . . . .	17
1.3.6 Perturbation theory . . . . .	23
1.4 DYNAMIC (FREQUENCY-DEPENDENT) RESPONSE PROPERTIES . . . . .	28
1.4.1 Quasienergy derivatives . . . . .	29
<b>2.0 CO<sub>2</sub>-IL CLUSTER MODEL DEVELOPMENT</b> . . . . .	33
2.1 SUMMARY . . . . .	33
2.2 INTRODUCTION . . . . .	34
2.3 METHODS . . . . .	39
2.3.1 Materials . . . . .	39
2.3.2 FTIR . . . . .	40
2.3.3 2D-IR . . . . .	40
2.3.3.1 Generation of femtosecond mid-IR pulses . . . . .	40
2.3.3.2 2D Spectrometer . . . . .	40

2.3.4	Global Fitting and Bootstrapping . . . . .	41
2.3.5	Computational Details . . . . .	42
2.4	RESULTS AND DISCUSSION . . . . .	42
2.4.1	Linear IR Spectroscopy Results . . . . .	42
2.4.2	Vibrational Frequency Calculations . . . . .	46
2.4.3	Frequency, Geometry, and Charge Transfer Discussion . . . . .	49
2.4.4	2D-IR Spectroscopy Overview . . . . .	54
2.4.5	2D-IR Shoulders and Cross-Peaks . . . . .	55
2.4.6	Peak Assignment . . . . .	58
2.4.7	Modelling of the Main Band . . . . .	63
2.4.8	Modelling of Shoulders and Cross-Peaks . . . . .	67
2.5	MOLECULAR INTERPRETATION . . . . .	71
2.6	CONCLUSIONS . . . . .	76
2.7	SUPPORTING INFORMATION . . . . .	77
2.7.1	Comparison of Global Fitting with Center Line Slope . . . . .	77
2.7.2	2D IR Spectra . . . . .	77
2.7.3	Computational Results . . . . .	77
<b>3.0</b>	<b>CO<sub>2</sub>-IL CLUSTER MODEL VALIDATION . . . . .</b>	<b>83</b>
3.1	SUMMARY . . . . .	83
3.2	INTRODUCTION . . . . .	84
3.3	COMPUTATIONAL METHODS . . . . .	88
3.3.1	Methods and Basis Sets . . . . .	88
3.3.2	Anharmonic Vibrational Frequency Calculations . . . . .	90
3.4	FREQUENCY DEPENDENCE ON MODEL APPROXIMATIONS . . . . .	92
3.4.1	Method and Basis Set Dependence . . . . .	92
3.4.2	Anharmonicity . . . . .	99
3.4.3	Molecular Mechanism of CO <sub>2</sub> -IL Interactions . . . . .	101
3.4.3.1	Role of Charge Transfer and Basis Set Superposition Error . . . . .	102
3.4.4	Energetics of CO <sub>2</sub> -IL Interactions and Dependence on Decomposition Approach . . . . .	106

3.4.5	Electrostatic Interactions with the Extended Solvent Phase . . . . .	111
3.4.6	Quantum Mechanical Interactions with the Solvation Shell . . . . .	113
3.4.7	Predicted Solvation Shift . . . . .	115
3.5	CONCLUSIONS . . . . .	115
3.6	ACKNOWLEDGEMENTS . . . . .	117
3.7	SUPPORTING INFORMATION . . . . .	117
3.7.1	Method and Basis Set Dependence of Harmonic Frequencies . . . . .	120
3.7.2	Potential Energy Surface Fitting . . . . .	122
3.7.3	$\nu_3$ Frequency Convergence with Increasing QM Size . . . . .	122
<b>4.0</b>	<b>CO<sub>2</sub>-IL SPECTROSCOPIC MAP DEVELOPMENT . . . . .</b>	<b>132</b>
4.1	SUMMARY . . . . .	132
4.2	INTRODUCTION . . . . .	133
4.3	COMPUTATIONAL METHODS . . . . .	136
4.4	SPECTROSCOPIC MAP FOR CO <sub>2</sub> VIBRATIONS . . . . .	138
4.5	PHYSICAL INTERPRETATION OF THE SPECTROSCOPIC MAP . . . . .	143
4.6	VALIDATION . . . . .	145
4.6.1	Experimental Frequency Distribution . . . . .	145
4.6.2	Calculated Frequency Distributions . . . . .	147
4.7	CONCLUSIONS . . . . .	153
4.8	ACKNOWLEDGEMENTS . . . . .	154
4.9	SUPPORTING INFORMATION . . . . .	154
4.9.1	Transition Dipole Moment Calculations . . . . .	154
<b>5.0</b>	<b>ALMO LINEAR RESPONSE . . . . .</b>	<b>158</b>
5.1	SUMMARY . . . . .	158
5.2	INTRODUCTION . . . . .	158
5.3	THEORY . . . . .	161
5.3.1	Linear response formalism . . . . .	162
5.3.2	Adaptation of linear response equations for absolutely localized MOs . . . . .	163
5.3.3	Decomposition of linear response properties into local contributions and interaction mechanisms . . . . .	164

5.4	METHODS . . . . .	166
5.5	RESULTS AND DISCUSSION . . . . .	168
5.5.1	Equilibrium distance . . . . .	168
5.5.1.1	Validation of analytic ALMO results . . . . .	169
5.5.1.2	CT mechanism analysis . . . . .	170
5.5.1.3	Basis set dependence . . . . .	171
5.5.2	Distance dependence . . . . .	172
5.6	CONCLUSIONS AND FUTURE WORK . . . . .	174
5.7	ACKNOWLEDGEMENTS . . . . .	177
5.8	SUPPORTING INFORMATION . . . . .	177
5.8.1	Input files . . . . .	179
5.8.2	Algorithm description . . . . .	186
<b>6.0</b>	<b>PSI4NUMPY . . . . .</b>	<b>189</b>
6.1	SUMMARY . . . . .	189
6.2	INTRODUCTION . . . . .	190
6.3	BASIC TOOLS . . . . .	194
6.3.1	Wavefunction Objects . . . . .	195
6.3.2	Integrals . . . . .	196
6.3.3	Coulomb and Exchange (JK) Matrix Objects . . . . .	197
6.4	RAPID DEVELOPMENT . . . . .	198
6.5	ACCESS AND CONTRIBUTIONS . . . . .	200
6.6	REFERENCE IMPLEMENTATIONS . . . . .	200
6.6.1	Jupyter Notebook integration . . . . .	202
6.7	CONCLUSIONS . . . . .	203
6.8	ACKNOWLEDGEMENTS . . . . .	205
6.9	SUPPORTING INFORMATION . . . . .	205
6.9.1	Hyperpolarizability Reference Implementation . . . . .	205
6.9.2	Hyperpolarizability Tutorial . . . . .	220
6.9.2.1	Introduction . . . . .	220
6.9.2.2	Notation . . . . .	222

6.9.2.3	Derivation . . . . .	222
6.9.2.4	Wigner's $2n + 1$ rule . . . . .	225
6.9.2.5	Final Expressions . . . . .	226
6.9.2.6	Computational Procedure . . . . .	227
6.9.2.7	References . . . . .	231
<b>7.0</b>	<b>FULL PUBLICATION LIST . . . . .</b>	<b>232</b>
<b>8.0</b>	<b>BIBLIOGRAPHY . . . . .</b>	<b>235</b>

## LIST OF TABLES

1	Connection between energy derivatives and molecular properties . . . . .	6
2	Connection between response functions and molecular properties . . . . .	7
3	Possible analytic and numerical permutations for Raman intensities . . . . .	13
4	Experimental and calculated CO <sub>2</sub> $\nu_3$ vibrational frequencies . . . . .	50
5	CO <sub>2</sub> geometry dependence on charge transfer . . . . .	52
6	Assignment of CO <sub>2</sub> features to 3rd or 5th order . . . . .	64
7	Best fit correlation function parameters, by ionic liquid . . . . .	69
8	CO <sub>2</sub> -IL geometries optimized allowing charge transfer . . . . .	81
9	CO <sub>2</sub> -IL geometries optimized without allowing charge transfer . . . . .	82
10	Functional and basis set dependence of CO <sub>2</sub> bond lengths . . . . .	94
11	Functional and basis set dependence of CO <sub>2</sub> $\nu_3$ frequencies . . . . .	95
12	Charge transfer dependence of CO <sub>2</sub> $\nu_3$ frequencies . . . . .	98
13	PES error in CO <sub>2</sub> $\nu_3$ . . . . .	100
14	Summary of DVR data from MD snapshots . . . . .	101
15	Comparison of ALMO and SAPT0 interaction energies for MD clusters . . . . .	107
16	Correlation between ALMO and SAPT0 terms . . . . .	109
17	Statistical distribution of harmonic frequencies for MD snapshots . . . . .	118
18	Histogram and weights of frequency distributions from MD snapshots . . . . .	119
19	Effect of empirical and self-consistent dispersion corrections in ALMO-EDA . . . . .	124
20	Solvatochromic shift dependence on charge transfer . . . . .	125
21	Solvatochromic shift dependence on counterpoise corrections . . . . .	126
22	Counterpoise correction analysis for [C <sub>1</sub> C <sub>1</sub> im][BF <sub>4</sub> ] . . . . .	127

23	Counterpoise correction analysis for $[C_1C_1im][PF_6]$ . . . . .	128
24	Fit parameters for DVR PES . . . . .	129
25	Fit procedure results for DVR convergence . . . . .	130
26	Parameters of the $CO_2 \nu_3$ spectroscopic map . . . . .	141
27	Decomposition of the spectroscopic map LJ component . . . . .	144
28	Transition dipole moments for gas phase stretching modes of $CO_2$ . . . . .	156
29	Argon—lithium cation dimer polarizabilities using HF/def2-SVP . . . . .	168
30	Argon—lithium cation dimer polarizabilities using HF/def2-SVPD . . . . .	169
31	Percentage changes in argon—lithium cation dimer $\alpha_{\parallel}$ due to the two CT mechanisms . . . . .	170
32	Higher-order ALMO-EDA charge transfer in argon—lithium cation dimer . . . . .	171
33	Point charge and basis function contributions to argon—lithium cation dimer polarizabilities . . . . .	173
34	ALMO-EDA results for argon—lithium cation dimer . . . . .	178
35	Analysis of ALMO-EDA terms from table 34. . . . .	178
36	Point charge and ghost function polarizability analysis . . . . .	179
37	MO coefficient derivatives and rotation matrices needed for energy derivatives . . . . .	225



## LIST OF FIGURES

1	Asymmetry in the 1st-order finite-difference polarizability . . . . .	16
2	Linear and 2D-IR of CO <sub>2</sub> in H <sub>2</sub> O . . . . .	38
3	FTIR of CO <sub>2</sub> in ionic liquids . . . . .	43
4	T-dependent FTIR of CO <sub>2</sub> in 1-butyl-3-methylimidazolium triflate . . . . .	45
5	ALMO decomposition of CO <sub>2</sub> $\nu_3$ frequency . . . . .	51
6	2D-IR spectra of CO <sub>2</sub> in ionic liquids . . . . .	56
7	CO <sub>2</sub> in IL bending mode kinetics . . . . .	57
8	Double-sided Feynman diagrams for 3rd and 5th order signals . . . . .	62
9	Experimental 2D-IR spectra of CO <sub>2</sub> in [Im <sub>4,1</sub> ][X] . . . . .	66
10	Example of global fitting of 2D-IR spectra . . . . .	68
11	Stochastic simulation spectrum . . . . .	71
12	$\tau_c$ -viscosity correlation with differing ionic liquid anion . . . . .	74
13	CLS overlaid on 2D-IR spectrum . . . . .	78
14	CLS fitting of 2D-IR spectrum . . . . .	79
15	Experimental 2D-IR spectra of CO <sub>2</sub> in all [Im <sub>4,1</sub> ][X] . . . . .	80
16	DVR contour plot . . . . .	93
17	Correlation of CO <sub>2</sub> $\nu_3$ frequencies with CT mechanisms . . . . .	103
18	Correlation of CO <sub>2</sub> $\nu_3$ frequencies with CP corrections . . . . .	104
19	Effect of QM and MM region size on CO <sub>2</sub> $\nu_3$ QM/MM frequencies . . . . .	112
20	DVR convergence with respect to number of QM ion pairs . . . . .	114
21	Spread of method and basis set dependence of MD snapshots . . . . .	121
22	Instantaneous method and basis set dependence of MD snapshots . . . . .	123

23	Self-correlation of QM/MM frequencies with increasing QM region size . . . . .	131
24	Correlation between $\mu_{01}$ and $\tilde{\nu}_3$ . . . . .	140
25	Spectroscopic map comparison with DVR . . . . .	142
26	Linewidth dependence on correlation time for fast motions . . . . .	148
27	Histograms of CO <sub>2</sub> $\tilde{\nu}_3$ from MD simulations . . . . .	149
28	Relationship between CO <sub>2</sub> $\tilde{\nu}_3$ and $1 + \cos \theta_{\text{OCO}}$ . . . . .	152
29	Correlation between DVR and <i>ab initio</i> transition dipole moments . . . . .	157
30	Distance dependence of analytic and numerical ALMO polarizabilities . . . . .	175
31	Distance dependence of CT restrictions on polarizabilities . . . . .	176
32	Distance dependence of $\alpha_{\parallel}$ for def2-SVP . . . . .	180
33	Distance dependence of $\alpha_{\parallel}$ for def2-SVPD . . . . .	181
34	Distance dependence of $\alpha_{\perp}$ for def2-SVP . . . . .	182
35	Distance dependence of $\alpha_{\perp}$ for def2-SVPD . . . . .	183
36	Location of PSI4NUMPY within the scientific computing ecosystem . . . . .	193
37	Extract from a PSI4NUMPY tutorial notebook . . . . .	202

**LIST OF ALGORITHMS**

1 Static linear response approach within fragment-localized formalism. . . . . 187  
2 Continuation of algorithm 1 . . . . . 188

## LIST OF LISTINGS

1	Sample Q-CHEM input file for “ALMO frz + pol” polarizability. Geometry is from HF/def2-SVPD. . . . .	179
2	Sample Q-CHEM input file for first-generation ALMO-EDA. Geometry is from HF/def2-SVPD. . . . .	184
3	Sample Q-CHEM input file for second-generation ALMO-EDA. Geometry is from HF/def2-SVP. . . . .	185
4	Basic example of PSI4 $\longleftrightarrow$ NUMPY interoperability . . . . .	194
5	Mutating a PSI4 Matrix directly from a NUMPY routine . . . . .	194
6	Initializing a PSI4 computation from Python . . . . .	195
7	Obtaining PSI4 wavefunction data in Python . . . . .	195
8	Computing atomic orbital-basis integrals from PSI4 via Python . . . . .	196
9	Calling PSI4’s AO-to-MO transformation from Python . . . . .	196
10	Building two-electron Fock matrix components using PSI4 via Python . . . . .	197
11	Performing an AO-to-MO transformation using NUMPY’s einsum . . . . .	198
12	Example of forming a coupled cluster intermediate using NUMPY and einsum	199

## PREFACE

I am surprised at how many people, places, and things are to blame for the creation of this dissertation. The first is my advisor, Daniel Secretary Lambrecht, who never got frustrated with me when I wasn't being productive, and mostly left me to my own devices, presumably because he didn't want to deal with me not knowing Vim and gnuplot. I will miss our discussions about how woodwind instruments are different from brass, how perfect the San Francisco weather is, how confusing the `NBasis/NBas6D/NB2/NB2car` mess is in Q-CHEM, and what directions quantum chemistry is and should be heading in. I promise to never use B3LYP/6-31G\*\* for anything ever again. More importantly, thank you for countering my occasionally poor communication skills and negative outlook with positive words of encouragement. I also thank Sean Garrett-Roe, who aside from always trying to make a collaboration with experimentalists as painless as possible, perhaps surpasses even me with fervent disgust toward Microsoft Word and was always a willing ear for my rants about computational chemists not using L<sup>A</sup>T<sub>E</sub>X. You are a model for how well-rounded a scientist should be that I strove to emulate. I would like to thank Ken Jordan for supervising me during my first summer at Pitt, and teaching me that some printers may prefer to work in atomic units.

I am indebted to Thomas Brinzer for all the time and effort he put into our collaboration, trying to get me to understand multidimensional spectroscopy, his always thoughtful critique of my work, and being a good friend. You are the model for how a graduate student should be that I strove to emulate.

I thank Evgeny Epifanovsky and Q-CHEM for giving me a two month ~~vacation~~internship in Pleasanton, which was the most valuable and enjoyable experience of my PhD.

I thank Daniel Smith for inviting me to a MolSSI workshop and for baiting me with

authorship on the `PSI4NUMPY` paper after watching me hack on `PSI4` in a Twitch stream.

As a strong believer in learning one's tools, I thank Emacs for making me faster at editing text than Ben Albrecht, who learned with me that people shouldn't share flash drives when printing posters. I thank Kevin Gasperich, Shiv Upadhyay, Amanda Dumi, and Daniel Burrill for joining me at weekly programming study groups that often got derailed as they learned how to push my buttons. I thank Keith Werling for being my original partner-in-crime as Daniel's other first student (though you are really the zeroth) and sharing my enjoyment of the Wabbajack. I thank Adam Gagorik for teaching me how to use Matplotlib and setting an example for getting work done as a graduate student. I thank Ali Sinan Sağlam for his sense of humor. I thank Daniel Bolt, Tyler Rohrs, and Halina Werner for tolerating some truly bad bowling strings while my mind was in other places, talking coffee, going to the symphony, and attending my own performances. I thank Forward Lanes (RIP), Row House Cinema, Hollywood Theater (RIP), the Cage, Voodoo Brewery, and Commonplace Coffee for providing places to go other than work.

I acknowledge Chemistry Stack Exchange for being incredibly distracting but also providing scientific material and Easter eggs that made it into this dissertation. In particular, I thank Brian Skinn and Zhe Lu for our lunch meetings and all the encouragement that I could make it out of academia. I thank cclib and particularly Karol Langer for allowing me to contribute to the project and helping me enter the world of open-source software development.

There are people from home who supported me as well. I thank Car Chat for coming to visit me and welcoming me every time I returned home, especially Brendan Reilly, who introduced me to anime and indirectly got me started playing music again. I thank Garen Chiloyan for listening to me complain about almost everything since the start of undergrad, sharing grad school gossip, and agreeing that Massachusetts really is the best. I thank Marshall Brennan for explaining what I should expect in grad school and being right about 99% of it.

I thank Adam Pratt and Tuguldur Togo Odbadrakh (Togie Stogie? Tuguldur Schmuğuldur? Togo Schmogo?) for our roller coaster ride, and I apologize for showing you my actual bad side.

My degree would not have been completed if not for the massive amount of music I listened to, primarily from Alfred Brendel, Francesco Corti, Passepied, and Judy Bedford, my original bassoon instructor who still inspires me. One of my biggest mistakes was not playing music for my first five years here, so I thank Brad Townsend, Barbara Hois, and Roger Zahab for taking a chance and encouraging me to play in all their ensembles. I am especially thankful to Linda Fisher for her patience while teaching me how to play and make reeds again.

Finally, but most importantly, I am eternally indebted to my parents, Gavra and John. Thank you for continuing to let me ask what's for dinner. You are my most passionate supporters, loudest critics, and best friends. Only you understand how difficult this was for me. I love you.

## 1.0 INTRODUCTION AND THEORETICAL BACKGROUND

A great strength of quantum chemistry is that it enables construction of molecular models using the building blocks of chemical intuition, such as interactions between functional groups, that can in theory be correlated with any information present in the wavefunction. Furthermore, the ability to calculate spectroscopic parameters from wavefunctions makes it possible to connect the features of a molecular model with molecular properties. This is of paramount importance to experimental spectroscopy, where there is not always an obvious “signature” that connects features in complex molecular systems to a distinct spectroscopic response. As a result, it is often not possible to draw conclusions from experimental results without the aid of quantum chemical calculations.

Finding correlation between molecular models and spectra is of great value, but there is also difficulty from the computational side in terms of identifying the *correct* molecular models and *why* they are correct in explaining experimental spectra. To form true structure-spectra relationships, the most desirable connection is through the building blocks of chemistry, functional groups, and identifying the impact of their interactions. A common approach in both the wet lab and computationally is to remove or modify functional groups in some systematic way, resulting in a qualitative connection between structural components and changes in spectra. The drawbacks of such chemical transformations are the inability to quantify the physics behind these chemically-intuitive changes, and potential side effects of perturbing a molecular model’s electronic and geometric structure due to cooperativity. One path to quantifying intermolecular interactions without such drastic chemical modifications is by fragmenting the system into groups of interest and decomposing their interaction energy into physically-motivated terms through a procedure called *energy decomposition analysis* (EDA). If EDA can be applied to spectra prediction, then the underlying causes



of spectral changes due to specific chemical motifs can be quantified. More importantly, spectra-property relationships can be formed, where changes in spectral response can be tied directly back to types of physical interactions, such as electrostatics, charge transfer, and dispersion.

The unifying theme of this work is that it is possible to identify the contribution of specific intermolecular interactions to spectroscopic response. The contributions of specific intermolecular interactions are in the language of energy decomposition analysis using absolutely localized molecular orbitals, abbreviated as ALMO-EDA. Details of the ALMO-EDA procedures are given in section 5.2.

The majority of this work presents applications of response theory to calculating spectroscopic properties, specifically vibrational frequencies in chapters 2, 3, 4 and dipole polarizabilities in chapter 5, where each property is decomposed in terms of contributions to the final response from distorting the molecular geometry, allowing the non-interacting molecular densities to interact and then relax, and finally allowing charge to flow unrestricted between molecules. Chapter 6 presents a new model for implementing quantum chemical methods, with the first dipole hyperpolarizability as an example (sections 6.9.1 and 6.9.2), and how the decomposition of molecular properties may be implemented and disseminated in the future.

The remainder of this introduction will cover the basic language of molecular response theory and its connection to macroscopic spectroscopic observables, with examples of which observables are related to which microscopic terms (section 1.2). It will also cover more general cases of how those microscopic terms appear in different forms of derivation, all of which are related and give identical final answers. Specifically, a connection will be made between phenomenological Hamiltonians (section 1.3.1), series expansions (section 1.3.2), energy derivatives (sections 1.3.3, 1.3.4, and 1.3.5), perturbation theory (section 1.3.6), and quasienergy derivatives (section 1.4), of which the latter two are appropriate for incorporation of time dependence, leading to dynamic properties.

For more background literature on the theoretical development and applications of response theory to molecular properties, see reviews by Gauss,<sup>1</sup> Neese,<sup>2</sup> Norman,<sup>3</sup> Helgaker<sup>4</sup> and books by Szabo & Ostlund,<sup>5</sup> McWeeny,<sup>6</sup> Yamaguchi,<sup>7</sup> and Barron.<sup>8</sup>

## 1.1 TERMS AND FUNDAMENTAL DEFINITIONS

Throughout the introduction and the remainder of this work, a number of terms appear and may seem to be used interchangeably. In quantum chemistry literature, the phrase *molecular property* is used to denote any quantity that can be calculated from the wavefunction. By this definition, not all molecular properties are *observables*, as they do not necessarily correspond to experimentally-measurable quantities. Such molecular properties may not also be uniquely defined. One of the most prominent examples of non-observable molecular properties is the calculation of atomic partial charges, with Mulliken<sup>5,9</sup> and Löwdin<sup>10</sup> population analyses being the two most commonly-known partitioning schemes for dividing the electron density into atomic contributions. However, this work associates molecular properties with observables, which may also be synonymous with *physical properties*. The term *chemical property* is avoided as it is more closely associated with reactivity than observable quantities.

The term *response* also requires disambiguation. Experimentally, *molecular response* refers to changes in a molecule’s electronic and geometric state due to incident electromagnetic radiation (spectroscopy) or physical manipulation (such as structural deformation leading to piezoelectric response).<sup>11–13</sup> Computationally, within the Born-Oppenheimer approximation, molecular response encompasses how the electronic state is altered due to both external and internal perturbations. Examples of external perturbations are electric or magnetic fields whose strengths may or may not fluctuate with time, and examples of internal perturbations are nuclear displacements, nuclear magnetic moments, and the total electronic spin. Although external perturbations have a clear correspondence to spectroscopy, internal perturbations are more a description of a molecule’s intrinsic structure.

However, in the context of this work, molecular response is sometimes used interchangeably with *response property*, which is intuitively a property associated with external applied fields, but here refers to any molecular property arising from the solution of the *response equations* or *response functions*.

The response equations relevant for this work are the coupled perturbed Hartree–Fock (CPHF), coupled perturbed Kohn–Sham (CPKS), or coupled perturbed self-consistent field (CPSCF) equations. These terms may be used interchangeably, as the equations are struc-

turally identical, with the only difference being the machinery of how the exchange-like terms are formed. These sets of equations are similar to those from time-dependent HF and KS (TDHF and TDKS/TDDFT) theory. In the TD equations, a non-Hermitian eigenvalue problem is solved for  $(\mathbf{G} - \omega\mathbf{\Delta})\mathbf{U} = \mathbf{0}$ , where  $\mathbf{G}$  is the orbital Hessian, the eigenvectors  $\mathbf{U}$  describe the transitions between the states of the system in the given molecular orbital (MO) basis, and the eigenvalues  $\omega$  are the system’s excitation energies (poles)<sup>a</sup>. In the CP equations, a set of linear equations are solved for  $\mathbf{G}\mathbf{U} = -\mathbf{V}$ , where  $\mathbf{V}$  represents the perturbation, and  $\mathbf{U}$  describes the modification to the unperturbed state’s orbitals caused by the perturbation. A basic implementation of the CPHF equations is given in section 6.9.1 using the electric dipole operator as the perturbation. Although the remainder of this work will be concerned with the properties arising from solution of the CP equations and not the TD equations, their combination gives rise to other important molecular properties, namely the residues of the response function (eigenvectors from the TD equation contracted with the operator of interest), describe the perturbation-mediated transitions between states (their moments).

The solution of the CPSCF equations to calculate the parameters that describe the wavefunction’s response to the perturbation  $\hat{V}$  determines the *linear response* of the system. Linear response is named as such not due to the linear form of the equations, but because the perturbation is linear in strength. This strength may be constant, static, or time-independent, or it may be oscillating, dynamic, frequency-, or time-dependent. This work is primarily concerned with static response properties, however an introduction to dynamic response is presented. The most common formulation of these equations is single-particle in nature, with both particle-hole and hole-particle terms, giving the *random phase approximation* (RPA).<sup>14,15</sup> If not the external perturbation but the response equations themselves are expanded using order analysis, the Hartree–Fock equations give the zeroth-order *polarization propagator*, RPA corresponds to the first-order polarization propagator, and the full two-particle terms are part of the second-order polarization propagator (SOPPA).<sup>16</sup>

---

<sup>a</sup> $\mathbf{\Delta}$  is an identity matrix, except the lower diagonal is negated, see (Neese 107).

## 1.2 CONNECTION BETWEEN MACROSCOPIC AND MICROSCOPIC PROPERTIES

There is often a direct connection between the macroscopic observables measurable by spectroscopic techniques and the molecular response calculated at the microscopic scale. Tables 1 and 2 give examples of the most common molecular properties of interest and their relationships to energy derivatives and response functions, respectively. Although the effort required for implementing general energy derivatives may be considerable, the entries in both tables are the true starting points for molecular property calculations based on quantum chemical wavefunctions.

### 1.3 STATIC (TIME-INDEPENDENT) RESPONSE PROPERTIES

The two primary ways to perform the derivation are

1. from a phenomenological Hamiltonian, similarly to using the correspondence principle when quantizing a classical expression (section 1.3.1), or
2. from series expansion of the energy with respect to one or more perturbations (section 1.3.2).

It is also possible to obtain expressions for static properties by the reduction of any expressions for dynamic properties to the static limit (zero frequency:  $\omega = 0$ ). The purpose of this section is to avoid some complexity in the derivation of time-dependent response by understanding the simpler case of static response first.

#### 1.3.1 Phenomenological approach

For a one-dimensional spring connecting a ball to a fixed object, Hooke's law is

$$F = -kx, \tag{1.1}$$

Table 1: Connection between specific energy derivatives and their respective molecular properties.  $F$  is an applied electric field,  $B$  is an applied magnetic field,  $X$  is a nuclear coordinate,  $m$  is a nuclear magnetic moment,  $J$  is a total rotational moment,  $I$  is a nuclear spin, and  $S$  is the intrinsic electronic spin. Adapted from Ref. [1] and [17].

Energy derivative	Molecular property
$\frac{dE}{dF_i}$	dipole moment; in a similar manner also multipole moments, electric field gradients, etc.
$\frac{dE}{dB_\alpha}$ $\frac{dE}{dX_i}$	magnetic dipole moment and higher-order magnetic multipoles forces on nuclei; stationary points on potential energy surfaces, equilibrium and transition state structures
$\frac{dE}{dm_{K_j}}$	spin density; hyperfine interaction constants
$\frac{d^2E}{dF_\alpha dF_\beta}$	(electric) polarizability
$\frac{d^2E}{dX_i dX_j}$	harmonic force constants; harmonic vibrational frequencies
$\frac{d^2E}{dX_i dF_\alpha}$	dipole derivatives; infrared intensities within the harmonic approximation
$\frac{d^2E}{dB_\alpha dB_\beta}$	magnetizability
$\frac{d^2E}{dm_{K_i} dB_\alpha}$	nuclear magnetic shielding tensor; relative NMR shifts
$\frac{d^2E}{dI_{K_i} dI_{L_j}}$	indirect spin-spin coupling constant
$\frac{d^2E}{dB_\alpha dJ_\beta}$	rotational $g$ -tensor; rotational spectra in magnetic field
$\frac{d^2E}{dI_{K_i} dB_\alpha}$	nuclear spin-rotation tensor; fine structure in rotational spectra
$\frac{d^2E}{dS_i dB_\alpha}$	electronic $g$ -tensor
$\frac{d^3E}{dX_i dF_\alpha dF_\beta}$	polarizability derivative; Raman intensities
$\frac{d^3E}{dF_\alpha d^2F_\beta}$	(first) electric hyperpolarizability
$\frac{d^3E}{dX_i dX_j dX_k}$	cubic force constants; vibrational corrections to distances and rotational constants
$\frac{d^4E}{dF_\alpha dF_\beta dF_\gamma dF_\delta}$	(second) electric hyperpolarizability
$\frac{d^4E}{dB_\alpha dB_\beta dB_\gamma dB_\delta}$	(second) hypermagnetizability
$\frac{d^4E}{dX_i dX_j dX_k dX_l}$	quartic force constants; anharmonic corrections to vibrational frequencies
$\frac{d^4E}{dF_\alpha dF_\beta dF_\gamma dX_i}$	hyper-Raman effects
$\frac{d^4E}{dF_\alpha dF_\beta dX_i dX_j}$	Raman intensities for overtone and combination bands
$\frac{d^4E}{dF_\alpha dF_\beta dB_\gamma dB_\delta}$	Cotton–Mutton effect

Table 2: Connection between specific (non)linear response functions and their respective molecular properties. Adapted from Ref. [3].

<b>Molecular Property</b>	<b>Definition</b>	<b>Type of response function</b>
polarizability	$\langle\langle\hat{\mu}; \hat{\mu}\rangle\rangle_{\omega}$	linear
magnetizability	$\langle\langle\hat{m}; \hat{m}\rangle\rangle_0$	linear
optical rotation	$\langle\langle\hat{\mu}; \hat{m}\rangle\rangle_{\omega}$	linear
electronic circular dichroism	$\langle\langle\hat{\mu}; \hat{m}\rangle\rangle_{\omega_f}$	single residue of linear
IR intensities	$\langle\langle\hat{\mu}; \partial\hat{H}_0/\partial R\rangle\rangle_{\omega}$	linear
NMR spin-spin coupling constants	$\langle\langle\hat{h}_{SD}; \hat{h}_{SD}\rangle\rangle_0,$ $\langle\langle\hat{h}_{FC}; \hat{h}_{FC}\rangle\rangle_0,$ $\langle\langle\hat{h}_{PSO}; \hat{h}_{PSO}\rangle\rangle_0$	linear linear linear
NMR chemical shifts	$\langle\langle\hat{l}_O; \hat{h}_{PSO}\rangle\rangle_0$	linear
EPR $g$ -tensor	$\langle\langle\hat{l}_O; \hat{h}_{SOC}\rangle\rangle_0$	linear
static first hyperpolarizability	$\langle\langle\hat{\mu}; \hat{\mu}, \hat{\mu}\rangle\rangle_{0,0}$	quadratic
second-harmonic generation	$\langle\langle\hat{\mu}; \hat{\mu}, \hat{\mu}\rangle\rangle_{\omega,\omega}$	quadratic
electro-optical Pockels effect	$\langle\langle\hat{\mu}; \hat{\mu}, \hat{\mu}\rangle\rangle_{\omega,0}$	quadratic
optical rectification	$\langle\langle\hat{\mu}; \hat{\mu}, \hat{\mu}\rangle\rangle_{\omega,-\omega}$	quadratic
Faraday rotation	$\langle\langle\hat{\mu}; \hat{\mu}, \hat{m}\rangle\rangle_{\omega,0}$	quadratic
magnetic circular dichroism	$\langle\langle\hat{\mu}; \hat{\mu}, \hat{m}\rangle\rangle_{\omega_f,0}$	single residue of quadratic
Raman intensities	$\langle\langle\hat{\mu}; \hat{\mu}, \partial\hat{H}_0/\partial R\rangle\rangle_{\omega,0}$	quadratic
static second hyperpolarizability	$\langle\langle\hat{\mu}; \hat{\mu}, \hat{\mu}, \hat{\mu}\rangle\rangle_{0,0,0}$	cubic
third-harmonic generation	$\langle\langle\hat{\mu}; \hat{\mu}, \hat{\mu}, \hat{\mu}\rangle\rangle_{\omega,\omega,\omega}$	cubic

where  $x$  is the displacement of the spring from equilibrium in meters,  $k$  is the spring constant in units of  $\text{N m}^{-1}$ , and  $F$  is the restoring force in units of newtons acting on the displaced spring by the object it is attached to. If the sign is reversed, then the equation can be viewed as describing the force of spring acting on the attached object; it is a direction change and a matter of convention.

Hooke's law can be generalized to multiple dimensions. For example, in three-dimensional space it can be written as

$$\begin{bmatrix} F_1 \\ F_2 \\ F_3 \end{bmatrix} = - \begin{bmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad (1.2)$$

which can represent either a single object or three one-dimensional springs. It can also be made  $3N$ -dimensional when describing the forces on  $N$  atoms (each with 3 Cartesian components) given their relative positions  $\mathbf{x}$  and the "stiffness" of their connectivity  $\mathbf{k}$ . In the most general  $N$ -dimensional form, it can be written as

$$F_i = - \sum_j^N k_{ij} x_j, \quad (1.3)$$

or in Einstein summation notation where repeated indices are implicitly summed (contracted) over as

$$F_i = -k_{ij} x_j, \quad (1.4)$$

where both  $i, j$  range from 1 to  $N$ , leading to vectors  $\mathbf{F}$  and  $\mathbf{x}$  of shape  $[N]$  and a matrix  $\mathbf{k}$  of shape  $[N, N]$ . From (1.2), (1.3), and (1.4), if off-diagonal elements of  $k$  are allowed to be nonzero, there can be coupling between springs. For example, if  $i = 1$ , in the case of coupling,

$$F_1 = -(k_{11}x_1 + k_{12}x_2 + k_{13}x_3), \quad (1.5)$$

which reduces to

$$F_1 = -k_{11}x_1 \quad (1.6)$$

in the absence of coupling, or the same result obtained in (1.1).

The force is also related to the energy. In one dimension,

$$F = -\nabla E \quad (1.7)$$

$$= -\frac{\partial E}{\partial x}, \quad (1.8)$$

where  $\nabla \equiv \partial/\partial x$ . In  $N$  dimensions, like (1.4), (1.7) is (using vector calculus)

$$F_i = -\frac{\partial E}{\partial x_i}. \quad (1.9)$$

Equating the  $F_i$  in (1.4) and (1.9) gives

$$-k_{ij}x_j = -\frac{\partial E}{\partial x_i}, \quad (1.10)$$

where the negative signs can be dropped. To solve for the stiffness coefficients, take the partial derivative of both sides with respect to  $x_j$ , using the product rule on the left hand side:

$$\left(\frac{\partial}{\partial x_j}\right)(k_{ij}x_j) = \left(\frac{\partial}{\partial x_j}\right)\left(\frac{\partial E}{\partial x_i}\right) \quad (1.11)$$

$$\left[\left(\frac{\partial}{\partial x_j}\right)(k_{ij})\right]x_j + k_{ij}\left[\left(\frac{\partial}{\partial x_j}\right)x_j\right] = \frac{\partial^2 E}{\partial x_j \partial x_i} \quad (1.12)$$

$$k_{ij} = \frac{\partial^2 E}{\partial x_j \partial x_i}. \quad (1.13)$$

This tells that the internal stiffness is related to the second derivative of the energy with respect to nuclear coordinate displacements. The internal stiffness matrix is the molecular Hessian, where each “spring constant” is called a force constant, which describes how a change or perturbation to one atomic coordinate affects a change in another atomic coordinate.

Another important property is that in general, due to Young’s theorem, the order of differentiation is not important, and the perturbations may be interchanged:

$$\frac{\partial^2 E}{\partial x_j \partial x_i} = \frac{\partial^2 E}{\partial x_i \partial x_j}, \quad (1.14)$$

leading to a symmetric matrix  $\mathbf{k}$ . In practice, this has implications for computational cost which will be discussed in section 1.3.5.



A similar derivation holds for the dipole polarizability,  $\alpha$ , which is the ratio of the induced dipole moment  $\mu$  of a system to the electric field  $F$  that produces this dipole moment. Both  $\mu$  and  $F$  are 3-dimensional vector quantities,

$$\vec{\mu} = \vec{\alpha} \cdot \vec{F}, \quad (1.15)$$

which can be expanded identically to (1.2) as

$$\begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix} \begin{bmatrix} F_1 \\ F_2 \\ F_3 \end{bmatrix}, \quad (1.16)$$

or in Einstein summation notation as

$$\mu_i = \alpha_{ij} F_j. \quad (1.17)$$

This is the most general case, where anisotropy may be present in the polarizability tensor, leading to nonzero off-diagonal elements.

Another definition of the molecular dipole moment induced by an external (applied) electric field is

$$\mu_i = -\frac{\partial E}{\partial F_i}, \quad (1.18)$$

which originates from the energy of a neutral dipole in an electric field,

$$E = -\mu_i F_i. \quad (1.19)$$

Following (1.10), equating (1.17) and (1.18) gives

$$\alpha_{ij} F_j = -\frac{\partial E}{\partial F_i} \quad (1.20)$$

The remaining steps follow identically to (1.11), (1.12), and (1.13):

$$\left( \frac{\partial}{\partial F_j} \right) (\alpha_{ij} F_j) = \left( \frac{\partial}{\partial F_j} \right) \left( -\frac{\partial E}{\partial F_i} \right) \quad (1.21)$$

$$\cancel{\left( \frac{\partial \alpha_{ij}}{\partial F_j} \right)} F_j + \alpha_{ij} \cancel{\left( \frac{\partial F_j}{\partial F_j} \right)} = -\frac{\partial^2 E}{\partial F_j \partial F_i} \quad (1.22)$$

$$\alpha_{ij} = -\frac{\partial^2 E}{\partial F_j \partial F_i} \quad (1.23)$$

### 1.3.2 Series expansion

More generally, the derivative terms in section 1.3.1 originate from series expansions of the energy in the presence of a perturbation; for (1.1), it is internal geometric displacements, and in (1.15) it is an external electric field. The energy in the presence of an arbitrary perturbation  $P$  is written as

$$\begin{aligned}
 E(P) &= \sum_{n=0}^{\infty} \frac{1}{n!} \left. \frac{\partial^n E}{\partial P^n} \right|_{P=a} \cdot (P-a)^n \\
 &= \sum_{n=0}^{\infty} \frac{1}{n!} E^{(n)}(a) \cdot (P-a)^n \\
 &= E^{(0)}(a) + E^{(1)}(a) \cdot (P-a) + \frac{1}{2} E^{(2)}(a) \cdot (P-a)^2 + \frac{1}{6} E^{(3)}(a) \cdot (P-a)^3 + \dots,
 \end{aligned} \tag{1.24}$$

where  $a$  is the point at which the derivative is taken. Choosing  $a \stackrel{!}{=} 0$  (expanding around the perturbation at zero strength) turns the Taylor series into a Maclaurin series<sup>b</sup>:

$$\begin{aligned}
 E(P) &= \sum_{n=0}^{\infty} \frac{1}{n!} \left. \frac{\partial^n E}{\partial P^n} \right|_{P=0} \cdot P^n \\
 &= \sum_{n=0}^{\infty} \frac{1}{n!} E^{(n)} \cdot P^n \\
 &= E^{(0)} + E^{(1)} \cdot P + \frac{1}{2} E^{(2)} \cdot P^2 + \frac{1}{6} E^{(3)} \cdot P^3 + \dots
 \end{aligned} \tag{1.25}$$

The perturbation  $P$  may have multiple components. For example, there are 3 possible Cartesian components to an external electric field  $\mathbf{F} = (F_x, F_y, F_z)$  and  $3N$  atomic coordinates. Generalizing the dimensionality of  $P$  to  $k$  and inserting into (1.25) gives

$$E(\mathbf{P}) = \sum_{n=0}^{\infty} \frac{1}{n!} \left. \frac{\partial^n E}{\partial \mathbf{P}^n} \right|_{\mathbf{P}=\mathbf{0}} \cdot \mathbf{P}^n \tag{1.26}$$

Considering specific examples, using (1.26), replacing  $\mathbf{P}$  with an external electric field  $\mathbf{F}$ , and switching to Einstein notation gives

$$E(\mathbf{F}) = E_0 - \mu_i \cdot F_i - \frac{1}{2} \alpha_{ij} \cdot F_i F_j - \frac{1}{6} \beta_{ijk} \cdot F_i F_j F_k - \frac{1}{24} \gamma_{ijkl} \cdot F_i F_j F_k F_l - \dots, \tag{1.27}$$

---

<sup>b</sup>The notation  $x \stackrel{!}{=} y$  means that  $x$  must be equal to  $y$  by definition.

where  $\mu_i$  is a component of the dipole (moment), expressed in operator form as

$$\hat{\mu} = (\hat{\mu}_x, \hat{\mu}_y, \hat{\mu}_z) = -e(\hat{x}, \hat{y}, \hat{z}), \quad (1.28)$$

$\alpha$  is a component of the polarizability,  $\beta$  is a component of the first hyperpolarizability,  $\gamma$  is a component of the second hyperpolarizability, etc., each describing an additional correction to how a system interacts with the external electric field.

It is also possible to consider multiple perturbations simultaneously. Adding another perturbation  $\mathbf{Q}$  to (1.26) gives

$$E(\mathbf{P}, \mathbf{Q}) = E^{(0,0)} + (E^{(1,0)} \cdot \mathbf{P} + E^{(0,1)} \mathbf{Q}) + \frac{1}{2} (E^{(2,0)} \cdot \mathbf{P}^2 + E^{(0,2)} \cdot \mathbf{Q}^2 + E^{(1,1)} \cdot \mathbf{P} \cdot \mathbf{Q}) + \dots, \quad (1.29)$$

where  $E^{(n,m)}$  refers to the energy correction that is simultaneously  $n$ th-order in the perturbation  $\mathbf{P}$  and  $m$ th-order in the perturbation  $\mathbf{Q}$ . The expected terms from both independent series expansions occur, but there is also a cross-term  $E^{(1,1)}$ . All mixed derivatives in Table 1 that are at least 2nd-order total correspond to such cross-terms. For example, consider adding an external magnetic field to (1.27):

$$E(\mathbf{F}, \mathbf{B}) = E_0 - \mu_i \cdot F_i - m_i \cdot B_i - \frac{1}{2} (\alpha_{ij} \cdot F_i F_j + \xi_{ij} \cdot B_i B_j + G_{ij} \cdot F_i B_j) - \dots, \quad (1.30)$$

where the magnetic field has introduced  $m_i$  as a component of the magnetic (dipole) moment,  $\xi_{ij}$  as a component of the magnetizability, and the mixed electric dipole–magnetic dipole polarizability  $G_{ij}$ , which is directly related to the optical rotation<sup>c</sup>, given as  $\langle\langle \hat{\mu}; \hat{m} \rangle\rangle_\omega$  in Table 2. The residues of the same response function give the rotatory strengths needed for electronic circular dichroism (ECD), shown as  $\langle\langle \hat{\mu}; \hat{m} \rangle\rangle_{\omega_f}$  in Table 2. Section 1.4 will show how frequency dependence can be properly introduced to this term and in general.

<sup>c</sup>See Eq. (2) in Ref. [18], specifically the  $G$  term.

### 1.3.3 Derivative evaluation

Up to this point, it has not been necessary to specify which energy expression is being differentiated: it is the energy expression for the chosen method (HF,  $\omega$ B97M-V, MP2, CCSD(T), ...). These derivatives may be evaluated in one of two ways:

1. numerically, by using a finite difference expression (usually based on central differences) for the desired derivative order, evaluating the energy (or some other property) at multiple perturbation strengths (step sizes) and directions, or
2. analytically, by differentiating the energy expression “on paper” and evaluating it directly.

It is also possible to combine numerical and analytic approaches to obtain higher-order derivatives. For example, in the calculation of Raman intensities, defined as  $\frac{\partial^3 E}{\partial X_A \partial F_i \partial F_j}$ , there are six unique ways to perform the derivative, shown in Table 3. More explicitly, the first row

Table 3: Possible permutations of analytic (a) and numerical (n) differentiation for each perturbation term in calculating Raman intensities. The two cancelled entries are not unique due to permutational symmetry.

$X_A$	$F_i$	$F_j$
a	a	a
a	a	n
<del>a</del>	<del>a</del>	<del>a</del>
n	a	a
a	n	n
n	a	n
<del>n</del>	<del>n</del>	<del>n</del>
n	n	n

corresponds to a fully analytic third derivative, the next three rows correspond to first-order finite difference of analytic second derivatives, the next three rows correspond to second-order finite difference of analytic first derivatives, and the last row corresponds to the third-order finite difference of energies. Due to symmetry in the electric field perturbation indices, two of the eight permutations are identical to others already present; for example, a/a/n and a/n/a are functionally identical.

In practice, the fully analytic third derivative is often not implemented, but second derivatives are, leading to two unique possibilities:

$$\frac{\partial^3 E}{\partial X_A \partial F_i \partial F_j} = \underbrace{\frac{\partial}{\partial X_A}}_{\text{numeric}} \underbrace{\left( \frac{\partial^2 E}{\partial F_i \partial F_j} \right)}_{\text{analytic}} = \frac{\partial \alpha_{ij}}{\partial X_A}, \quad (1.31)$$

$$= \underbrace{\frac{\partial}{\partial F_i}}_{\text{numeric}} \underbrace{\left( \frac{\partial^2 E}{\partial X_A \partial F_j} \right)}_{\text{analytic}} = \frac{\partial}{\partial F_i} \left( \frac{\partial \mu_j}{\partial X_A} \right). \quad (1.32)$$

In (1.31) there are  $2 \times (3N \text{ atomic coordinates}) = 6N$  atom-displaced polarizability calculations. This is closer to the textbook definition of Raman intensities, which are the change in molecular polarizability along each normal mode coordinate.<sup>19</sup> In (1.32), the dipole gradient (needed for IR intensities) is calculated analytically for  $2 \times (3 \text{ field directions}) = 6$  finite electric field calculations.

### 1.3.4 Finite difference for numerical derivatives

To perform numerical differentiation for molecular properties, first consider the definition of a (first) derivative,

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}, \quad (1.33)$$

where  $f(x)$  is the function of an independent variable  $x$  being differentiated with respect to that variable. If  $h$  is set to a small finite number ( $h > 0$ ),

$$f'(x) \approx f'(x|h) = \frac{f(x+h) - f(x)}{h}, \quad (1.34)$$

the exact (analytic) derivative is approximated using a step size  $h$ . (1.34) is the *forward* (finite) difference, as the step is taken by incrementing the independent variable. More common is *central* difference,

$$f'(x|h) = \frac{f(x + \frac{1}{2}h) - f(x - \frac{1}{2}h)}{h}. \quad (1.35)$$

Replace  $f$  with the molecular energy, and let  $h$  be the strength of an applied electric field along the  $z$ -direction. (1.35) becomes

$$\mu_z(h_z) = \frac{E(\frac{1}{2}h_z) - E(-\frac{1}{2}h_z)}{h_z}, \quad (1.36)$$

the  $z$ -component of the electric dipole moment.  $x$  disappears because the derivative is taken at zero field. (1.36) may be useful for methods that do not commonly have analytic derivatives of any order, such as CCSDT. Of more interest is replacing the energies with analytic dipole moments to give an element of the polarizability tensor:

$$\alpha_{xz}(h_x) = \frac{\mu_z(\frac{1}{2}h_x) - \mu_z(\frac{1}{2}h_x)}{h_x}. \quad (1.37)$$

(1.37) is an example of a mixed analytic/numerical derivative as discussed in section 1.3.3. A fully-numeric polarizability calculation would require applying 2nd-order central difference,

$$f''(x|h) = \frac{f(x+h) - 2f(x) - f(x-h)}{h^2}, \quad (1.38)$$

to give

$$\alpha_{zz}(h_z) = \frac{E(h_z) - 2E_0 - E(h_z)}{h_z^2}. \quad (1.39)$$

As discussed in section 5.3.3 and Ref. [1], there are considerable disadvantages to performing numerical differentiation of the wavefunction. These include the presence of finite-difference error due to step size sensitivity, the inability to handle frequency-dependent perturbations, the inability to handle response to applied magnetic fields without complex energies, and the poor calculation time scaling (especially for geometric derivatives). A particularly insidious example of error related to the step size is given in Figure 1. Many tests of numerical derivative errors consider relative or absolute accuracy in comparison to the exact analytic result. Figure 1 does not compare the error between the numerical and analytic results, but between matrix elements of the numerical result. In this case, the fundamental structure of the electric field response is violated: asymmetry implies that the polarizability is dependent on the order of which the probing electric fields are applied, and (1.14) does not hold, giving an unphysical result.

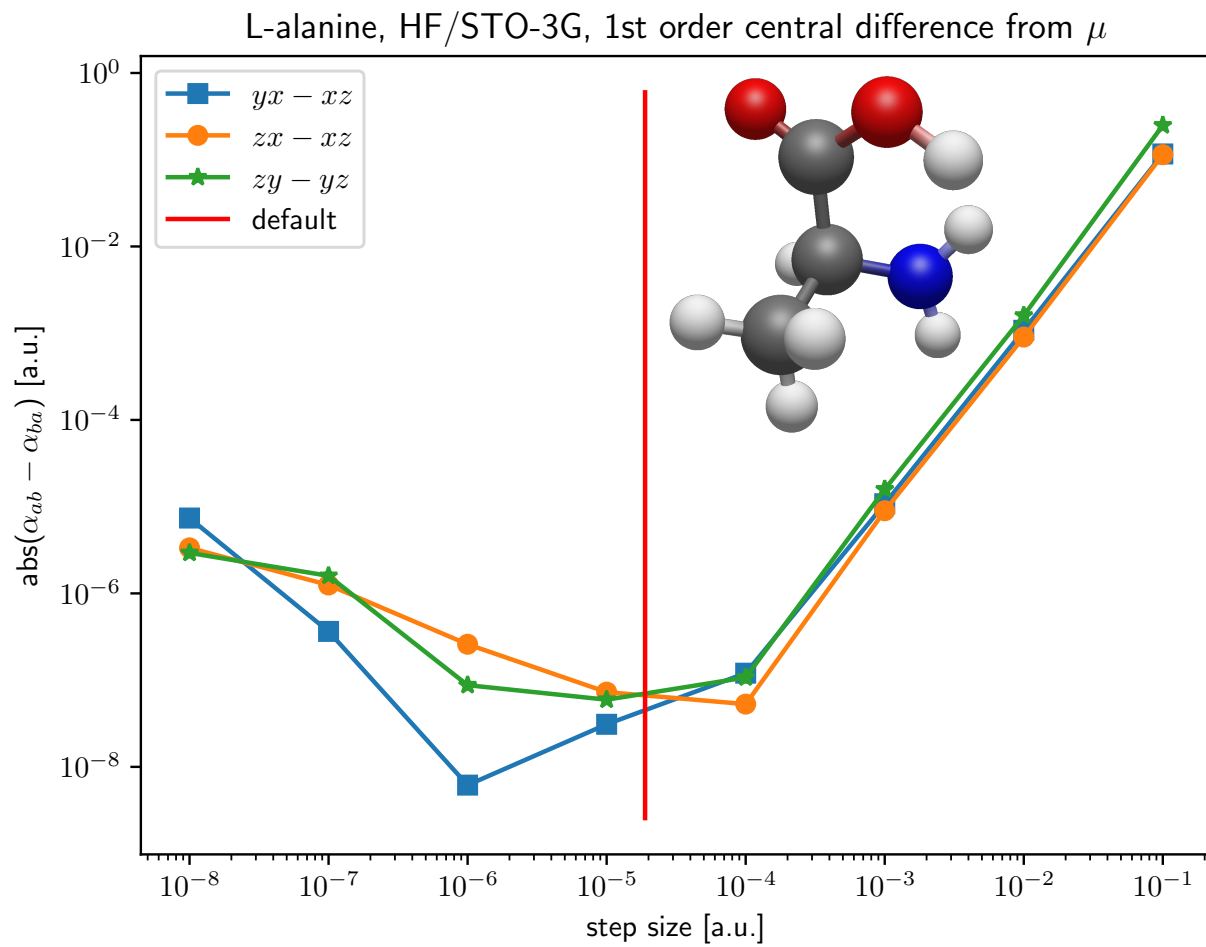


Figure 1: Effect of numerical noise on the off-diagonal matrix elements of the polarizability tensor. Nonzero differences indicate asymmetry, and the polarizability tensor is supposed to be symmetric. The red bar indicates the default step size for the applied electric field in Q-CHEM 5.1, set at  $1.88973 \times 10^{-5}$  a.u..

### 1.3.5 Analytic derivative theory

As mentioned in section 1.3.3, the first requirement for evaluating analytic energy derivatives is to form the necessary mathematical expression. In the most general case, there are both derivatives of the atomic orbital (AO) basis integrals themselves and of the density matrix, which leads to derivatives of the MO coefficients. To illustrate some of the mechanics of differentiation, consider the derivative of the MO-basis matrix representation of the electron-nuclear attraction operator  $\hat{V}_{eN}$  with respect to a nuclear coordinate  $X_A$ , which is a term needed for the molecular gradient:

$$\frac{\partial V_{ij}}{\partial X_A} = \frac{\partial}{\partial X_A} \left( \sum_{\mu\nu}^{\text{AO}} C_{\mu i} C_{\nu j} V_{\mu\nu} \right) \quad (\text{Yamaguchi eq. 3.80})$$

$$= \sum_{\mu\nu}^{\text{AO}} \left( \frac{\partial C_{\mu i}}{\partial X_A} C_{\nu j} V_{\mu\nu} + C_{\mu i} \frac{\partial C_{\nu j}}{\partial X_A} V_{\mu\nu} + C_{\mu i} C_{\nu j} \frac{\partial V_{\mu\nu}}{\partial X_A} \right), \quad (\text{Yamaguchi eq. 3.81})$$

where the third (last) term is the true AO integral derivative, and the first two terms, the MO coefficient derivatives, come from differentiating the density matrix, which is defined as

$$P_{\mu\nu}^{\text{RHF}} = \sum_i^{\text{d.o.}} C_{\mu i} C_{\nu i} \quad (1.40)$$

in the AO basis.

The AO integral derivative can be further expanded. Using  $\mu, \nu$  rather than  $\chi_\mu, \chi_\nu$  so they refer to both AO basis functions *and* their matrix indices,

$$\frac{\partial V_{\mu\nu}}{\partial X_A} = \frac{\partial}{\partial X_A} \langle \mu | \hat{V} | \nu \rangle \quad (\text{Yamaguchi eq. 3.24})$$

$$= \left\langle \frac{\partial \mu}{\partial X_A} | \hat{V} | \nu \right\rangle + \left\langle \mu \left| \frac{\partial \hat{V}}{\partial X_A} \right| \nu \right\rangle + \left\langle \mu | \hat{V} \left| \frac{\partial \nu}{\partial X_A} \right. \right\rangle, \quad (\text{Yamaguchi eq. 3.25})$$

where the first and third terms are derivatives of basis functions and the second term is a derivative of the operator itself. Although AO integral derivatives are a necessary component of most derivative expressions, they do not play a direct role in response equations, and do not need to be discussed further.



It is convenient to rewrite the MO coefficient derivatives,

$$\frac{\partial C_{\mu i}}{\partial X_A} = \sum_m^{\text{MO}} U_{mi}^{X_A} C_{\mu m}, \quad (\text{Yamaguchi eq. 3.7})$$

where the index  $m$  runs over all occupied and unoccupied/virtual MOs. Although  $X_A$  is being used for the perturbation, all parts of this derivation hold for any general perturbation. The key insight is that the effect of a perturbation on the MO coefficients can be written as the contraction of the unmodified MO coefficients with a unitary matrix describing single-particle excitations from occupied to virtual MOs, as well as deexcitations from virtual to occupied MOs. In matrix form, this is

$$\mathbf{C}^{(X_A)} = \mathbf{C}^{(0)} (\mathbf{U}^{(X_A)})^T, \quad (1.41)$$

where the dimension of  $\mathbf{U}$  is  $[N_{\text{orb}}, N_{\text{orb}}]$ .

Now consider the derivatives of the MO coefficients/density matrix in the context of the Hartree–Fock equations. Starting from the restricted Hartree–Fock electronic energy expression,<sup>d</sup>

$$E_{\text{elec}}^{\text{RHF}} = 2 \sum_i^{\text{d.o.}} h_{ii} + \sum_{ij}^{\text{d.o.}} \{2(ii|jj) - (ij|ij)\}, \quad (\text{Yamaguchi eq. 4.1})$$

the first derivative with respect to a nuclear displacement  $X_A$  is<sup>e</sup>

$$\frac{\partial E_{\text{elec}}^{\text{RHF}}}{\partial X_A} = 2 \sum_i^{\text{d.o.}} h_{ii}^{X_A} + \sum_{ij}^{\text{d.o.}} \{2(ii|jj)^{X_A} - (ij|ij)^{X_A}\} - 2 \sum_i^{\text{d.o.}} S_{ii}^{X_A} \epsilon_i, \quad (\text{Yamaguchi eq. 4.21})$$

where terms with the superscript  $X_A$  indicate the a derivative of only the AO term.

Notice that the MO coefficient derivatives do not appear in the final HF gradient expression. They disappear due to Wigner’s  $2n + 1$  rule. From page 25 of Ref. [7]:

---

<sup>d</sup> $\mu, \nu, \lambda, \sigma, \dots$  are AO indices,  $i, j, k, l, \dots$  are occupied MO indices, and  $p, q, r, s, \dots$  are general MO indices. ( $\hat{h} = \hat{H}^{\text{core}}$ )  $\equiv \hat{T}_e + \hat{V}_{N_e}$  is the one-electron core Hamiltonian operator, which itself is the sum of the electronic kinetic energy and electron-nuclear attraction energy operators, here in matrix representation, similarly to (Yamaguchi eq. 3.80).  $(pq|rs)$  is an MO-basis two-electron (repulsion) integral in Mulliken notation (see Table 2.2 of Ref. [5]).

<sup>e</sup>Additionally, see section C.3 of Szabo & Ostlund<sup>5</sup> and Ref. [20].

When the wavefunction is determined up to the  $n$ th order, the expectation value (electronic energy) of the the system is resolved, according to the results of perturbation theory, up to the  $(2n + 1)$ st order. This principle is called Wigner's  $2n + 1$  theorem.<sup>21,22</sup>

More explicitly, we have the zeroth-order wavefunction, so we must be able to calculate the first-order correction to the energy. Worded differently, any first derivative of the energy can be calculated without differentiating MO coefficients, which is only required for second derivatives, such as the molecular Hessian or the dipole polarizability.

Differentiating (Yamaguchi eq. 4.1) with respect to  $X_A$  and collecting terms with  $\mathbf{U}$  gives

$$\frac{\partial E_{\text{elec}}^{\text{RHF}}}{\partial X_A} = 2 \sum_i^{\text{d.o.}} h_{ii}^{X_A} + \sum_{ij}^{\text{d.o.}} \{2(ii|jj)^{X_A} - (ij|ij)^{X_A}\} + 4 \sum_m^{\text{all}} \sum_i^{\text{d.o.}} U_{mi}^{X_A} F_{im},$$

(Yamaguchi eq. 4.16)

where the Fock matrix is defined as

$$F_{pq} = h_{pq} + \sum_k^{\text{d.o.}} \{2(pq|kk) - (pk|qk)\}$$

(Yamaguchi eq. 4.6)

$$= h_{pq} + 2J_{pq} - K_{pq},$$

and the Coulomb and exchange matrices  $\mathbf{J}$  and  $\mathbf{K}$  have also been introduced. Using the RHF variational conditions, the Fock matrix from a converged calculation is diagonal in the MO basis, corresponding to the MO energies

$$F_{pq} = \delta_{pq} \epsilon_{pq},$$

(Yamaguchi eq. 4.7)

so (Yamaguchi eq. 4.16) simplifies to

$$\frac{\partial E_{\text{elec}}^{\text{RHF}}}{\partial X_A} = 2 \sum_i^{\text{d.o.}} h_{ii}^{X_A} + \sum_{ij}^{\text{d.o.}} \{2(ii|jj)^{X_A} - (ij|ij)^{X_A}\} + 4 \sum_m^{\text{all}} \sum_i^{\text{d.o.}} U_{mi}^{X_A} \epsilon_{im},$$

(Yamaguchi eq. 4.17 modified)

which can be further simplified as

$$\frac{\partial E_{\text{elec}}^{\text{RHF}}}{\partial X_A} = 2 \sum_i^{\text{d.o.}} h_{ii}^{X_A} + \sum_{ij}^{\text{d.o.}} \{2(ii|jj)^{X_A} - (ij|ij)^{X_A}\} + 4 \sum_i^{\text{d.o.}} U_{ii}^{X_A} \epsilon_{ii}.$$

(Yamaguchi eq. 4.17)

Now one of the most important tricks in quantum chemistry is used. Given the orthonormality of the MOs,

$$S_{pq} = \delta_{pq},$$

(Yamaguchi eq. 3.44)

we must have (see footnote b)

$$\frac{\partial S_{pq}}{\partial X_A} \stackrel{!}{=} 0. \quad (\text{Yamaguchi eq. 3.45})$$

Expanding the left-hand side of (Yamaguchi eq. 3.45) similarly to (Yamaguchi eq. 3.81) gives

$$\frac{\partial S_{pq}}{\partial X_A} = \sum_{\mu\nu}^{\text{AO}} C_{\mu p} C_{\mu q} \frac{\partial S_{\mu\nu}}{\partial X_A} + \sum_m^{\text{all}} (U_{mp}^{X_A} S_{mq} + U_{mq}^{X_A} S_{pm}) \quad (\text{Yamaguchi eqs. 3.40 + 3.43})$$

$$= S_{pq}^{X_A} + \sum_m^{\text{all}} (U_{mp}^{X_A} S_{mq} + U_{mq}^{X_A} S_{pm}). \quad (\text{Yamaguchi eq. 3.43})$$

The sum over all MOs can be eliminated by reusing the orthonormality condition, so in the first term  $m \stackrel{!}{=} q$  and for the second term  $m \stackrel{!}{=} p$ , and the overlap matrix in the MO basis is unity for those terms, giving

$$\frac{\partial S_{pq}}{\partial X_A} = S_{pq}^{X_A} + U_{qp}^{X_A} + U_{pq}^{X_A} \stackrel{!}{=} 0. \quad (\text{Yamaguchi eq. 3.46})$$

Recognizing that we only need diagonal terms, this can be rewritten as

$$U_{pp}^{X_A} = -\frac{1}{2} S_{pp}^{X_A}, \quad (\text{Yamaguchi eq. 4.20})$$

which is then plugged back into the first derivative expression to give

$$\begin{aligned} \frac{\partial E_{\text{elec}}^{\text{RHF}}}{\partial X_A} &= 2 \sum_i^{\text{d.o.}} h_{ii}^{X_A} + \sum_{ij}^{\text{d.o.}} \{2(ii|jj)^{X_A} - (ij|ij)^{X_A}\} + 4 \sum_i^{\text{d.o.}} \left(-\frac{1}{2} S_{ii}^{X_A}\right) \epsilon_{ii} \\ &= 2 \sum_i^{\text{d.o.}} h_{ii}^{X_A} + \sum_{ij}^{\text{d.o.}} \{2(ii|jj)^{X_A} - (ij|ij)^{X_A}\} - 2 \sum_i^{\text{d.o.}} S_{ii}^{X_A} \epsilon_{ii}. \end{aligned}$$

(Yamaguchi eq. 4.21 [rederived])

Since is it almost always advantageous to avoid MO transformations and work in the AO basis, the last term can be rewritten

$$\begin{aligned}
\sum_i^{\text{d.o.}} S_{ii}^{X_A} \epsilon_{ii} &= \sum_i^{\text{d.o.}} \sum_{\mu\nu}^{\text{AO}} C_{\mu i} C_{\mu i} \frac{\partial S_{\mu\nu}}{\partial X_A} \epsilon_{ii} \\
&= \sum_i^{\text{d.o.}} \sum_{\mu\nu}^{\text{AO}} C_{\mu i} C_{\mu i} \epsilon_{ii} S_{\mu\nu}^{X_A} \\
&= \sum_{\mu\nu}^{\text{AO}} W_{\mu\nu} S_{\mu\nu}^{X_A}
\end{aligned} \tag{Yamaguchi eq. 4.24}$$

to use the energy-weighted density matrix  $\mathbf{W}$ , also sometimes called  $\mathbf{Q}$ :<sup>f</sup>

$$\frac{\partial E_{\text{elec}}^{\text{RHF}}}{\partial X_A} = 2 \sum_{\mu\nu}^{\text{AO}} P_{\mu\nu} h_{\mu\nu}^{X_A} + \sum_{\mu\nu\lambda\sigma}^{\text{AO}} P_{\mu\nu} P_{\lambda\sigma} \{2(\mu\nu|\lambda\sigma)^{X_A} - (\mu\lambda|\nu\sigma)^{X_A}\} - 2 \sum_{\mu\nu}^{\text{AO}} Q_{\mu\nu} S_{\mu\nu}^{X_A} + V_{NN}^{X_A} \tag{Szabo & Ostlund eq. C.12}$$

Again, the elimination of the  $\mathbf{U}$  matrix is one of the most important results in quantum chemistry, as it means the coupled-perturbed SCF equations described [previously](#) do not need to be solved for first derivatives of SCF wavefunctions. This is why density or MO coefficient derivatives are not present in the gradient expression.

Differentiation of ([Yamaguchi eq. 4.21](#)) once more with respect to another nuclear displacement  $Y_B$  is

$$\begin{aligned}
\frac{\partial^2 E_{\text{tot}}^{\text{RHF}}}{\partial X_A \partial Y_B} &= 2 \sum_i^{\text{d.o.}} h_{ii}^{X_A Y_B} + \sum_{ij}^{\text{d.o.}} \{2(ii|jj)^{X_A Y_B} - (ij|ij)^{X_A Y_B}\} \\
&\quad - 2 \sum_i^{\text{d.o.}} S_{ii}^{X_A Y_B} \epsilon_i - 2 \sum_i^{\text{d.o.}} \sum_p^{\text{all}} \{U_{ip}^{X_A} U_{ip}^{Y_B} + U_{ip}^{Y_B} U_{ip}^{X_A} - S_{ip}^{X_A} S_{ip}^{Y_B} - S_{ip}^{Y_B} S_{ip}^{X_A}\} \epsilon_i \\
&\quad + 4 \sum_p^{\text{all}} \sum_i^{\text{d.o.}} (U_{pi}^{Y_B} F_{pi}^{X_A} + U_{pi}^{X_A} F_{pi}^{Y_B}) + 4 \sum_p^{\text{all}} \sum_i^{\text{d.o.}} U_{pi}^{X_A} U_{pi}^{Y_B} \epsilon_p \\
&\quad + 4 \sum_p^{\text{all}} \sum_i^{\text{d.o.}} \sum_q^{\text{all}} \sum_j^{\text{d.o.}} U_{pi}^{X_A} U_{qj}^{Y_B} \{4(pi|qj) - (pq|ij) - (pj|i q)\} \\
&\quad - 3(X_A - X_B)(Y_A - Y_B) \frac{Z_A Z_B}{R_{AB}^5},
\end{aligned} \tag{Yamaguchi eqs. 4.54, 4.55, 3.127}$$

---

<sup>f</sup>The convention in Szabo & Ostlund is to absorb the RHF factor of 2 into the density matrix, which is not done here.

where the nucleus-nucleus repulsion energy derivative is included for completeness. This is the final expression for the molecular Hessian<sup>g</sup> derived in (1.13). From here and Table 37, we can see that evaluating the  $\mathbf{U}$  matrices (forming derivatives of the MO coefficients) is unavoidable.

However, the case of the molecular Hessian is a general one, because the AOs are perturbation-dependent: for a fixed electron position, the amplitude of a basis function will change if it moved, and they are typically atom-centered. In the case that the basis set is *not* perturbation dependent, as is most often in the case in electric field perturbations<sup>h</sup>, (Yamaguchi eqs. 4.54, 4.55, 3.127) reduces to

$$\frac{\partial^2 E_{\text{tot}}^{\text{RHF}}}{\partial F_\alpha \partial F_\beta} = -4 \sum_a^{\text{virt}} \sum_i^{\text{d.o.}} U_{ai}^{F_\beta} h_{ai}^{F_\alpha}, \quad (\text{Yamaguchi eq. 17.54})$$

where  $\frac{\partial^2 E_{\text{tot}}^{\text{RHF}}}{\partial F_\alpha \partial F_\beta}$  is the  $\alpha\beta$ -component of the static polarizability tensor (1.23), and the one-electron term  $h_{ai}^{F_\alpha}$  is the  $\alpha$ -component of the dipole operator in the occupied-virtual MO basis (the *property gradient*, see the  $\mathbf{P}/\mathbf{Q}$  matrix elements in section 5.3.1 and line 371). This specific term originates from the Fock matrix derivatives on line three, as the complete Hamiltonian now includes the perturbation (see (1.43)), which is the only term that survives the differentiation. In general, when the AO basis is perturbation independent, the energy derivative with respect to two arbitrary perturbations  $\lambda$  and  $\theta$  can be written as<sup>2</sup>

$$\frac{\partial^2 E}{\partial \lambda \partial \theta} = \sum_{\mu\nu} \left( P_{\mu\nu} \frac{\partial^2 h_{\mu\nu}}{\partial \lambda \partial \theta} + \frac{\partial P_{\mu\nu}}{\partial \theta} \frac{\partial h_{\mu\nu}}{\partial \lambda} \right), \quad (\text{Neese eq. 77})$$

where the first term is evaluated as an expectation value and the second term requires solution of the response equations. The derivative of the density matrix is related to the  $\mathbf{U}$  matrices via<sup>i</sup>

$$\frac{\partial \mathbf{P}}{\partial \theta} = \mathbf{C}^{(0)} (\mathbf{C}^{(\theta)})^T + \mathbf{C}^{(\theta)} (\mathbf{C}^{(0)})^T \quad (1.42)$$

and (1.41).

---

<sup>g</sup>Not mass-weighted.

<sup>h</sup>See Ref. [23] for a discussion of using electric field-dependent functions.

<sup>i</sup>(Neese eq. 77) is for real perturbations; in the case of imaginary perturbations, the sum changes to a difference.

The form of (Yamaguchi eq. 17.54) and (Neese eq. 77) as shown above should make it clear that for a second derivative of an HF wavefunction, only one set of  $\mathbf{U}$  matrices is needed. This leads to a potential computational savings. Considering the IR intensities  $\frac{d^2E}{dX_A dF_\alpha}$ , calculating  $\frac{d}{dX_A}(\mu_\alpha)$  would require  $3N$   $\mathbf{U}$  matrices (one for each nuclear displacement), but using (1.14) to calculate  $\frac{d}{dF_\alpha}\left(\frac{dE}{dX_A}\right)$  would only require 3  $\mathbf{U}$  matrices (one for each external field component).

### 1.3.6 Perturbation theory

In Rayleigh–Schrödinger perturbation theory<sup>j</sup>, the exact Hamiltonian  $\hat{H}$  of a system under an applied perturbation  $\hat{V}$  can be written as

$$\hat{H} = \hat{H}^{(0)} + \lambda \hat{V}, \quad (1.43)$$

where  $\hat{H}^{(0)}$  is the Hamiltonian in the absence of the perturbation and  $\lambda \in [0, 1]$  controls the strength of the perturbation. Note that it is not yet necessary to specify the exact form of  $\hat{V}$ . The main assumption in perturbation theory, worded in two ways, is that the unperturbed Hamiltonian is an acceptable approximation to the exact Hamiltonian, and the perturbation is small. This assumption allows for a power (Maclaurin) series expansion of the wavefunction  $|\Psi_i\rangle$  and its energy  $\mathcal{E}_i$  for a given state  $i$ , where increasing orders account for better approximations to the exact (perturbed) energy:<sup>k</sup>

$$|\Psi_i\rangle = |\psi_i^{(0)}\rangle + \lambda |\psi_i^{(1)}\rangle + \lambda^2 |\psi_i^{(2)}\rangle + \dots \quad (1.44)$$

$$\mathcal{E}_i = E_i^{(0)} + \lambda E_i^{(1)} + \lambda^2 E_i^{(2)} + \dots \quad (1.45)$$

Combining (1.43), (1.44), and (1.45) into the Schrödinger equation,

$$\hat{H} |\Psi_i\rangle = \mathcal{E}_i |\Psi_i\rangle, \quad (1.46)$$

---

<sup>j</sup>See Szabo & Ostlund<sup>5</sup> page 322; identical notation is followed throughout, except the summation index  $n$  is generally replaced with  $k$ .

<sup>k</sup>While one hopes the series is convergent, it is often not the case, so the series is often truncated at the second-order correction to the energy, which may still be useful. See Ref. [24] for a series convergence study. It is unclear if the same issue exists when  $\hat{V}$  corresponds to an operator other than  $\hat{V}_{ee}$ , such as for external fields.

gives

$$\begin{aligned} (\hat{H}^{(0)} + \lambda \hat{V}) & \left[ |\psi_i^{(0)}\rangle + \lambda |\psi_i^{(1)}\rangle + \lambda^2 |\psi_i^{(2)}\rangle + \dots \right] \\ & = \left[ E_i^{(0)} + \lambda E_i^{(1)} + \lambda^2 E_i^{(2)} + \dots \right] \left[ |\psi_i^{(0)}\rangle + \lambda |\psi_i^{(1)}\rangle + \lambda^2 |\psi_i^{(2)}\rangle + \dots \right], \end{aligned} \quad (1.47)$$

where the  $\{\lambda\}$  are now also useful for collecting terms of like orders. The zeroth-order terms give the Schrödinger equation for the unperturbed energy,

$$\hat{H}^{(0)} |\psi_i^{(0)}\rangle = E_i^{(0)} |\psi_i^{(0)}\rangle, \quad (1.48)$$

but equating all terms that are first order in  $\lambda$  on both sides gives

$$\hat{H}^{(0)} |\psi_i^{(1)}\rangle + \hat{V} |\psi_i^{(0)}\rangle = E_i^{(0)} |\psi_i^{(1)}\rangle + E_i^{(1)} |\psi_i^{(0)}\rangle, \quad (1.49)$$

where  $\lambda$  has been dropped for readability since it is present in front of each term. (1.49) can be simplified through integration using the bra  $\langle \psi_i^{(0)} |$ , which does not change the order from  $\lambda^1$ :

$$\begin{aligned} \langle \psi_i^{(0)} | \left( \hat{H}^{(0)} |\psi_i^{(1)}\rangle + \hat{V} |\psi_i^{(0)}\rangle \right) & = \langle \psi_i^{(0)} | \left( E_i^{(0)} |\psi_i^{(1)}\rangle + E_i^{(1)} |\psi_i^{(0)}\rangle \right) \\ \langle \psi_i^{(0)} | \hat{H}^{(0)} |\psi_i^{(1)}\rangle + \langle \psi_i^{(0)} | \hat{V} |\psi_i^{(0)}\rangle & = \langle \psi_i^{(0)} | E_i^{(0)} |\psi_i^{(1)}\rangle + \langle \psi_i^{(0)} | E_i^{(1)} |\psi_i^{(0)}\rangle \\ & = E_i^{(0)} \langle \psi_i^{(0)} | \psi_i^{(1)}\rangle + E_i^{(1)} \langle \psi_i^{(0)} | \psi_i^{(0)}\rangle. \end{aligned} \quad (1.50)$$

It is now important to know what orthonormality conditions exist between the set of all corrected states  $\{|\psi_i^{(n)}\rangle\}$ . For the unperturbed state, which is usually the Hartree–Fock ground state,

$$\langle \psi_i^{(0)} | \psi_i^{(0)}\rangle = 1, \quad (1.51)$$

and the choice of *intermediate normalization* is made,

$$\langle \psi_i^{(0)} | \Psi_i \rangle \stackrel{!}{=} 1, \quad (1.52)$$

which upon expanding the ket using (1.44) leads to

$$\langle \psi_i^{(0)} | \psi_i^{(n)}\rangle = 0 \quad (1.53)$$

for any correction state where  $n > 0$ . Returning to (1.50), this first allows for simplification of the right-hand side,

$$\begin{aligned} \langle \psi_i^{(0)} | \hat{H}^{(0)} | \psi_i^{(1)} \rangle + \langle \psi_i^{(0)} | \hat{V} | \psi_i^{(0)} \rangle &= E_i^{(0)} \langle \cancel{\psi_i^{(0)} | \psi_i^{(1)}} \rangle^{\color{red}0} + E_i^{(1)} \langle \cancel{\psi_i^{(0)} | \psi_i^{(0)}} \rangle^{\color{green}1} \\ &= E_i^{(1)}, \end{aligned} \quad (1.54)$$

and the first term on the left-hand side can be simplified using the hermiticity of the Hamiltonian followed by (1.53),

$$\begin{aligned} \langle \psi_i^{(0)} | \hat{H}^{(0)} | \psi_i^{(1)} \rangle &= \langle \psi_i^{(1)} | \hat{H}^{(0)} | \psi_i^{(0)} \rangle^* \\ &= \langle \psi_i^{(1)} | E_i^{(0)} | \psi_i^{(0)} \rangle^* \\ &= E_i^{(0)} \langle \psi_i^{(1)} | \psi_i^{(0)} \rangle^* \\ &= E_i^{(0)} \langle \cancel{\psi_i^{(0)} | \psi_i^{(1)}} \rangle^{\color{red}0} \\ &= 0. \end{aligned} \quad (1.55)$$

The final form of (1.50) is now

$$\langle \psi_i^{(0)} | \hat{V} | \psi_i^{(0)} \rangle = E_i^{(1)}, \quad (1.56)$$

revealing that the first-order correction to the energy is the expectation value of the perturbation operator over the zeroth-order wavefunction. For context, when using perturbation theory to approximate the correlation energy of system on top of the mean-field wavefunction,  $\hat{V} \equiv \hat{V}_{ee} = \frac{1}{|\vec{r}_1 - \vec{r}_2|} = \frac{1}{r_{12}}$ , the electron-electron repulsion operator. However, the perturbation operator may be any one- or two-electron operator, and (1.56) is exact as long as  $|\psi_i^{(0)}\rangle$  is a variationally-optimized state<sup>1</sup>. The key insight is that to calculate the first-order correction to the energy, only the zeroth-order wavefunction is required. This means that if  $\hat{V}$  is replaced with an operator related to a molecular property, it can be calculated as an expectation value without needing the perturbed wavefunction. This is the same result as

---

<sup>1</sup>Hartree–Fock and most density functional approximations that do not contain a perturbative correction (such as double hybrids) satisfy this criterion. Additionally, it is important for ALMO-EDA, where the polarized but CT-disallowed intermediate state  $|\psi_{\text{pol}}\rangle$  is variational, but the frozen density state  $|\psi_{\text{frz}}\rangle$  is *not*. All the work found in chapter 5 starts from  $|\psi_{\text{pol}}\rangle$ , so the use of a Lagrangian to account for orbital relaxation (leading to additional terms) is unnecessary.



in (Yamaguchi eq. 4.21), and resembles the result of the Hellmann–Feynman theorem when the AO basis is not dependent on the perturbation.

The generalization of (1.56) is that for  $n > 0$ , the  $n$ th order correction to the energy is given by

$$E_i^{(n)} = \langle \psi_i^{(0)} | \hat{V} | \psi_i^{(n-1)} \rangle, \quad (1.57)$$

so to find the second-order correction to the energy, the first-order correction to the wavefunction is required. The general rule is that given the  $n$ th order correction to the wavefunction, the  $2n + 1$ th order correction to the energy can be calculated. This is known as Wigner’s  $2n + 1$  rule (see section 6.9.2.4). From (1.57), the first form of the second-order correction is

$$E_i^{(2)} = \langle \psi_i^{(0)} | \hat{V} | \psi_i^{(1)} \rangle, \quad (1.58)$$

where the problem now becomes the calculation of the perturbed wavefunction  $|\psi_i^{(1)}\rangle$ . The strategy is to expand it as a linear combination of eigenfunctions of  $\hat{H}^{(0)}$  that are orthogonal to the unperturbed state (in line with (1.53)),

$$\langle k | \psi_i^{(0)} \rangle \stackrel{!}{=} 0, \quad (1.59)$$

and form a complete orthonormal set  $\{|\psi_k^{(0)}\rangle\} \equiv \{|k\rangle\}$ ,

$$|\psi_i^{(1)}\rangle = \sum_k c_k^{(1)} |k\rangle, \quad (1.60)$$

leading to

$$\langle k | \psi_i^{(1)} \rangle = c_k^{(1)}. \quad (1.61)$$

Using (1.53) and the fact that the set  $\{|k\rangle\}$  is complete, the resolution of the identity can be inserted:

$$|\psi_i^{(1)}\rangle = \sum_k |k\rangle \langle k | \psi_i^{(1)} \rangle \quad (1.62)$$

To calculate the first-order correction to the wavefunction, first rearrange (1.49) to collect all terms with the same ket,

$$(E_i^{(0)} - \hat{H}^{(0)}) |\psi_i^{(1)}\rangle = (\hat{V} - E_i^{(1)}) |\psi_i^{(0)}\rangle, \quad (1.63)$$

and multiply (1.63) on the left with  $\langle k|$  to give

$$E_i^{(0)} \langle k|\psi_i^{(1)}\rangle - \langle k|\hat{H}^{(0)}|\psi_i^{(1)}\rangle = \langle k|\hat{V}|\psi_i^{(0)}\rangle - E_i^{(1)} \langle k|\psi_i^{(0)}\rangle, \quad (1.64)$$

where the last term cancels due to  $\langle a|b\rangle = \delta_{ab}$ ,  $\{a, b\} \in k$ , and  $|k\rangle \neq |\psi_i^{(0)}\rangle$  from (1.59). To deal with the second term, use (1.60), canceling terms again due to orthonormality, and finally inserting (1.61):

$$\begin{aligned} \langle k|\hat{H}^{(0)}|\psi_i^{(1)}\rangle &= \langle k|\hat{H}^{(0)}\left(\sum_a c_a^{(1)}|a\rangle\right) \\ &= \langle k|\hat{H}^{(0)}\left(c_a^{(1)}|a\rangle + c_b^{(1)}|b\rangle + \dots + c_k^{(1)}|k\rangle + \dots\right) \\ &= \langle k|\left(E_a^{(0)}c_a^{(1)}|a\rangle + E_b^{(0)}c_b^{(1)}|b\rangle + \dots + E_k^{(0)}c_k^{(1)}|k\rangle + \dots\right) \\ &= E_a^{(0)}c_a^{(1)}\langle k|a\rangle + E_b^{(0)}c_b^{(1)}\langle k|b\rangle + \dots + E_k^{(0)}c_k^{(1)}\langle k|k\rangle + \dots \\ &= E_k^{(0)}c_k^{(1)} \\ &= E_k^{(0)}\langle k|\psi_i^{(1)}\rangle. \end{aligned} \quad (1.65)$$

Plugging (1.65) back into (1.64) gives

$$\left(E_i^{(0)} - E_k^{(0)}\right) \langle k|\psi_i^{(1)}\rangle = \langle k|\hat{V}|\psi_i^{(0)}\rangle \quad (1.66)$$

$$\langle k|\psi_i^{(1)}\rangle = \frac{\langle k|\hat{V}|\psi_i^{(0)}\rangle}{E_i^{(0)} - E_k^{(0)}}, \quad (1.67)$$

which upon inserting into (1.62), gives the final form of the first-order correction to the perturbed wavefunction:

$$|\psi_i^{(1)}\rangle = \sum_{k \neq i} |k\rangle \frac{\langle k|\hat{V}|\psi_i^{(0)}\rangle}{E_i^{(0)} - E_k^{(0)}}. \quad (1.68)$$

Inserting (1.68) into (1.58) then gives the second-order correction to the perturbed energy:

$$\begin{aligned} E_i^{(2)} &= \langle \psi_i^{(0)}|\hat{V}\left\{\sum_{k \neq i} |k\rangle \frac{\langle k|\hat{V}|\psi_i^{(0)}\rangle}{E_i^{(0)} - E_k^{(0)}}\right\} \\ &= \sum_{k \neq i} \frac{\langle \psi_i^{(0)}|\hat{V}|k\rangle \langle k|\hat{V}|\psi_i^{(0)}\rangle}{E_i^{(0)} - E_k^{(0)}}. \end{aligned} \quad (1.69)$$

In the case where  $\hat{V}$  is operator with multiple components, such as the dipole operator in (1.28), the energy becomes a rank-2 tensor, the polarizability matrix:

$$\alpha_{ab} = - \sum_{k \neq i} \frac{\langle \psi_i^{(0)} | \hat{\mu}_a | k \rangle \langle k | \hat{\mu}_b | \psi_i^{(0)} \rangle}{E_i^{(0)} - E_k^{(0)}}. \quad (1.70)$$

(1.70) is called a *sum-over-states* expansion due to the explicit summation over the set of excited states  $|\psi_k^{(0)}\rangle$ . Although the derivation assumes that all states are exact and the summation is infinite, the actual states are formed from the basis of excited Slater determinants using the converged SCF wavefunction as the reference state  $|\psi_i^{(0)}\rangle$ .

#### 1.4 DYNAMIC (FREQUENCY-DEPENDENT) RESPONSE PROPERTIES

Up to this point, only static perturbations have been considered, where the strength of an applied field does not oscillate or vary with time. This is perfectly satisfactory for many properties, such as those that are intrinsic to the system or do not depend on field strength. For example, it does not make sense to have a field strength directly associated with a change in nuclear positions; this is not the same thing as having nuclear positions change in the presence of an applied field. Similarly, from Table 1, rotational and nuclear magnetic moments do not have a strength associated with them. This holds generally for “internal” perturbations. However, many non-linear optical processes in particular take advantage of oscillating fields, so frequency- or time-dependent molecular response properties must be calculable.

In order to deal with frequency-dependent response, a slightly different path must be taken, in part because the total energy of the system is no longer stationary with time, so energy derivatives as encountered in Section 1.3.5 are not physically well-defined. There are numerous possible derivation frameworks which cannot be done justice here. Equivalent results arise from time-dependent perturbation theory, quasienergy derivatives,<sup>1,3,25,26</sup> and the polarization propagator,<sup>6</sup> to name a few. Ideally, there is a time-dependent extension of Section 1.3.2 that at some step allows for identification of coefficients in a series expansion with response functions. Similarly, in the static limit ( $\omega \rightarrow 0$ ), results from analytic

derivative theory should be recovered, making time-dependent response theory a more general framework for response properties. The following derivation is borrowed heavily from Ref. [26], with expansion using Ref. [1].

### 1.4.1 Quasienergy derivatives

The starting point is the now the time-dependent Schrödinger equation (TDSE),

$$\hat{H}(t) |\Psi(t)\rangle = i \frac{\partial}{\partial t} |\Psi(t)\rangle, \quad (1.71)$$

which, unlike (1.46) does not have the energy on the right-hand side. It is useful to factor out the time-dependent phase of the wavefunction,

$$|\Psi(t)\rangle = e^{-i\mathcal{F}(t)} |\tilde{\Psi}(t)\rangle, \quad (\text{Toulouse eq. 48})$$

where  $\Psi(t)$  is the original unfactored wavefunction,  $\mathcal{F}(t)$  is a phase factor, and  $\tilde{\Psi}(t)$  is the *phase-isolated wavefunction*.<sup>m</sup> Rearrange the TDSE (1.71),

$$\left[ \hat{H}(t) - i \frac{\partial}{\partial t} \right] |\Psi(t)\rangle = 0, \quad (\text{Toulouse eq. 47})$$

and use the phase-isolated wavefunction<sup>n</sup>

$$\left[ \hat{H}(t) - i \frac{\partial}{\partial t} \right] |\tilde{\Psi}(t)\rangle = \dot{\mathcal{F}}(t) |\tilde{\Psi}(t)\rangle, \quad (\text{Toulouse eq. 49})$$

where  $\dot{\mathcal{F}}(t)$  now resembles the energy from the time-independent Schrödinger equation (TISE). This “energy” is still not stationary with time, however it will reappear later as part of a stationary theory.

Before going further, the Hamiltonian must be specified. The partitioning introduced in (1.43) is modified:

$$\hat{H}(t) = \hat{H}^{(0)} + \hat{V}(t), \quad (1.72)$$

---

<sup>m</sup>This notation is consistent with Ref. [1]. In Ref. [26],  $\tilde{\Psi}(t)$  is the total wavefunction and  $\Psi(t)$  is the phase-isolated wavefunction. In Ref. [25],  $|\bar{0}\rangle$  is the total wavefunction,  $|\tilde{0}\rangle$  is the phase-isolated wavefunction, and  $|0\rangle$  is the stationary wavefunction, recovered from  $|\tilde{0}(t \rightarrow 0)\rangle$ .

<sup>n</sup> $\dot{A}(t) = dA(t)/dt$

so that time-dependence is introduced only to the perturbation and not the standard molecular Hamiltonian.<sup>o</sup> The prototypical perturbation usually represents an external electric field that is slowly turned on in order for the system to gradually respond and may oscillate with time at a characteristic frequency.<sup>p</sup> Representing the perturbation in terms of Fourier components,

$$\hat{V}(t) = x_1 \hat{B} e^{-i\omega t} + x_2 \hat{A} e^{+i\omega t}, \quad (\text{Toulouse eq. 46 [modified]})$$

shows that the perturbation, and therefore the Hamiltonian, is periodic in time ( $\hat{H}(t+T) = \hat{H}(t)$ ) with a period of  $T = 2\pi/\omega$ . Other perturbation shapes, such as step functions, can be represented using a more general form of (Toulouse eq. 46 [modified]) with more Fourier components:

$$\hat{V}(t) = \sum_{k=-N}^N e^{-i\omega_k t} \sum_X \epsilon_X(\omega_k) \hat{X}, \quad (\text{Gauss eq. 106})$$

where  $k$  is the Fourier component index,  $\omega_k$  are the frequencies, and  $\epsilon_X(\omega_k)$  and  $\hat{X}$  are the corresponding perturbation strengths and operators. This gives the required integration bounds for (Toulouse eq. 49), leading to

$$\begin{aligned} \mathcal{Q} &= \frac{1}{T} \int_0^T dt \dot{\mathcal{F}}(t) \\ &= \frac{1}{T} \int_0^T dt \frac{\langle \tilde{\Psi}(t) | \left[ \hat{H}(t) - i \frac{\partial}{\partial t} \right] | \tilde{\Psi}(t) \rangle}{\langle \tilde{\Psi}(t) | \tilde{\Psi}(t) \rangle}, \end{aligned} \quad (\text{Toulouse eq. 50})$$

where  $\mathcal{Q}$  is the *time-averaged quasienergy*. The second part of (Toulouse eq. 50) is the variational condition for  $\tilde{\Psi}(t)$ , making  $\mathcal{Q}$  stationary with respect to fluctuations in  $\tilde{\Psi}(t)$ . In the time-independent case, the quasienergy reduces to the (time-independent) energy, and the phase-isolated wavefunction reduces to the (time-independent) wavefunction.

The time-averaged quasienergy can now be expanded with respect to the perturbation parameters,

$$\mathcal{Q}(x_1, x_2) = \mathcal{Q}^{(0)} + \mathcal{Q}^{(10)} x_1 + \mathcal{Q}^{(01)} x_2 + \frac{1}{2} \mathcal{Q}^{(20)} x_1^2 + \mathcal{Q}^{(11)} x_1 x_2 + \frac{1}{2} \mathcal{Q}^{(02)} x_2^2 + \dots, \quad (\text{Toulouse eq. 51})$$

---

<sup>o</sup>Sometimes  $\hat{V}(t)$  is notated as  $\hat{H}'(t)$  or  $\hat{H}^{(1)}(t)$ ; all are equivalent as long as the perturbation and time partitioning in the full Hamiltonian are the same.

<sup>p</sup>In the case that the response of the system is slower than the speed of which the perturbation is turned on, a different (undesirable) solution may be found if the perturbation is strong enough.

analogous to the series expansion in Section 1.3.2. However, the expansion coefficients  $\partial^n \mathcal{Q}/\partial x^n$  are still not in a position to be directly calculable for an approximate theory due to the presence of the integral in (Toulouse eq. 50), unlike in the static case, so the time-dependent theory must be further derived for exact states.

Using the time-dependent form of (1.56),

$$\begin{aligned} \mathcal{Q}^{(10)} &= \frac{1}{T} \int_0^T dt \langle \tilde{\Psi}_0 | \hat{B} e^{-i\omega t} | \tilde{\Psi}_0 \rangle \\ &= \langle \tilde{\Psi}_0 | \hat{B} | \tilde{\Psi}_0 \rangle \delta[\omega]. \end{aligned} \quad (\text{Toulouse eq. 53})$$

The rule used in (Toulouse eq. 53) is

$$\frac{1}{T} \int_0^T dt e^{-i\omega t} = \delta[\omega] = \begin{cases} \delta_{\omega,0} & |T| < \infty \\ \delta(\omega) & T \rightarrow \infty \end{cases}, \quad (1.73)$$

where we define  $\delta[\omega]$  as the Kronecker delta for finite  $T$  and the Dirac delta for  $T \rightarrow \infty$ . More explicitly,  $\delta[\omega = 0] = (1/T) \int_0^T dt e^0 = 1$ .  $\delta[\omega \neq 0]$  would lead a non-zero value for exponential inside (1.73), which when expanded as sines and cosines, gives rise to terms that are not physically admissible as Fourier components. The conclusion is that only static first-order properties are non-zero, where the physical intuition is that absorption or emission of a photon with  $\omega \neq 0$  would violate energy conservation. Using (1.58) and (1.52),

$$\mathcal{Q}^{(20)} = \frac{1}{T} \int_0^T dt 2 \langle \tilde{\Psi}_0 | \hat{B} e^{-i\omega t} | \tilde{\Psi}^{(10)}(t) \rangle, \quad (\text{Toulouse eq. 54})$$

where the first-order wavefunction correction is (similar to (1.60))

$$|\tilde{\Psi}^{(10)}(t)\rangle = \sum_{n \neq 0} c_n^{(10)}(t) e^{-i\omega_{n0} t} |\tilde{\Psi}_n\rangle, \quad (\text{Toulouse eq. 55})$$

where  $\omega_{n0} = E_n - E_0$  are the unperturbed system excitation energies. The first-order wavefunction expansion coefficients are found similarly to (1.61),

$$c_n^{(10)}(t) = -i \int_{-\infty}^t dt' \langle \tilde{\Psi}_n | \hat{B} e^{-i\omega t'} e^{\gamma t'} | \tilde{\Psi}_0 \rangle e^{i\omega_{n0} t'}, \quad (\text{Toulouse eq. 56})$$

where a factor of  $e^{\gamma t}$  with  $\gamma \rightarrow 0^+$  adiabatically switches on the perturbation from  $t' \rightarrow -\infty$  and imposes the initial condition  $c_n^{(10)}(t \rightarrow -\infty) = 0$ . Performing the integration in (Toulouse eq. 56) gives

$$c_n^{(10)}(t) = \frac{\langle \tilde{\Psi}_n | \hat{B} e^{-i\omega t} | \tilde{\Psi}_0 \rangle e^{i\omega_{n0} t}}{\omega - \omega_{n0} + i\gamma}, \quad (\text{Toulouse eq. 57})$$

where the factor  $e^{\gamma t}$  is dropped from the numerator, but now supplies a damping factor for when the field oscillation comes close to a resonance (excitation energy). Now (Toulouse eq. 54), (Toulouse eq. 55), and (Toulouse eq. 57) are combined to give

$$\begin{aligned} \mathcal{Q}^{(20)} &= \frac{1}{T} \int_0^T dt 2 \sum_{n \neq 0} \frac{\langle \tilde{\Psi}_0 | \hat{B} e^{-i\omega t} | \tilde{\Psi}_n \rangle \langle \tilde{\Psi}_n | \hat{B} e^{-i\omega t} | \tilde{\Psi}_0 \rangle}{\omega - \omega_{n0} + i\gamma} \\ &= 2 \sum_{n \neq 0} \frac{\langle \tilde{\Psi}_0 | \hat{B} | \tilde{\Psi}_n \rangle \langle \tilde{\Psi}_n | \hat{B} | \tilde{\Psi}_0 \rangle}{\omega - \omega_{n0} + i\gamma} \delta[2\omega], \end{aligned} \quad (\text{Toulouse eq. 58})$$

which is only finite for  $\omega = 0$  similarly to (Toulouse eq. 53). Performing the same steps for the mixed term,

$$\begin{aligned} \mathcal{Q}^{(11)} &= \frac{1}{T} \int_0^T dt \left[ \langle \tilde{\Psi}_0 | \hat{A} e^{+i\omega t} | \tilde{\Psi}^{(10)}(t) \rangle + \langle \tilde{\Psi}_0 | \hat{B} e^{-i\omega t} | \tilde{\Psi}^{(01)}(t) \rangle \right] \\ &= \frac{1}{T} \int_0^T dt \left[ \frac{\langle \tilde{\Psi}_0 | \hat{A} e^{+i\omega t} | \tilde{\Psi}_n \rangle \langle \tilde{\Psi}_n | \hat{B} e^{-i\omega t} | \tilde{\Psi}_0 \rangle}{\omega - \omega_{n0} + i\gamma} + \frac{\langle \tilde{\Psi}_0 | \hat{B} e^{-i\omega t} | \tilde{\Psi}_n \rangle \langle \tilde{\Psi}_n | \hat{A} e^{+i\omega t} | \tilde{\Psi}_0 \rangle}{-\omega - \omega_{n0} + i\gamma} \right] \\ &= \sum_{n \neq 0} \frac{\langle \tilde{\Psi}_0 | \hat{A} | \tilde{\Psi}_n \rangle \langle \tilde{\Psi}_n | \hat{B} | \tilde{\Psi}_0 \rangle}{\omega - \omega_{n0} + i\gamma} - \sum_{n \neq 0} \frac{\langle \tilde{\Psi}_0 | \hat{B} | \tilde{\Psi}_n \rangle \langle \tilde{\Psi}_n | \hat{A} | \tilde{\Psi}_0 \rangle}{\omega + \omega_{n0} - i\gamma}, \end{aligned} \quad (\text{Toulouse eq. 59})$$

which is in general nonzero for all  $\omega$ . Physically, it corresponds to the absorption of a virtual photon due to  $\hat{B}$  that probes the excited states of the system, followed by emission of a photon due to  $\hat{A}$ , which is allowed due to net energy conservation.  $\mathcal{Q}^{(11)}$  is the *linear response function* of operators  $\hat{A}$  and  $\hat{B}$  in the spectral representation and it is often denoted as  $\mathcal{Q}^{(11)} = \langle\langle \hat{A}; \hat{B} \rangle\rangle_\omega$  (see Table 2). Physically, it can be viewed as the change in the expectation value of the operator  $\hat{A}$  due to the perturbation  $\hat{B} e^{-i\omega t}$ . Due to the pole structure introduced by the denominator (with de-excitation energies arising from the second term), (Toulouse eq. 59) contains information about all the excitation energies of the system. Higher-order derivatives of the quasienergy lead to time-dependent non-linear response functions: third derivatives correspond to quadratic response, and fourth derivatives correspond to cubic response, and so on.

## 2.0 CARBON CAPTURE FROM CARBON DIOXIDE’S POINT-OF-VIEW

The text in this chapter has been adapted from Brinzer, T.; Berquist, E. J.; Ren, Z.; Dutta, S.; Johnson, C. A.; Krisher, C. S.; Lambrecht, D. S.; Garrett-Roe, S. Ultrafast Vibrational Spectroscopy (2D-IR) of CO<sub>2</sub> in Ionic Liquids: Carbon Capture from Carbon Dioxide’s Point of View. *J. Chem. Phys.* **2015**, *142*, 212425, DOI: [10.1063/1.4917467](https://doi.org/10.1063/1.4917467), and its erratum Brinzer, T.; Berquist, E. J.; Dutta, S.; Johnson, C. A.; Krisher, C. S.; Lambrecht, D. S.; Garrett-Roe, S.; Ren, Z. Erratum: "Ultrafast vibrational spectroscopy (2D-IR) of CO<sub>2</sub> in ionic liquids: Carbon capture from carbon dioxide’s point of view" [J. Chem. Phys. 142, 212425 (2015)]. *J. Chem. Phys.* **2017**, *147*, 049901, DOI: [10.1063/1.4995447](https://doi.org/10.1063/1.4995447). The author’s contribution to the work included performing all *ab initio* calculations and writing the corresponding sections 2.4.2, 2.4.3, and 2.3.5. These calculations consisted of constructing molecular models, performing geometry optimizations followed by harmonic frequency calculations using DFT, and designing the application of ALMO-EDA to vibrational frequency analysis.

## 2.1 SUMMARY

The CO<sub>2</sub>  $\nu_3$  asymmetric stretching mode is established as a vibrational chromophore for ultrafast two-dimensional infrared (2D-IR) spectroscopic studies of local structure and dynamics in ionic liquids, which are of interest for carbon capture applications. CO<sub>2</sub> is dissolved in a series of 1-butyl-3-methylimidazolium-based ionic liquids ([Im<sub>4,1</sub>][X]), where [X]<sup>−</sup> is the anion from the series hexafluorophosphate (PF<sub>6</sub><sup>−</sup>), tetrafluoroborate (BF<sub>4</sub><sup>−</sup>), bis-(trifluoromethylsulfonyl)imide (Tf<sub>2</sub>N<sup>−</sup>), triflate (TfO<sup>−</sup>), trifluoroacetate (TFA<sup>−</sup>), dicyan-



amide ( $\text{DCA}^-$ ), and thiocyanate ( $\text{SCN}^-$ ). In the ionic liquids studied, the  $\nu_3$  center frequency is sensitive to the local solvation environment and reports on the timescales for local structural relaxation. Density functional theory calculations predict charge transfer from the anion to the  $\text{CO}_2$  and from  $\text{CO}_2$  to the cation. The charge transfer drives geometrical distortion of  $\text{CO}_2$ , which in turn changes the  $\nu_3$  frequency. The observed structural relaxation timescales vary by up to an order of magnitude between ionic liquids. Shoulders in the 2D-IR spectra arise from anharmonic coupling of the  $\nu_2$  and  $\nu_3$  normal modes of  $\text{CO}_2$ . Thermal fluctuations in the  $\nu_2$  population stochastically modulate the  $\nu_3$  frequency and generate dynamic cross-peaks. These timescales are attributed to the breakup of ion cages that create a well-defined local environment for  $\text{CO}_2$ . The results suggest that the picosecond dynamics of  $\text{CO}_2$  are gated by local diffusion of anions and cations.

## 2.2 INTRODUCTION

There is a pressing need to develop next-generation materials to capture  $\text{CO}_2$  from fossil-fuel burning power plants. Commercial carbon capture technologies are inefficient and greatly increase the cost of energy.<sup>29</sup> Novel materials, including metal-organic frameworks,<sup>30,31</sup> polymers,<sup>32,33</sup> and ionic liquids,<sup>34-36</sup> have been proposed as transformational technologies. In each case, however, a lack of tools to interrogate at a molecular scale how  $\text{CO}_2$  interacts with the sorbents, what local structures it forms, and for how long those structures persist has limited the ability to optimize these materials for this important task.

Our strategy to investigate the local environment of  $\text{CO}_2$  is to use a vibration of the  $\text{CO}_2$  itself. We hypothesized that the antisymmetric stretching vibration of  $\text{CO}_2$ , the  $\nu_3$  mode, could be a sensitive probe of the carbon capture process and that ultrafast two-dimensional vibrational spectroscopy (2D-IR) could give new insight into the local structure and dynamics around the  $\text{CO}_2$ . The central goal of this report is to establish the  $\nu_3$  mode of  $\text{CO}_2$  as a viable platform for ultrafast multidimensional spectroscopy of carbon capture in ionic liquids.

Ionic liquids, sometimes called room temperature molten salts, excite particular interest for carbon capture due to their chemical tunability. Ionic liquids are organic salts that are

molten at or below 100 °C. Each formula unit consists of an anion-cation pair without a surrounding solvent. The physical and chemical properties of ionic liquids can be manipulated by changing the specific anion and cation pair, which provides tremendous flexibility. The chemical space spanned by possible ionic liquids is estimated at  $10^{18}$  anion-cation combinations, including binary and ternary mixtures,<sup>37</sup> only a tiny fraction of which has been explored. Even without optimization, many ionic liquids have CO<sub>2</sub> solubility and selectivity comparable to molecular solvents,<sup>38,39</sup> and chemical modification can improve these properties further.<sup>40,41</sup> Ionic liquids exhibit negligible vapor pressure and are generally thermally and hydrolytically stable at the operating temperature ( $\sim 50$  °C) of post-combustion carbon capture.

The promise of chemical tunability of ionic liquid properties has spurred efforts to design improved ionic liquids, both synthetically<sup>34,40–45</sup> and *in silico*.<sup>46,47</sup> Additionally, other strategies, such as creating mixtures of multiple ionic liquids,<sup>48,49</sup> or mixtures of ionic liquids with molecular solvents<sup>50,51</sup> show promise. Despite the progress that has been made, there remains a lack of fundamental understanding of the local solute-solvent interactions between CO<sub>2</sub> and the ionic liquid sorbents.

Molecular modelling provides valuable insights but faces challenges due to the time- and length-scale mismatches between the simulations and bulk thermodynamic experiments. For example, electronic structure theory can predict CO<sub>2</sub> binding motifs and energies,<sup>52–55</sup> molecular dynamics simulations of CO<sub>2</sub> in ionic liquids can provide atomistic detail of local structure and dynamics in the condensed phase,<sup>56–61</sup> and Monte Carlo simulations can calculate important thermodynamic properties.<sup>57,62–66</sup> However, it remains a challenge to directly compare these results to macroscopic experimental observables such as viscosity or CO<sub>2</sub> solubility.

Ultrafast spectroscopy naturally provides observables that are compatible with the time- and length-scales of molecular modelling. Ultrafast two-dimensional infrared spectroscopy (2D-IR), a coherent third-order spectroscopy, can investigate femtosecond to picosecond molecular dynamics in the condensed phase at equilibrium.<sup>67–70</sup>

A 2D-IR experiment, in essence, measures the distribution of vibrational frequencies in an ensemble of molecules at two points in time, separated by a controllable delay. When the

delay is short, the local environment does not have time to reorganize, causing correlation of initial and final frequencies, and stretching the 2D-IR peak along the frequency diagonal. As the delay time is increased, the system loses memory of its initial configuration and the peaks become rounder. Thus, 2D-IR spectra encode the two-point frequency fluctuation correlation function through the change in their shape with the population time,  $t_2$ , of the experiment

$$c_2(t_2) = \langle \delta\omega(t_2)\delta\omega(0) \rangle \quad (2.1)$$

where  $\delta\omega(t)$  is the offset of the instantaneous vibrational frequency  $\omega(t)$  from the average  $\langle\omega\rangle$

$$\delta\omega(t) = \omega(t) - \langle\omega\rangle \quad (2.2)$$

This time-correlation function is the essential information content of any specific peak in 2D-IR spectroscopy, which can be compared with the available time-correlation functions from molecular simulations.

Cross-peaks in 2D-IR also contain valuable information. Vibrational coupling of modes, population transfer, and chemical exchange can all generate cross-peaks in a 2D-IR spectrum. These mechanisms can be distinguished by their spectral kinetics. Vibrational coupling, or mixing of the local modes, causes cross-peaks that are seen even at the earliest population times, because they result from direct vibrational transitions, for example the coupling of the symmetric and antisymmetric stretching modes of water.<sup>71</sup> Coherence transfer between bright modes can additionally create beat frequencies in both their diagonal and cross-peaks.<sup>71</sup> Population transfer arises from the exchange of excited state population during  $t_2$ , for example, that of the Amide I modes of small peptides.<sup>72</sup> Chemical exchange results from the change in frequency of a single molecule as it moves between different local environments, for example, the exchange of free and complexed phenol-benzene.<sup>70</sup> Both population transfer and chemical exchange give rise to dynamic cross-peaks, whose relative intensity shows an early minimum, increasing with population time.

Linear spectroscopies can provide insight into structure and dynamics in ionic liquids. For example, Raman spectroscopy has been used to study CO<sub>2</sub> in [Im<sub>4,1</sub>][TFA],<sup>73</sup> or the effects of ionic liquid water content on CO<sub>2</sub> solubility.<sup>50</sup> Similarly, NMR spectroscopy has been used to probe the solvation of CO<sub>2</sub> in imidazolium ionic liquids at high pressures.<sup>74,75</sup>

A number of studies have also examined structure and dynamics in ionic liquids on an ultrafast timescale using various chromophores, including IR pump-probe measurements of small anions in ionic liquids,<sup>76</sup> 2D-IR studies of water<sup>71</sup> and thiocyanate<sup>77</sup> in imidazolium ionic liquids, electronic spectroscopy of solvated laser dyes,<sup>78,79</sup> and optical Kerr effect spectroscopies.<sup>80-83</sup> Each of these spectroscopies is sensitive to the spectral density of low frequency vibrations (i.e., the Fourier transform of  $c_2(t)$ ), but they differ in the coupling of their respective chromophores to the low frequency modes.

The difficulty for most ultrast spectroscopies lies in connecting the average solvent dynamics to the dynamics specifically around the CO<sub>2</sub>. Time-dependent Stokes shift and fluorescence upconversion experiments<sup>79</sup> can measure dynamics over a broad range of timescales; however, it is an open question how relevant the dynamics of a comparatively large laser dye are to those of a small molecule such as CO<sub>2</sub>. Optical Kerr effect spectroscopies<sup>80,84</sup> are sensitive to the many body Raman polarizability tensor, which is difficult to connect to specific motions around a solute in most cases. In our approach, our spectroscopic probe is guaranteed to access the solvation environment of CO<sub>2</sub> (because it *is* CO<sub>2</sub>) and is guaranteed to be local to the CO<sub>2</sub> (because it is an isolated vibration of the CO<sub>2</sub>). On the other hand, the observed solvation dynamics will arise from only those local motions which most strongly couple into the  $\nu_3$  mode; there is no guarantee that the accessible solvent motions are important. This report explores what information this otherwise optimal probe provides.

For our present purposes, a good spectroscopic probe should have the following properties: a high transition dipole moment ( $\epsilon \sim 800$  to  $1000 \text{ M}^{-1} \text{ cm}^{-1}$ ); a fundamental frequency in a spectral region that is free of fundamentals or strong combination bands (for most ionic liquids from  $1600$  to  $2600 \text{ cm}^{-1}$ ); sensitivity to the local environment; a lifetime,  $T_1$ , that is long enough to measure dynamics on a tens of picoseconds timescale; and, finally, a lineshape at least partially inhomogeneously broadened, so that motional narrowing (homogeneous broadening) does not average away the embedded information.

Given these requirements, the  $\nu_3$  band of CO<sub>2</sub> is a natural choice. Nonlinear spectroscopy on the  $\nu_3$  band ( $\epsilon \sim 1000 \text{ M}^{-1} \text{ cm}^{-1}$ ) has been demonstrated in water,<sup>85,86</sup> where the  $T_1$  time is  $\sim 10$  ps. In addition, the mode absorbs in a free spectroscopic window of most ionic liquids ( $\sim 2350 \text{ cm}^{-1}$ ).

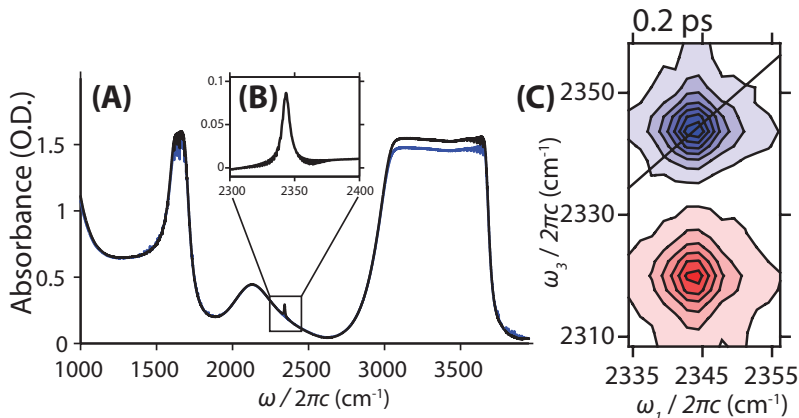


Figure 2: (A) FTIR spectrum of H<sub>2</sub>O with (blue line) and without (black line) CO<sub>2</sub> dissolved in it. CO<sub>2</sub>  $\nu_3$  lies in the overtones and combinations bands from H<sub>2</sub>O librational modes. (B) Inset of  $\nu_3$  with water background subtracted shows the Lorentzian character of the peak. (C) Purely absorptive 2D-IR spectrum of  $\nu_3$  in H<sub>2</sub>O with  $t_2 = 0.2$  ps shows that, in water, the line is nearly completely in the limit of homogeneous dynamics.

The remaining question is how sensitively CO<sub>2</sub> reports on structural relaxation of its local environment in ionic liquids. CO<sub>2</sub> in water is an important point of reference because, in water, CO<sub>2</sub> is insensitive to its environment. The CO<sub>2</sub>  $\nu_3$  frequency shift from the gas phase to the condensed phase is very small, which indicates low sensitivity to the condensed phase local environment. In addition, the CO<sub>2</sub>  $\nu_3$  line is narrow and Lorentzian (Figure 2). Even at early population times, the 2D spectra show almost no diagonal character (Figure 2C), indicating that peak shape is determined by dephasing time,  $T_2$ , and that a deeper analysis of the frequency fluctuations (Equation 2.1) is not possible. This finding is consistent with previous three pulse photon echo peak shift (3PEPS) experiments.<sup>85</sup> In ionic liquids, the timescales of solvent motion are expected to be slower than in water, and the coupling to the environment may well be larger. Both of these effects would increase the ability to observe dynamics around CO<sub>2</sub>.

The other vibrational modes of CO<sub>2</sub>, the symmetric stretch,  $\nu_1$ , and the doubly degenerate bend,  $\nu_2$  and  $\bar{\nu}_2$ , present distinct challenges for vibrational spectroscopy. The  $\nu_2$  modes

are sensitive to the chemical environment of the CO<sub>2</sub>,<sup>87,88</sup> but are located in the crowded fingerprint region. The Raman active  $\nu_1$  is readily measured, but the lineshape is dominated by the Fermi resonance with the overtone of the  $\nu_2$ ;<sup>73</sup> furthermore, it is a dark mode for IR.

Here, we demonstrate that the CO<sub>2</sub>  $\nu_3$  mode reports on its local solvent environment by showing the sensitivity of the  $\nu_3$  frequency and dynamics to variation in solvent anion in a series of imidazolium ionic liquids; furthermore, we show that  $\nu_3$  reports a broad range of solvation timescales in these ionic liquids. We employ electronic structure calculations to investigate the mapping of vibrational frequencies onto structures of CO<sub>2</sub>-anion-cation clusters. We show that, despite apparent complexity, the CO<sub>2</sub> linear and 2D-IR lineshapes can be interpreted using simple and accurate physical models. Finally, we establish correlations between the measured dynamics of CO<sub>2</sub> and the macroscopic properties of its ionic liquid solvent.

The paper is organized as follows. Initially, we present the analysis and interpretation of the linear CO<sub>2</sub> spectrum (Section 2.4.1), including a discussion of its temperature dependence. Next, we describe our results from electronic structure calculations on CO<sub>2</sub>-anion-cation clusters and its relationship to our experimental data (Section 2.4.2), followed by a simple model of CO<sub>2</sub> that is able to reproduce the observed trends (Section 2.4.3). We then present an overview of the 2D-IR spectra (Section 2.4.4), including a kinetic analysis of the observed shoulders and cross-peaks (Section 2.4.5), assignments of the spectral features (Section 2.4.6), and modelling of the main  $\nu_3$  peak (Section 2.4.7) and shoulders and cross-peaks (Section 2.4.8). Finally, we present a discussion of the physical interpretation of our results (Section 2.5), conclusions (Section 2.6) and methods (Section 2.3).

## 2.3 METHODS

### 2.3.1 Materials

Ionic liquids were obtained from Ionic Liquids Technologies, Inc (IoLiTec), and used without further purification (except for [Im<sub>4,1</sub>][BF<sub>4</sub>], which was purified prior to experiments using

the procedure described by Giernoth and Bankmann.<sup>89)</sup> Ionic liquid samples were stored in ambient conditions; however, prior to experiments, samples dried under vacuum at 50 mTorr.

### 2.3.2 FTIR

FTIR spectra were measured using a N<sub>2(g)</sub>-purged Nicolet 6700 FTIR instrument (ThermoFisher Scientific). Ionic liquid samples were loaded with CO<sub>2</sub> (99.8% purity, Matheson TRIGAS) in an airtight custom glass vial, with a septum that allows access to the CO<sub>2</sub>-loaded ionic liquid. An aliquot of the liquid sample was sandwiched between two 2 mm-thick CaF<sub>2</sub> optical windows (Crystran Ltd, UK) separated by a 25  $\mu$ m Teflon spacer, which were held in a brass sample cell. The cell was assembled in a glove bag to limit adsorption of atmospheric water. Spectra were obtained for both the neat ionic liquid and for the ionic liquid loaded with CO<sub>2</sub>.

For temperature-dependent measurements, the sample was temperature-controlled by using a cooling/heating recirculating chiller (Fisher Isotemp) to control the temperature of the sample cell holder. The sample temperature was monitored by measuring the temperature at the optical window using a thermocouple (National Instruments USB-TC01 J-type).

### 2.3.3 2D-IR

**2.3.3.1 Generation of femtosecond mid-IR pulses** The experiments utilize a commercial Ti:Sapphire chirped pulse amplifier laser system ( $\lambda = 805$  nm, 5 kHz repetition rate, 120 fs pulse duration) (Coherent Vitesse / Coherent Legend Elite).

A home-built optical parametric amplifier (OPA) generates the mid-IR pulses ( $\lambda = 12$  to  $2 \mu$ m), corresponding to around  $830$  to  $5000 \text{ cm}^{-1}$ . The OPA design leads to noise suppression in the resulting mid-IR pulses.<sup>90</sup> The spectral bandwidth after the OPA is about  $200 \text{ cm}^{-1}$ . For these experiments, we tune the OPA wavelength to  $4.3 \mu$ m. Mid-IR pulse energy entering the 2D spectrometer is approximately  $2.2 \mu$ J per pulse.

**2.3.3.2 2D Spectrometer** The 2D-IR spectrometer uses a pump-probe geometry,<sup>91</sup> in which the first two mid-IR pulses travel collinearly to the sample. A Mach-Zehnder in-

terferometer controls the coherence time  $t_1$  between these pulses. A delay stage after the interferometer controls the population time  $t_2$  between the second and third pulses. The signal, which contains both rephasing and non-rephasing components, is emitted in the direction of the probe pulse ( $\vec{k}_3$ ), which also serves as a local oscillator to heterodyne the signal.

A 150 line/mm grating in a single monochromator disperses the signal in  $\omega_3$  onto a liquid N<sub>2</sub>-cooled  $2 \times 32$  channel mercury cadmium telluride (MCT) detector. The signal in  $\omega_1$  is indirectly acquired by scanning along  $t_1$  using the Mach-Zehnder interferometer, and then applying a numerical Fourier transform to the resulting  $t_1$ -dependent signal at each data point in  $\omega_3$  (which also removes the transient absorption signal). The delay changes as the interferometer scans along  $t_1$  are acquired by comparing the interference pattern generated by a He:Ne beam which travels a parallel path through the interferometer.

A series of spectra in  $t_2$  are then acquired by varying the population time between the second and third laser pulses, and then obtaining a spectrum in  $\omega_1$  and  $\omega_3$  at each population time.

### 2.3.4 Global Fitting and Bootstrapping

Global fitting of spectra utilizes a third-order response function formalism in the semi-impulsive limit, including the Condon approximation and also the approximation of the cumulant expansion truncated after second order. A constrained nonlinear optimization algorithm (`fmincon` MATLAB) is used to minimize the magnitude of the sum of squared error between each data point in the set of spectra and a corresponding data point in a calculated spectrum. A bootstrapping algorithm<sup>92</sup> establishes the error of the global fitting result. The global fitting algorithm was run using synthetic data sets composed of a random selection of the original data points taken with replacement. The distribution of the fitting parameters after 100 iterations provided the error estimate.



### 2.3.5 Computational Details

All calculations were performed with a development version of the Q-Chem program package,<sup>93</sup> employing the B3LYP density functional,<sup>94,95</sup> the 6-31G(d,p) basis set,<sup>96,97</sup> and a (100,302) grid for the numerical quadrature. All SCF calculations were tightly converged to below  $10^{-9}$  a.u. for the DIIS error.<sup>98</sup> Gas-phase geometry optimizations of free CO<sub>2</sub> and the CO<sub>2</sub>-ionic liquid clusters were converged to changes of  $1 \times 10^{-8}$  a.u. in the energy,  $1 \times 10^{-6}$  a.u. in the nuclear displacement, and  $1 \times 10^{-6}$  a.u. in the gradient. Optimized structures were confirmed as minima via harmonic frequency calculations using analytic Hessians. Frequencies were scaled by a factor of 0.9627 (Table 6 Merrick et al.).<sup>99</sup>

For the calculations including the electrostatic and polarization effects of the ionic liquid with charge transfer disabled (“ $+\Delta\omega_{\text{FRZ}} + \Delta\omega_{\text{POL}}$ ” in Figure 5), absolutely localized molecular orbitals (ALMOs)<sup>100</sup> were employed, with the ionic liquid constituents as one combined fragment and the CO<sub>2</sub> as another fragment. Solution of the ALMO equations is requested by setting `frgm_method = gia`. Charge transfer between fragments is disabled by setting `frgm_lpcorr = 0`. For the calculation of complementary occupied-virtual pairs (COVPs), an energy decomposition analysis (ALMO-EDA) calculation is performed where charge transfer is allowed (`frgm_lpcorr = rs_exact_scf`).

## 2.4 RESULTS AND DISCUSSION

### 2.4.1 Linear IR Spectroscopy Results

The linear spectra of CO<sub>2</sub> establish that the frequency of the  $\nu_3$  mode is sensitive to the anion, and set the stage for discussing the spectroscopic features present in the 2D-IR spectra.

The  $\nu_3$  vibration absorbs strongly around  $2340 \text{ cm}^{-1}$  in a spectral region with no strong solvent absorbances (Figure 3A). The lineshape of the  $\nu_3$  band (Figure 3B) appears mostly Lorentzian, with a low frequency shoulder.

Changing the anion causes both the full-width at half-maximum (FWHM) and center frequency of  $\nu_3$  to change (Figure 3C). We varied the anion, rather than the cation, because the

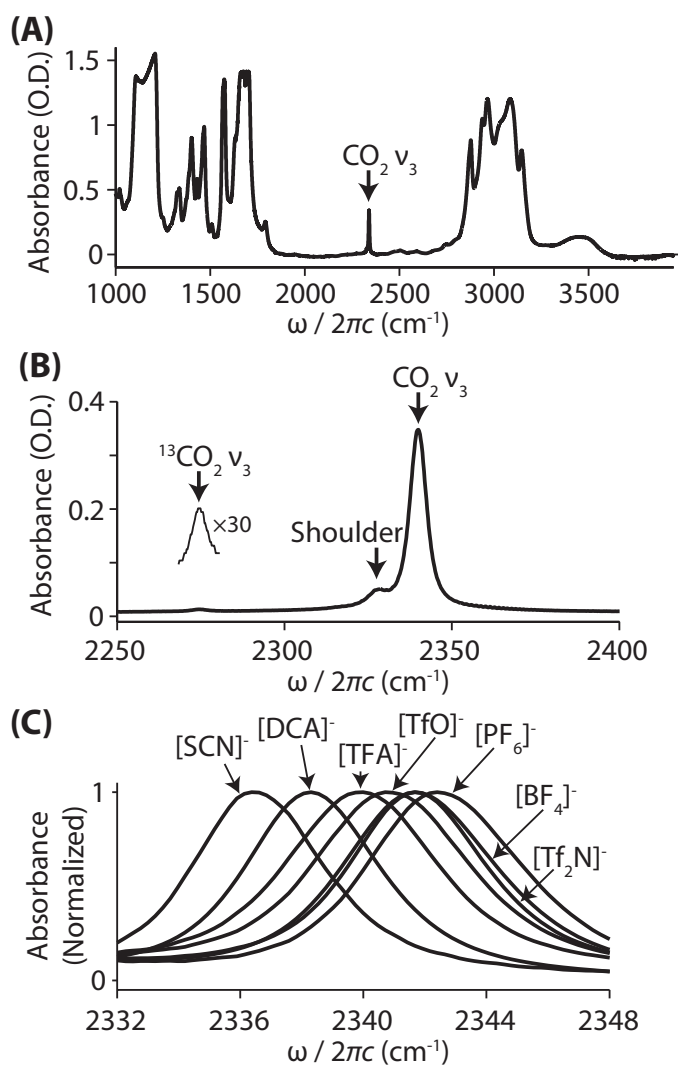


Figure 3: a) Absorption spectrum of  $\text{CO}_2$  in  $[\text{Im}_{4,1}][\text{TFA}]$  shows the intense antisymmetric stretch absorption at  $2340 \text{ cm}^{-1}$ ; b) the background subtracted spectrum is Lorentzian with a shoulder at  $2328 \text{ cm}^{-1}$ , the  $\nu_3$  band of the  $^{13}\text{C}$  isotopomer is located at  $2280 \text{ cm}^{-1}$ ; c) The average vibrational frequency of the  $\nu_3$  absorption of  $\text{CO}_2$  shifts for different anions and the same  $[\text{Im}_{4,1}^+]$  cation (background subtracted).

anion dominates CO<sub>2</sub> solubility in ionic liquids.<sup>39,53,101</sup> Anions studied were hexafluorophosphate (PF<sub>6</sub><sup>-</sup>), tetrafluoroborate (BF<sub>4</sub><sup>-</sup>), bis-(trifluoromethylsulfonyl)imide (Tf<sub>2</sub>N<sup>-</sup>), triflate (TfO<sup>-</sup>), trifluoroacetate (TFA<sup>-</sup>), dicyanamide (DCA<sup>-</sup>), and thiocyanate (SCN<sup>-</sup>). The cation in all experiments was 1-butyl-3-methylimidazolium ([Im<sub>4,1</sub>]).

The  $\nu_3$  center frequency progressively redshifts from a maximum of 2342.5 cm<sup>-1</sup> ([PF<sub>6</sub>]<sup>-</sup>) to a minimum of 2336 cm<sup>-1</sup> ([SCN]<sup>-</sup>). The shoulder moves with the main absorption band and stays  $\sim 12$  cm<sup>-1</sup> lower in frequency. Qualitatively, smaller, harder anions like [SCN]<sup>-</sup> and [DCA]<sup>-</sup> create a larger redshift than bulkier, softer anions like [Tf<sub>2</sub>N]<sup>-</sup> and [TfO]<sup>-</sup>, which is most likely a function of increased anionic charge density. The increased charge density could red-shift the CO<sub>2</sub> center frequency through an increased local electric field (Stark effect), through charge transfer, or through inductive effects. Quantum chemistry calculations (Section 2.4.2) help to disentangle the driving forces behind this qualitative trend.

The shoulder on the low frequency side of the main  $\nu_3$  transition could arise from several possible mechanisms, a “hot-band”, a multiple-quantum transition, or different chemical environments. The temperature dependence of this feature is important in discriminating among these possibilities.

Temperature-dependent FTIR demonstrates that the shoulder on the low frequency side of the main  $\nu_3$  transition is a hot band of the  $\nu_3$  mode (Figure 4A). Increasing temperature causes a decrease in intensity of the main peak and an increase in intensity of the shoulder, while conserving oscillator strength. The relative magnitude of the shoulder ( $\sim 10\%$  of the main band at room temperature) is similar to the expected relative excited state bending mode ( $\nu_2$ ) population predicted by a Boltzmann distribution.

We fit the spectra in Figure 4A to two Voigt profiles (one for the main peak, and one for the first shoulder seen on the 2D-IR).

A van’t Hoff analysis of the logarithm of the relative peak heights against  $1/T$  gives an activation energy of  $810 \pm 30$  cm<sup>-1</sup>. This value is near the energy of the  $\nu_2$  bending vibration (667 cm<sup>-1</sup>). Residual gas lines and nonlinearity of the detector contribute to the systematic error of this measurement. Nevertheless, the temperature dependence strongly suggests that the first shoulder is due to one quantum of the bending mode which is excited thermally and

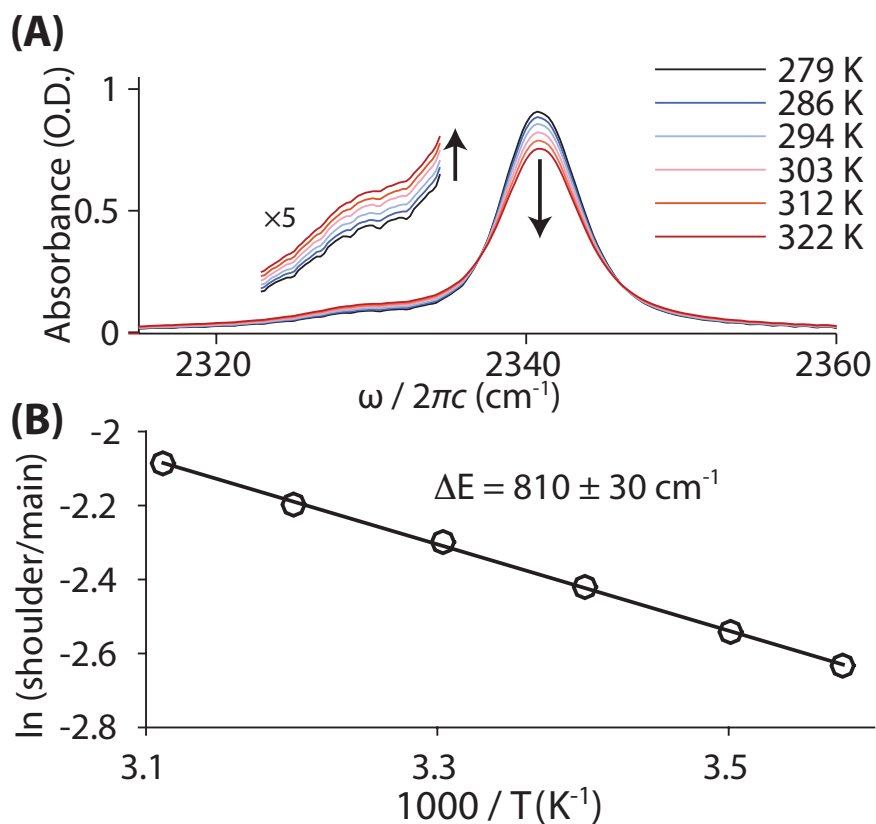


Figure 4: (A) Temperature dependence of the  $\nu_3$  spectrum in  $[\text{Im}_{4,1}][\text{TfO}]$  shows transition of intensity from the main peak to the shoulder with increasing temperature, indicating a temperature-dependent two-state transition. (B) Relative intensity of shoulder to the main band (based on fitting to Voigt profiles) follows a van't Hoff temperature dependence, indicating an energy barrier of around  $800 \text{ cm}^{-1}$ , which closely follows the prediction from the temperature dependence of bending mode population.

is anharmonically coupled to the  $\nu_3$  mode.

The temperature dependence and hot-band assignment are important components of our interpretation of the shape of the 2D-IR spectra (Section 2.4.5) and their time-dependence (Section 2.4.8).

## 2.4.2 Vibrational Frequency Calculations

Electronic structure calculations provide a rationale for the observed trend in vibrational frequencies.

Harmonic frequency calculations reproduce the general trend that  $\nu_3$  progressively red-shifts with decreasing anion size (Table 4). We simplify the solvated CO<sub>2</sub> structure to a gas-phase cluster consisting of one CO<sub>2</sub> with one cation-anion pair, with 1,3-dimethylimidazolium (Im<sub>1,1</sub>) as the cation. When scaled with the appropriate factor,<sup>99</sup> the simulations calculate vibrational frequencies within a few wavenumbers of the experimental value. The ordering of anions mostly agrees with experiment as well, the only outliers being [TFA]<sup>-</sup> and [SCN]<sup>-</sup>, which are located only 2.3 cm<sup>-1</sup> apart.

The level of agreement between experiment and theory is good, given that the condensed phase environment is neglected. That such a simple representation reproduces the general trends so well suggests that the interactions of CO<sub>2</sub> are dominated by local effects in its immediate surroundings. Future work will address the condensed phase effects by sampling representative structures from molecular dynamics simulations and repeating the analysis in the context of larger solvation shells.

Encouraged by the fact that the electronic structure calculations reproduce the experimental trends in  $\nu_3$  frequencies, we decompose the calculated vibrational frequencies into different components using absolutely localized molecular orbitals (ALMO), in analogy to ALMO energy decomposition analysis (ALMO-EDA).<sup>100,102,103</sup>

Unlike standard quantum chemical calculations, where individual molecular orbitals may delocalize over more than a single fragment, each ALMO is composed of atomic orbitals from a single fragment. This constraint allows us to control charge transfer between fragments and to quantify the interaction between fragments in physically intuitive terms.

The total vibrational frequencies  $\omega_{\text{tot}}$  and vibrational shifts  $\Delta\omega_{\text{int}}$  from gas phase (“free”) CO<sub>2</sub> for each mode  $\nu$  may be written as:

$$\omega_{\text{tot}} = \omega_{\text{free}} + \Delta\omega_{\text{int}}, \quad (2.3)$$

where

$$\Delta\omega_{\text{int}} = \Delta\omega_{\text{GEOM}} + \Delta\omega_{\text{FRZ}} + \Delta\omega_{\text{POL}} + \Delta\omega_{\text{CT}}. \quad (2.4)$$

$\Delta\omega_{\text{GEOM}}$  (geometric distortion) corresponds to the change in frequency caused by distortion of fragments from their free geometries to their cluster geometries,  $\Delta\omega_{\text{FRZ}}$  (the frozen orbital interaction) results from the combined electrostatic interaction and Pauli repulsion between the filled, unrelaxed orbitals of each fragment,  $\Delta\omega_{\text{POL}}$  (polarization) is due to the relaxation of a fragment’s orbitals in the field of the other fragments, and  $\Delta\omega_{\text{CT}}$  (charge transfer) is from occupied-virtual orbital donation between orbitals of different fragments.<sup>104</sup>

The cluster environment can affect the vibrational frequency of CO<sub>2</sub> in two different ways. The first depends on the anharmonic potential surface of a free CO<sub>2</sub> molecule. In a harmonic system, the spring constant,  $k$ , uniquely determines the vibrational frequency for all nuclear positions (i.e., the curvature of a quadratic potential is constant). CO<sub>2</sub>’s potential energy surface, however, is inherently anharmonic. Any change in geometry will thus cause a change in the  $\nu_3$  vibrational frequency. In other words, the cluster can change the  $\nu_3$  frequency just by shifting the location of minimum of the CO<sub>2</sub> potential ( $\Delta\omega_{\text{GEOM}}$ ). The second results from changes in the local curvature of CO<sub>2</sub>’s potential energy surface due to interactions with the surrounding cluster ( $\Delta\omega_{\text{FRZ}} + \Delta\omega_{\text{POL}} + \Delta\omega_{\text{CT}}$ ).

We focus on the following grouping of terms: (1) distortion of isolated CO<sub>2</sub> to its cluster geometry ( $\Delta\omega_{\text{GEOM}}$ ), (2) the combined frozen orbital and polarization contributions through use of ALMOs ( $\Delta\omega_{\text{FRZ}} + \Delta\omega_{\text{POL}}$ ), and (3) charge transfer between the fragments in the cluster ( $\Delta\omega_{\text{CT}}$ ).

The final frequency,  $\omega_{\text{tot}}$ , correlates most strongly with  $\Delta\omega_{\text{GEOM}}$  ( $R^2 = 0.96$ ), indicating that geometrical distortion of CO<sub>2</sub> dominates the frequency shift. Depending on the particular geometry that CO<sub>2</sub> adopts in the cluster,  $\Delta\omega_{\text{GEOM}}$  can vary by  $\pm 6 \text{ cm}^{-1}$

(Figure 5).<sup>a</sup> Electrostatics and charge transfer both change the local curvature of the potential energy surface; however, their effects are mostly uniform across the ionic liquids studied. When we turn off charge transfer between fragments,  $\nu_3$  blue-shifts on average by  $\langle \Delta\omega_{\text{FRZ}} + \Delta\omega_{\text{POL}} \rangle = +2.8 \pm 0.6 \text{ cm}^{-1}$ . Modelling the cluster geometry as a field of point charges increases this blue shift to  $+4.4 \pm 0.6 \text{ cm}^{-1}$  (Supplementary Information), which may indicate that a point-charge representation of the cluster over-polarizes the QM region.<sup>105,106</sup> Allowing charge transfer red-shifts the frequencies by  $\langle \Delta\omega_{\text{CT}} \rangle = -3.5 \pm 0.8 \text{ cm}^{-1}$ . Thus, both electrostatics and charge transfer change the local curvature of the  $\nu_3$  potential energy surface. The effects, however, are uniform and nearly cancel. The net result is that the inherent anharmonicity of the  $\text{CO}_2$  potential energy surface ultimately dominates the  $\nu_3$  frequency.

The geometrical distortion of the  $\text{CO}_2$  is driven by charge transfer. Using ALMOs and forbidding charge transfer during relaxation of the cluster removes the variance in  $\Delta\omega_{\text{GEOM}}$  (Figure 5B). Once again, both electrostatic and charge transfer interactions act uniformly on the frequency, leading to a negligible variance in  $\omega_{\text{tot}}$ . Without the geometrical distortion due to charge transfer, the vibrational frequencies for all clusters remain within  $\pm 2 \text{ cm}^{-1}$  and no longer follow the experimental trend.

The essential point is that charge transfer and electrostatics only affect the vibrational frequency indirectly, through their coupling into the equilibrium  $\text{CO}_2$  nuclear geometry. The direct effects of electrostatics and charge transfer on the potential energy surface (and thus the effective spring constant) of  $\nu_3$  are uniform across ionic liquids studied and counteract each other.

This DFT-based analysis may over-emphasize charge transfer to some extent. It is well known that DFT over-delocalizes electrons due to self-interaction error, which might exaggerate the amount of charge transfer. In the context of gas-phase anion clusters, both DFT<sup>107</sup> and post-Hartree–Fock methods<sup>108</sup> have also identified charge transfer as a driver of geometrical distortion of  $\text{CO}_2$ , so we expect that our qualitative picture will be robust with respect to the theoretical method. Future work will explore the quantitative differences between

---

<sup>a</sup>Since  $\omega$  normally refers to an angular frequency, this quantity might more rigorously be called  $\Delta\omega_{\text{GEOM}}/2\pi c$ ; however, for convenience of notation, we chose to use spectroscopic units uniformly in this section. That is, all frequencies are given in wavenumbers, rather than  $\text{rads}^{-1}$ .

the different methods. Furthermore, although our cluster results agree well with experiment, we expect that the extended solvation environment neglected here may introduce screening effects not captured at the cluster level.<sup>109</sup> A more sophisticated study of solvation effects could determine whether our finding, that charge transfer dominates the geometric effects which differentiate CO<sub>2</sub> vibrational frequency shifts, is transferable to bulk ionic liquids.

These calculations demonstrate that the  $\nu_3$  frequency shift reflects a complex interplay between geometry, electrostatics, and charge transfer. The geometrical distortion is the most important factor in determining the final variation in vibrational frequency with anion identity; however, the equilibrium geometry of the CO<sub>2</sub> itself ultimately depends on inductive effects, such as charge transfer.

### 2.4.3 Frequency, Geometry, and Charge Transfer Discussion

To gain chemical insight into the charge transfer process, one can analyze the donor and acceptor orbitals of CO<sub>2</sub> and the ionic liquid. This analysis sheds light on the nature of the geometrical distortion, which, in turn, leads to a simple model for the vibrational frequencies in terms of a few geometrical parameters.

We employ a complementary occupied-virtual orbital pair (COVP) analysis<sup>103</sup> to gain insight into the electronic structure effects underlying the charge transfer between CO<sub>2</sub> and the solvent. COVPs provide a way to visualize charge transfer effects, where “each COVP corresponds to an occupied orbital on one molecule donating charge to one specific (complementary) virtual orbital on the other molecule.”<sup>110</sup> These orbital pairs are constructed from a singular value decomposition of the occupied-virtual mixing matrix  $\mathbf{X}$  that describes charge transfer between the (polarized) fragments upon removing the ALMO fragment localization constraint. While  $\mathbf{X}$  contains, in general, excitations between all possible occupied-virtual pairs, the COVP representation of  $\mathbf{X}$  is diagonal and thus provides the most compact possible basis for describing charge transfer. The associated singular values allow one to assess the contribution of a particular COVP to charge transfer, and typically one finds that only one or a few orbital pairs dominate the effect (as shown below in this study). The COVP analysis thus allows to conveniently identify and visualize the electron donor/acceptor orbitals



Table 4: Experimental and calculated  $\nu_3$  vibrational frequencies ( $\text{cm}^{-1}$ ) for  $\text{CO}_2$  in various imidazolium (experimental [Im<sub>4,1</sub>], calculations [Im<sub>1,1</sub>]) ionic liquids. Calculations are carried out using a gas phase anion-cation- $\text{CO}_2$  cluster at 0 K. The level of agreement between calculations and experimental results indicates that interactions of  $\text{CO}_2$  are dominated by local effects in its immediate surroundings.

Anion	Expt. Freq.	Calc. Freq.	Calc. Freq. (scaled)
[PF <sub>6</sub> ] <sup>-</sup>	2342.5	2437.7	2346.8
[Tf <sub>2</sub> N] <sup>-</sup>	2341.7	2435.8	2344.9
[BF <sub>4</sub> ] <sup>-</sup>	2341.7	2434.7	2343.9
[TfO] <sup>-</sup>	2340.9	2431.9	2341.2
[TFA] <sup>-</sup>	2339.9	2429.3	2338.7
[DCA] <sup>-</sup>	2338.4	2430.5	2339.8
[SCN] <sup>-</sup>	2336.5	2430.0	2339.4

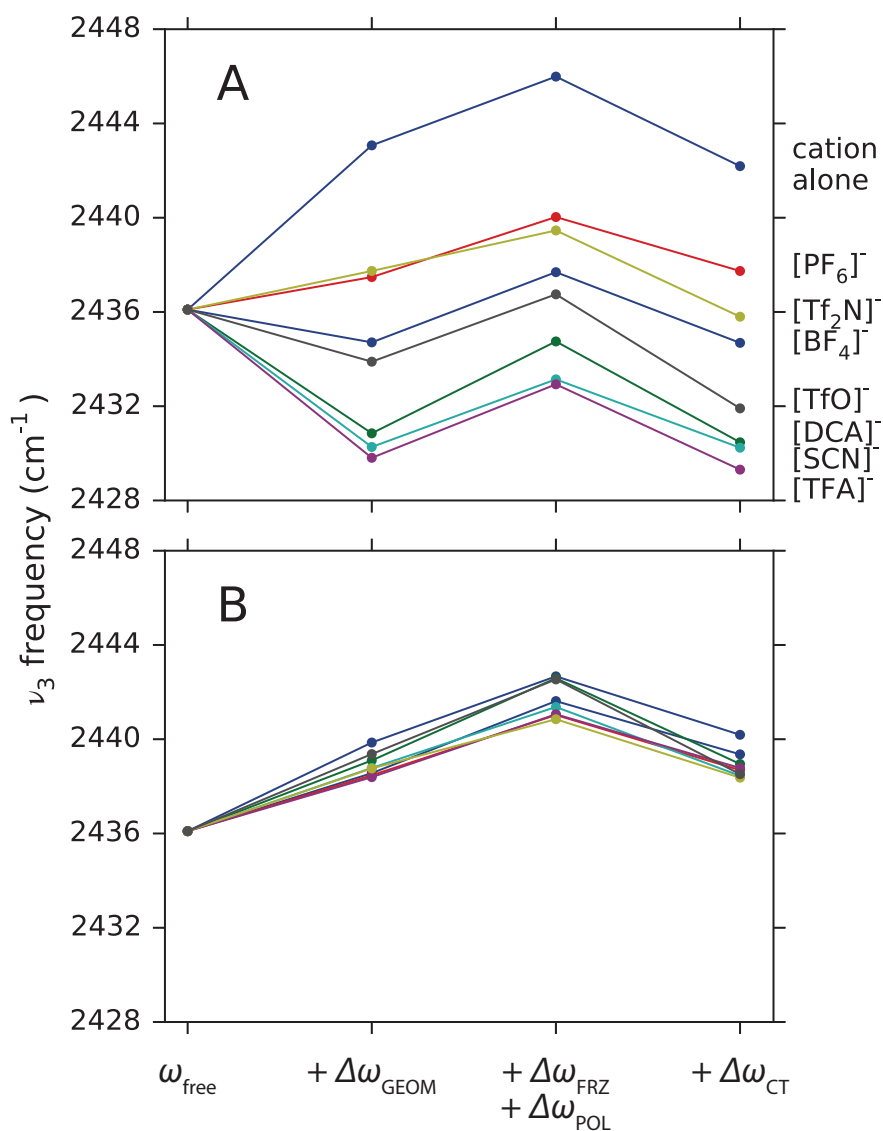


Figure 5: Decomposition of the geometric, electrostatic, and charge transfer contributions to  $\text{CO}_2$   $\nu_3$  vibrational frequency shifts. (A) The cluster geometry (including  $\text{CO}_2$ ) was optimized while allowing charge transfer. (B) The cluster geometry was optimized using ALMOs to forbid charge transfer between fragments. The final frequency ( $\omega_{\text{tot}} = \omega_{\text{free}} + \Delta\omega_{\text{GEOM}} + \Delta\omega_{\text{FRZ}} + \Delta\omega_{\text{POL}} + \Delta\omega_{\text{CT}}$ ) correlates most strongly with the change in frequency from geometric distortion ( $\Delta\omega_{\text{GEOM}}$ ).

Table 5: Charge transfer from the anion to CO<sub>2</sub> and from CO<sub>2</sub> to the cation both contribute to the geometrical distortion of the CO<sub>2</sub>. The most important geometrical degrees of freedom are the bend angle,  $\theta$  and the sum of the two carbonyl bond lengths,  $L$ .

cluster	CT to CO <sub>2</sub> (me <sup>-</sup> )	CT from CO <sub>2</sub> (me <sup>-</sup> )	$\theta$ (°)	$L$ (Å)
[Im <sub>1,1</sub> ] <sup>+</sup>	0.079	2.251	0.07	2.3374
[PF <sub>6</sub> ] <sup>-</sup>	3.336	1.012	3.79	2.3379
[Tf <sub>2</sub> N] <sup>-</sup>	2.493	1.603	2.64	2.3380
[BF <sub>4</sub> ] <sup>-</sup>	4.558	1.264	4.51	2.3385
[TfO] <sup>-</sup>	3.009	2.339	4.01	2.3389
[TFA] <sup>-</sup>	5.131	1.618	5.22	2.3394
[DCA] <sup>-</sup>	3.376	2.259	4.99	2.3394
[SCN] <sup>-</sup>	3.317	1.323	4.35	2.3395

pair(s) that significantly contribute to charge transfer.

Charge transfer to the CO<sub>2</sub> originates from occupied orbitals on the anions. The charge is accepted by virtual orbitals on the CO<sub>2</sub>. The virtual orbitals are a linear combination of the LUMO and LUMO+1 orbitals, which have  $\sigma^*$  and  $\pi^*$  character, respectively. Charge transfer from the anion has a strong linear correlation to the bend angle ( $R^2 = 0.84$ ). Mechanistically, bending the CO<sub>2</sub> allows  $\sigma^*$  and  $\pi^*$  to mix, lowers the energy of the acceptor orbital, and also maximizes the spatial overlap with the donor orbital. The amount of charge transferred (3 to 5 me<sup>-</sup>) is small, but the resulting bend angle (3 to 5°) can be substantial (Table 5).

CO<sub>2</sub> also donates charge back to the ionic liquid cluster, primarily to the cation. The amount of charge donated by the CO<sub>2</sub> to the cation (-1 to -2 me<sup>-</sup>) is typically less than the charge donated from the anion, but nevertheless has specific consequences for the resulting geometry. The donor orbital from the CO<sub>2</sub> is a mixture of  $\sigma$ -bonding and  $\pi$ -nonbonding character. The charge donation from the CO<sub>2</sub> is linearly correlated to the carbonyl bond length differences ( $R^2 = 0.85$ ). For the [Im<sub>1,1</sub>][TfO] cluster, which has the highest charge

transfer from CO<sub>2</sub> to the cation (2.32 me<sup>-</sup>), the  $\sigma$  character of the donating orbital causes the carbonyl nearest the cation to lengthen from the gas-phase 1.169 Å to 1.175 Å while the distal carbonyl contracts to 1.163 Å.

The effects of charge transfer to and from the CO<sub>2</sub> can be incorporated into a simple model of the vibrational frequencies. The model is constructed in a vibrational local-mode basis. The effective one exciton vibrational Hamiltonian is

$$H^{(1)} = \begin{pmatrix} h\omega_1 & \beta \\ \beta & h\omega_2 \end{pmatrix} \quad (2.5)$$

where  $\omega_i$  is the local mode frequency of carbonyl  $i$  and  $\beta$  is the coupling between the two local modes. The diagonalized hamiltonian gives symmetric and antisymmetric linear combinations of the local modes as the vibrational eigenstates with energies  $h\omega_s$  and  $h\omega_a$ . The splitting between symmetric and antisymmetric vibrations is twice the coupling constant,  $h(\omega_s - \omega_a) = 2\beta$ , and the average frequencies in local and normal modes are also equal  $h(\omega_a + \omega_s)/2 = h(\omega_1 + \omega_2)/2 = \alpha$ .

The bend in the CO<sub>2</sub> determines the change in the coupling constant between the local modes,  $\beta$  ( $R^2 = 0.94$ ). The motion of the central carbon atom is the primary motion that couples the two carbonyls. When they are collinear, the motion of one carbonyl directly influences the other. Bending the CO<sub>2</sub> means the carbonyls are no longer collinear, which means that the projection of one local vibration on the other decreases. This decrease, in turn, decreases the effective coupling constant.

The sum of the CO<sub>2</sub> bond lengths  $L$ , is correlated to the  $\alpha$  ( $R^2 = 0.998$ ). The dependence of  $\alpha$  on the geometry of the molecule is not as straightforward than that of  $\beta$ . In the strong coupling limit,  $\beta \gg |h\omega_2 - h\omega_1|$ , the symmetric and antisymmetric stretching frequencies only depend on the average frequencies,  $(h\omega_1 + h\omega_2)/2$ . As such, the geometrical asymmetry induced by charge transfer from the CO<sub>2</sub>, which weakens one bond and strengthens the other, is largely averaged out. Only a weak correlation between the charge donated from the CO<sub>2</sub> and the  $\alpha$  remains. The sum of the bond lengths,  $L$ , however, reports exactly this difference in total bonding strength. Though the changes in  $L$  are minute ( $\sim 0.001$  Å), the effect on the vibrational frequencies is not.

Because of the strong linear correlation with these two geometrical variables, we propose that the simplest model of the scaled vibrational frequencies is  $\alpha(L)$  and  $\beta(\theta)$ , where

$$\alpha(L) = 9523.8 \text{ cm}^{-1} - (3415.3 \text{ cm}^{-1} \text{ \AA}^{-1}) L \quad (2.6)$$

and

$$\beta(\theta) = -515.7 \text{ cm}^{-1} + (1.12 \text{ cm}^{-1} \text{ deg}^{-1}) \theta, \quad (2.7)$$

where  $\alpha$  and  $\beta$  are in units of  $\text{cm}^{-1}$ ,  $L$  is in  $\text{\AA}$ , and  $\theta$  is in degrees. The one exciton Hamiltonian

$$H^{(1)} = \begin{pmatrix} \alpha(L) & \beta(\theta) \\ \beta(\theta) & \alpha(L) \end{pmatrix}, \quad (2.8)$$

reproduces the scaled harmonic frequencies with an RMS error of  $1.2 \text{ cm}^{-1}$  and the experimental frequencies with an RMS error of  $4 \text{ cm}^{-1}$ .<sup>b</sup>

Following a detailed investigation of the charge transfer, the participating orbitals, and the effects of charge transfer on the  $\text{CO}_2$  geometry, we have put forward a simple model that successfully reproduces the calculated vibrational frequencies and, to a reasonable extent, the experimental frequencies, with no free parameters.

#### 2.4.4 2D-IR Spectroscopy Overview

While linear spectroscopy shows the sensitivity of  $\text{CO}_2$  to its local environment, it cannot address the ultrafast dynamics of the chromophore. 2D-IR, however, directly reports on the dynamic structural relaxation around  $\text{CO}_2$ . 2D-IR of  $\text{CO}_2$  in  $[\text{Im}_{4,1}][\text{TFA}]$  introduces many of the features that are general across all of the ionic liquids studied.

The 2D spectra of the  $\nu_3$  mode of  $\text{CO}_2$  in  $[\text{Im}_{4,1}][\text{TFA}]$  (Figure 6A and B) show the main  $\nu_3$  band, two diagonal shoulders, and cross-peaks between them. The observed peaks cannot be explained without also keeping track of the states of the  $\text{CO}_2$  symmetric stretch and bending modes in the vibrational state. The total vibrational wavefunction is specified by  $|\nu_1\nu_2^l\nu_3\rangle$ , where  $\nu_1$  is the number of quanta in the symmetric stretch,  $\nu_2$  is the number

---

<sup>b</sup>See section 2.7 for 2D-IR spectra of  $\text{CO}_2$  in each ionic liquid, as well as a comparison of center line slope and global fitting results for correlation times, and more detailed computational results. Additionally, there is information on fitting of the one exciton Hamiltonian model for unscaled frequencies, and vibrational frequency calculations of  $\text{CO}_2$  in a field of point-charges that approximates the cluster geometry.

of quanta in the bending mode,  $l$  is the vibrational angular momentum quantum number of the bending mode, and  $\nu_3$  is the number of quanta in the antisymmetric stretch.

The main band consists of a pair of intense peaks corresponding to the  $|00^00\rangle \rightarrow |00^01\rangle$  and  $|00^01\rangle \rightarrow |00^02\rangle$  (1a and 1b) transitions, separated by the anharmonicity of  $\nu_3$  ( $\sim 24 \text{ cm}^{-1}$ ). Due to ground state bleach, 1a appears as a negative (blue) feature, while 1b, from excited state absorption, is a positive (red) feature. The  $\nu_3$  shoulder appears as a pair of small peaks (2a and b), shifted along the diagonal by  $-12 \text{ cm}^{-1}$  in  $\omega_1$  and  $\omega_3$ . A second apparent shoulder, not seen in FTIR, presents as a pair of smaller peaks (1e and 1f), shifted along the diagonal from the main band by  $-24 \text{ cm}^{-1}$  in  $\omega_1$  and  $\omega_3$ . At early population times, there is an apparent cross-peak between the main peak and second shoulder (1c). Cross-peaks grow in between the first shoulder and main peak (2-1a and b, 1-2b) over  $t_2$ . The expected cross-peak 1-2a cannot be seen due to cancellation with the overwhelming opposite signal from the 1b.

The combination of  $\text{CO}_2$ 's high molar absorptivity ( $\sim 1000 \text{ M}^{-1} \text{ cm}^{-1}$ ) and the  $\varepsilon^2$  dependence of the third-order signal causes the 2D signal from  $\text{CO}_2$  to dominate the spectrum in intensity. Contributions from solvent overtones in the background are smaller than the level of noise in the spectrum. Thus, the 2D signal reflects the vibrational modes of  $\text{CO}_2$ , rather than solvent background, and the solvent can only affect the 2D spectrum through intermolecular couplings with  $\text{CO}_2$  vibrational modes.

The unambiguous spectral diffusion of 1a and 1b over the 50 ps population time demonstrates that the line is not entirely in the motional narrowing (homogeneous) limit. That is,  $\nu_3$  will allow us to resolve the dynamics of structural relaxation around  $\text{CO}_2$ .

The observations of a complex pattern of peaks and of spectral diffusion on a tens of picoseconds timescale for  $\text{CO}_2$  in  $[\text{Im}_{4,1}][\text{TFA}]$  are general features of the spectra in all of the tested ionic liquids and are described in detail in the next two sections.

#### 2.4.5 2D-IR Shoulders and Cross-Peaks

Careful analysis of the relative kinetics of the diagonal peaks and cross-peaks (Figure 7) provides insight into how the coupling of  $\text{CO}_2$  vibrational modes and their stochastic dynamics

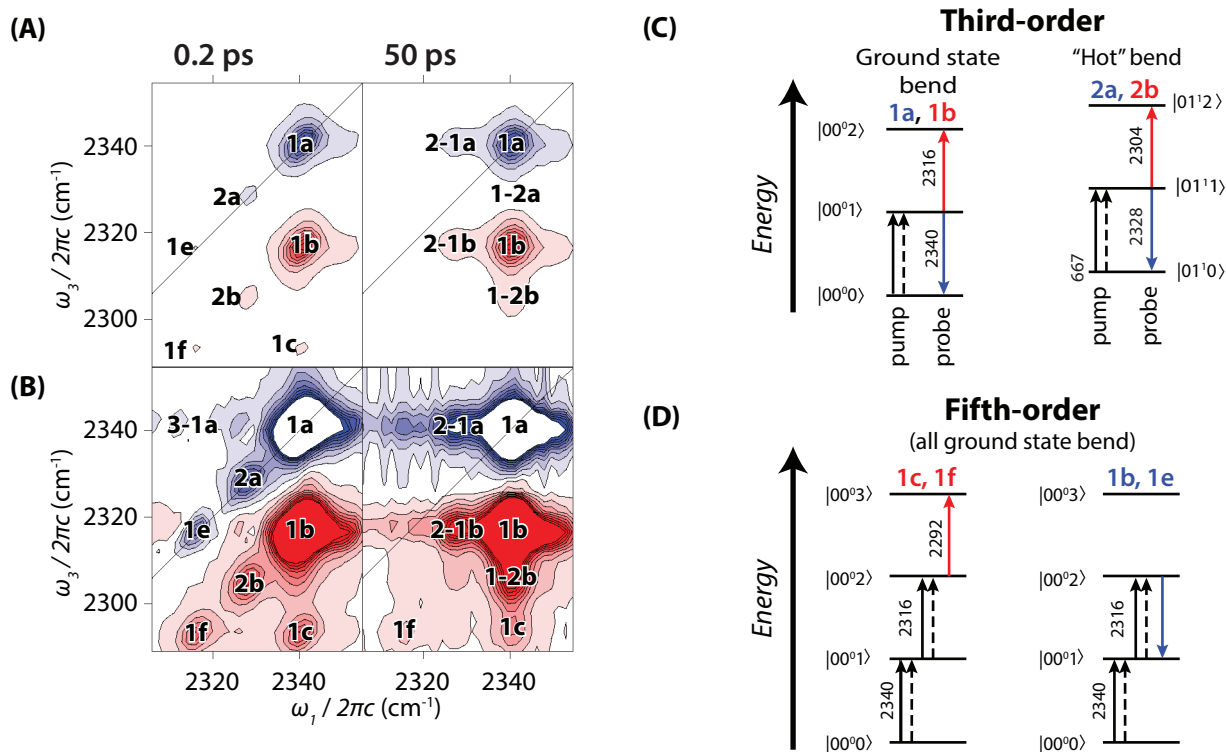


Figure 6: **(A)** 2D-IR spectrum of CO<sub>2</sub> in [Im<sub>4,1</sub>][TFA] at  $t_2 = 0.2$  and 50 ps. The peak labels correspond to transitions in (C) and (D). **(B)** The same spectrum as (A) with contours limited to 10% of the maximum in the  $z$ -direction. Structure of the diagonal shoulders and cross-peaks can be seen much more readily. **(C)** and **(D)** Vibrational energy level diagrams for observed third-order (C) and fifth-order (D) bands of CO<sub>2</sub> in [Im<sub>4,1</sub>][TFA]. Quantum numbers correspond to  $|\nu_1\nu_2^l\nu_3\rangle$ . Transition frequencies are labeled in wavenumbers (cm<sup>-1</sup>), and a label corresponding to the peaks in A and B. The color of the label indicates whether the expected peak is negative (blue) or positive (red).

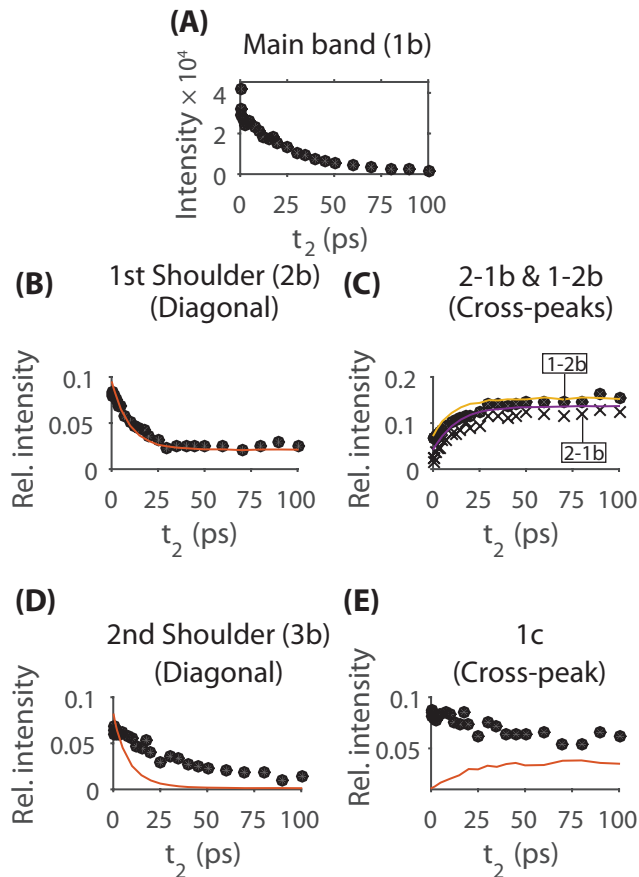


Figure 7: Kinetic analysis of [Im<sub>4,1</sub>][TFA]. (A) Absolute intensity of the main peak (1b) with increasing  $t_2$ . Decreased intensity results from both orientational and population relaxation. (B)-(E) show intensities relative to A. Discrete data points represent experimental data, while lines show data from stochastic simulations (Section 2.4.8). (B) and (C), which result from the first diagonal shoulder (2b) and its cross-peaks with the main band (1-2b / 2-1b) show good agreement between experimental and simulated kinetics. (D) and (E), which result from the second diagonal shoulder (3b) and the apparent cross-peak (1c) show distinct kinetics, which cannot result from the same stochastic bending mode fluctuations, and point to a direct transition.



create the observed spectrum.

Orientalional and vibrational relaxation of CO<sub>2</sub> cause a decrease in spectral intensity over  $t_2$ . The main peak (1b) shows an initial rapid decrease in intensity from orientation relaxation, followed by a slower decay over  $\sim 100$  ps, from vibrational relaxation processes. Polarization-controlled studies can remove the contribution from orientational relaxation, and provide an explicit assessment of vibrational relaxation rate.<sup>69,111</sup>

The shoulders and cross-peaks are significantly weaker than the main peak, and also relax over  $t_2$ ; however, analysis of their kinetics relative to the main band reveals the underlying stochastic dynamics of CO<sub>2</sub> vibrational modes.

The first diagonal shoulder (2b) decays more quickly than the main band, decreasing to a minimum intensity at  $\sim 25$  ps followed by a (relative) steady state (Figure 7). Cross-peaks between the first shoulder and the main band (2-1b and 1-2b) start at a minimum intensity, and then increase, before reaching a relative steady state by 25 ps. This behavior points to dynamic exchange between  $|00^01\rangle$  and  $|01^11\rangle$ , which give rise to the diagonal peaks (Section 2.4.8).

The second diagonal shoulder (3b) decreases in intensity relative to the main band, but more slowly than the first shoulder. The apparent cross-peak 1c decreases in relative intensity over  $t_2$  with slower kinetics than the second shoulder. This behavior contrasts with that of the cross-peaks 1-2b and 2-1b, and indicates that 1c arises from a direct vibrational transition, rather than dynamic exchange.

### 2.4.6 Peak Assignment

The Dunham expansion for anharmonically coupled vibrational modes provides a theoretical framework for building an analysis of coupled vibrational modes:

$$E = \sum_i \hbar\omega_i \left( n_i + \frac{1}{2} \right) + \sum_{ij} x_{ij} \left( n_i + \frac{1}{2} \right) \left( n_j + \frac{1}{2} \right) \quad (2.9)$$

where  $\omega_i$  is the frequency of the mode,  $n_i$  is the number of quanta in the mode, and  $x_{ij}$  are anharmonic coupling constants. It directly follows that the energy of a particular mode is

given by:

$$\nu_{n_k \rightarrow n_k+1} = \nu_k + 2n_k x_{kk} + \sum_{i \neq k} x_{ij} n_i, \tag{2.10}$$

which implies that transition energy will decrease by  $x_{ij}$  for every quantum of energy in an anharmonically coupled mode.<sup>112</sup> Gas phase vibrational calculations predict  $x_{23} \approx -12 \text{ cm}^{-1}$  per quantum in  $\nu_2$ .<sup>113,114</sup>

The first shoulder can then be explained by anharmonic coupling of excited state bending modes with the asymmetric stretching mode. This coupling causes a frequency shift of  $-12 \text{ cm}^{-1}$ , and creates a shoulder on the linear and the 2D spectra (peaks 2a/2b on 2D-IR). Stochastic fluctuations in thermally populated bending modes cause dynamic 2D-IR cross-peaks.

This mechanism for dynamic cross-peaks is not unique to  $\text{CO}_2$  in ionic liquids; however, it is a nearly ideal system in which to observe and analyze it. The narrow total linewidth ( $\sim 6 \text{ cm}^{-1}$ ), combined with a large anharmonic coupling constant ( $x_{23} = -12 \text{ cm}^{-1}$ ), leads to clear segregation of the resulting frequencies into distinct peaks. The difference in energy between the ground and first excited state  $\nu_2$  modes is only around  $3k_B T$ , which is low enough to be thermally accessible, but high enough to be quantized. Finally, the rate of stochastic fluctuations in bending mode states is slow enough to preserve distinct diagonal bands at short population times, but fast enough to allow dynamic cross-peaks over the timescale of the experiment ( $\sim 100 \text{ ps}$ ).

Alternative hypotheses can be ruled out. There are two possibilities that need to be addressed. First, the shoulder is often assigned to a multiquantum transition. Second, the shoulder and cross-peaks could be due to chemical exchange.

The first  $\nu_3$  shoulder in our spectra is also present in the FTIR of  $\text{CO}_2$  dissolved in organic liquids and many polymers, and is often attributed to a multiquantum transition  $\nu_3 + \nu_2 - \nu_2$ ,<sup>87,115–118</sup> where the difference in energy is attributed to splitting of normally degenerate  $\text{CO}_2$   $\nu_2$  and  $\bar{\nu}_2$  bending modes. (This combination should perhaps be written as  $\nu_3 + \nu_2 - \bar{\nu}_2$  or  $\nu_3 + \bar{\nu}_2 - \nu_2$ , but is typically presented without distinguishing between the two bending modes.) This hypothesis fails on several accounts. First, such a multiquantum transition is harmonically forbidden, so the transition dipole moment should be significantly lower than

for the one quantum  $\nu_3$  transition. The intensity of the shoulder, however, directly follows a Boltzmann distribution for the  $\nu_2$  bending modes, which implies that the magnitude of the transition dipole moment is roughly equal for both the fundamental and this multiquantum transition.

Second, with this mechanism, a shoulder on the high frequency side of the fundamental should accompany the observed shoulder on the low frequency side. That is, there is no clear reason why the transition  $\nu_3 + \nu_2 - \bar{\nu}_2$ , would be strongly allowed, but  $\nu_3 + \bar{\nu}_2 - \nu_2$  would not. Third, this hypothesis assumes that the splitting of the bending modes is an identical  $12\text{ cm}^{-1}$ , but this is not the case.<sup>88</sup>

Finally, in the 2D spectrum, this combination band would give rise to cross-peaks between the first shoulder and the main band (since each a  $\text{CO}_2$  molecule could undergo either transition with excitation), which would be present at the earliest population times, and would only relatively decay (rather than relatively increase) with  $t_2$ . This behavior contradicts the observed spectral kinetics.

The temperature dependence of the shoulder (Figure 4) also excludes different chemical environments (such as multiple equilibrium geometries of a  $\text{CO}_2$ -anion interaction) undergoing chemical exchange. In this hypothesis, the most intense feature would be due to free  $\text{CO}_2$  and the shoulder to  $\text{CO}_2$  with a stronger chemical interaction. The growth of the cross-peaks would correspond to the exchange of these populations. However, the free  $\text{CO}_2$  band should be entropically favored and increase with increasing temperature while the shoulder should be enthalpically favored and decrease with increasing temperature. The opposite is observed, so this hypothesis can be ruled out. Additionally, the equal relative energy spacing of the diagonal shoulders from the main band for  $\text{CO}_2$  in every ionic liquid studied strongly suggests that the additional peaks arise from the  $\text{CO}_2$  itself, rather than from distinct chemical environments.

In contrast, the explanation of anharmonic coupling between  $\nu_2$  and  $\nu_3$  fits the temperature dependence, the transition dipole scaling, and accounts for the cross-peak kinetics. In our picture, each step requires only a one quantum transition for each peak, explains the presence of a shoulder on only one side of the fundamental, and correctly reproduces the cross-peak kinetics (over  $t_2$ ) between the first shoulder and main band due to thermal

excitation and de-excitation of the bend.

The apparent cross-peak 1c, as well as the second apparent shoulder (1e and 1f) are fifth-order signals, as can be shown by pump power-dependence, frequencies, and sign show that it is a fifth-order signal, as are several other features. Assignments of the  $\nu_3$  2D-IR spectrum of carbon dioxide in 1-butyl-3-methylimidazolium trifluoroacetate [Im<sub>4,1</sub>][TFA], are given in Figure 6 and Table 6).

The magnitude of third-order signal is linear in pump light intensity, since there are two pump electric field interactions, while that of fifth-order signal (with four electric field interactions) is quadratic. We assessed the magnitude of each peak when the pump power was changed by a factor of two (Table 6). There is a clear distinction between the pump power dependence of the third order signals (1a, 1b, 2a, 2b, and their population exchange cross-peaks) and the fifth order signal (1c, 1e, 1f, and 31a). The reported intensity ratio is the ratio of volumes of a single peak when the pump power doubles. Thus, for the main ‘red’ peak (1b:  $\nu_3$  excited state absorption with ground state  $\nu_2$ ), peak volume goes up by a factor of  $2.2 \pm 0.1$  when pump power doubles. Peak 1c’s volume, however, increases by a factor of  $3.3 \pm 0.2$ .

The fifth-order perturbative pathways have previously been described by Garrett-Roe and Hamm for 3D-IR (five IR pulse) spectroscopy.<sup>86</sup> The assignment of a “name” for each fifth-order pathway (Table 4 & Figure 8) follows the scheme used by Garrett-Roe and Hamm, where pathways are described by listing the vibrational quantum numbers of the three coherent states that contribute to them. Non-rephasing diagrams are shown for all pathways.

Since a 2D-IR experiment only has two coherence times (which give two frequency axes), we can only resolve two of the three coherences in the fifth-order pathway. When up-pumping occurs during the first two (“pump”) pulses, either  $t_1$  or  $t_3$  will be unresolved. Thus, each pathway can give two spectral peaks on a 2D-IR spectrum, if the coherent frequencies in  $t_1$  and  $t_3$  differ. For fifth-order pathways in Table 4, the coherence noted in parenthesis does not contribute to the observed peak, since oscillation of the first coherence will not be observed when there are multiple electric field interactions during a single 100 fs pulse. That is, the pathway 10|21|32 contributes to peaks 1c and 1f. Peak 1f results from up-pumping during the first pump pulse, (10) |21|32, and thus only the  $|2\rangle\langle 1|$  and  $|3\rangle\langle 2|$  coherences are

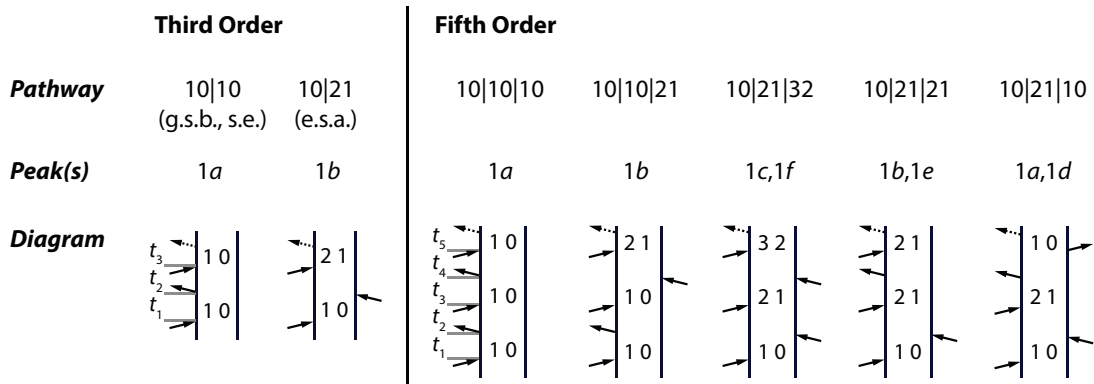


Figure 8: Double-sided Feynman diagrams, pathway labels, and peaks for third- and fifth-order peaks in the CO<sub>2</sub> 2D-IR spectrum (non-rephasing pathways shown). Labels for pathways correspond to those used by Garrett-Roe and Hamm in their description of purely absorptive 3D-IR spectra.<sup>86</sup> In a three optical pulse experiment (like 2D-IR), only two of the three coherences of a fifth-order signal can be resolved. Thus, depending on whether up-pumping occurs during the first or second pump pulse, either  $t_1$  or  $t_3$  will be unresolved in the fifth-order pathways. This effect can lead to multiple peaks on the 2D-IR spectrum from a single fifth-order pathway.

observed. Peak 1c results from up-pumping during the second pulse,  $10| \langle 21 | \langle 32$ , and thus only the  $|1\rangle \langle 0|$  and  $|3\rangle \langle 2|$  coherences are observed. The center frequency and sign of the peak amplitude for each observed fifth order peak matches that predicted by the corresponding fifth-order pathway (Table 4).

The peaks 1c, 1e, 1f, and 3-1a are fifth-order signals, not cascading third-order signals. Two of the peaks, 1c and 1f, cannot be generated by a cascade, because they involve walking up the vibrational ladder to a  $|3\rangle \langle 2|$  coherence. Given the presence of these unambiguously direct fifth-order signals, we would expect to find spectra contributions from other fifth-order pathways. Peaks 1e and 3-1a are located where we would predict additional fifth-order signal. The correspondence of the sign of the signal to those predicted by a direct fifth-order pathway is also important, because the relative signs of cascading third-order signals and fifth-order signals are opposite (due to the difference of  $i^2$  in the pre-factor in the infrared). Furthermore, the relative magnitudes follow the predictions based on the various pathways through population states and harmonic transition dipole moment scaling.<sup>86</sup> Finally, the relative peak intensities, including those of direct third-order signals, do not substantially vary with the concentration of the chromophore. Cascaded signals scale proportionally to  $c^2$ , and thus would not scale with the other peaks on the spectrum.

#### 2.4.7 Modelling of the Main Band

The 2D peak of a specific vibrational mode encodes the frequency-fluctuation correlation function (Equation 2.1) in its lineshape. Lineshape analysis can extract the timescale of structure relaxation around that mode, by quantifying spectral diffusion, or change in diagonal character (or ellipticity), of a peak over  $t_2$ .

The intense  $\nu_3$  peak clearly exhibits spectral diffusion in each ionic liquid studied (Figure 9). Qualitatively, at early population times, the main  $\nu_3$  peaks have diagonal character. As a function of the population time,  $t_2$ , the peaks become rounder. The rate of  $\nu_3$  spectral diffusion varies in the ionic liquids tested, indicating a broad range of timescales for structural relaxation in the different solvents. The rate is slowest in  $[\text{Im}_{4,1}][\text{PF}_6]$  (Figure 9A) and fastest in  $[\text{Im}_{4,1}][\text{DCA}]$  (Figure 9C). The vibrational relaxation time is slow enough to allow

Table 6: Peak parameters related to the assignment of peaks in the 2D-IR spectrum of CO<sub>2</sub> in [Im<sub>4,1</sub>][TFA] to third-order or fifth-order signal. Pathways are labeled with the non-rephasing coherences they exhibit. Parenthetical coherences are not observed due to the timing of up-pumping (thus, 10|(21)|32 indicates a pathway with observed 2D-IR frequencies of  $\omega_1, \omega_3 = \omega_{01}, \omega_{23}$ ).

Peak Label	Center ( $\omega_1, \omega_3$ )	Peak Vol. Ratio <sup>1</sup>	Sign	Order	Pathway <sup>2</sup>
1a	(2341.5, 2341.5)	$2.20 \pm 0.01$	-	3	10 10 (g.s.b./s.e.)
1b	(2341.5, 2317.5)	$2.20 \pm 0.02$	+	3	10 21 (e.s.a.)
2a	(2329.5, 2329.5)	$1.90 \pm 0.03$	-	3	‘hot’ g.s.b./s.e.
2b	(2329.5, 2305.5)	$2.5 \pm 0.1$	+	3	‘hot’ e.s.a.
21a	(2329.5, 2341.5)	$1.7 \pm 0.2$	-	3	pop. exch.
21b	(2329.5, 2317.5)	$2.40 \pm 0.04$	+	3	pop. exch.
1e	(2317.5, 2317.5)	$3.8 \pm 0.2$	-	5	(10) 21 21
1f	(2317.5, 2293.5)	$3.5 \pm 0.1$	+	5	(10) 21 32
1c	(2341.5, 2293.5)	$3.3 \pm 0.2$	+	5	10 (21) 32
31a	(2317.5, 2341.5)	$3.4 \pm 0.4$	+	5	(10) 21 10

<sup>1</sup> Factor by which the volume of the indicated peak increased when pump power was doubled.

<sup>2</sup> g.s.b./s.e.: Ground state bleach / stimulated emission. e.s.a.: Excited state absorption. ‘Hot’ peaks refer to peaks resulting from thermal excitation of the  $\nu_2$  bending mode. Numerals for Feynman pathways of fifth-order peaks from Garrett-Roe and Hamm.<sup>86</sup> Subscripted ‘a’ and ‘b’ indicate whether up-pumping occurred on the first or second pulse.

us to measure over 100 ps of dynamics.

We used a global fitting algorithm to quantify the rate of spectral diffusion for  $\nu_3$  in each ionic liquid (Figure 9). The main peak in the 2D spectrum is sufficiently separated from the shoulders that it can be treated independently. Simulated spectra were calculated using a third-order response function formalism in the semi-impulsive limit. The frequency-fluctuation correlation function

$$c_2(t) = \frac{\delta(t)}{T_2} + \Delta^2 \exp\left(-\frac{t}{\tau_c}\right), \quad (2.11)$$

corresponds to a physical system in which CO<sub>2</sub> senses two distinct timescales of motion. Fast motions, in the homogeneous limit, are modeled by first term, which describes a loss of correlation that is too fast to be quantified. In the time domain lineshape function, this leading term describes exponential decay in the ensemble response from dephasing, (with time constant  $T_2$ ); the resulting lineshape in frequency space is a Lorentzian with FWHM of  $(\pi T_2)^{-1}$ .

The second term corresponds to processes in the spectral diffusion regime, which create a Kubo lineshape. In the slow modulation limit (where  $\tau_c \Delta \gg 1$ ), correlations do not change over the timescale of the molecular response. The resulting lineshape function describes a time-domain Gaussian with a variance of  $\Delta^{-2}$ ; the corresponding lineshape in frequency space is a Gaussian with FWHM of  $2.355\Delta$ .

The analytical lineshape function:

$$g(t) = \frac{t}{T_2} + \Delta^2 \tau_c^2 \left[ \exp\left(-\frac{t}{\tau_c}\right) + \frac{t}{\tau_c} - 1 \right], \quad (2.12)$$

can be used to calculate 2D spectra. Normalization of spectra before fitting removes any contribution from vibrational or orientational relaxation, which we do not model. A constrained nonlinear optimization algorithm globally fits the calculated spectra to experimental spectra by minimizing the sum of squares difference between the data and calculation. The algorithm optimizes  $T_2$ ,  $\tau_c$ , and  $\Delta$ , in addition to the central frequency, anharmonicity, the  $|00^01\rangle \rightarrow |00^02\rangle$  transition dipole moment, and the phase. The resulting spectral diffusion time,  $\tau_c$ , shows good agreement with that obtained by center line slope (CLS) analysis (section 2.7.1).



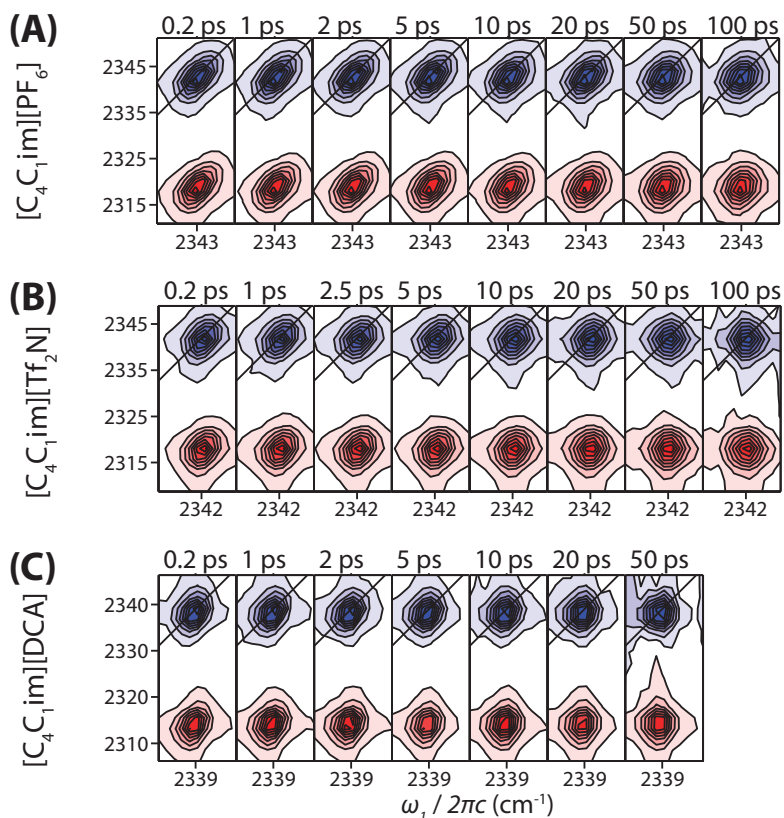


Figure 9: Experimental 2D-IR spectra of  $\text{CO}_2$  in  $[\text{Im}_{4,1}]$  (A)  $[\text{PF}_6]$ , (B)  $[\text{Tf}_2\text{N}]$ , and (C)  $[\text{DCA}]$  show the range of timescales for spectral diffusion in  $\nu_3$ . Spectral diffusion results from local structural relaxation around  $\text{CO}_2$ . Spectral modelling quantifies the timescale of this structural relaxation, and indicates that the timescale varies by up to an order of magnitude between ionic liquid solvents, and that  $\text{CO}_2$  dynamics are likely gated by anion dynamics in these ionic liquids. Spectra for  $\text{CO}_2$  in other ionic liquids tested can be found in figure 15.

Experimental and optimized calculated spectra agree in terms of overall lineshape and rate of spectral diffusion (Figure 10). The resulting lineshape parameters can then be used as input for spectral modelling that treats the shoulders and cross-peaks of the spectrum (Section 2.4.8).

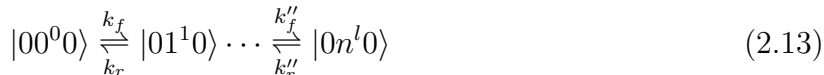
The lineshape parameters (Table 7) for  $\nu_3$ , combined with insights from computational modeling of CO<sub>2</sub>-anion-cation clusters, help to refine a physical picture of the solvation environment of CO<sub>2</sub> in ionic liquids. The timescale of frequency fluctuation correlations ( $\tau_c$ ) for CO<sub>2</sub> varies by up to an order of magnitude between the solvents, from  $13 \pm 3$  ps in [Im<sub>4,1</sub>][DCA] to  $104 \pm 10$  ps in [Im<sub>4,1</sub>][PF<sub>6</sub>]. The inhomogeneous width ( $\Delta$ ) which is largest in PF<sub>6</sub><sup>-</sup> and smallest in DCA<sup>-</sup> reflects the diversity of local environments reported by CO<sub>2</sub> in an ionic liquid. The dephasing time ( $T_2$ ), which varies from 2.6 to 3.4 ps, is longer than typical dephasing times in molecular solvents.

A quantitative analysis based on lineshape theory has allowed us to determine the dynamical timescales, dephasing times, and inhomogeneous linewidths for  $\nu_3$  in the six ionic liquids studied.

### 2.4.8 Modelling of Shoulders and Cross-Peaks

Having quantified the change in shape of the main 2D  $\nu_3$  band, we now turn to modelling of the dynamics encoded in the diagonal shoulders and cross-peaks.

Fluctuations in the CO<sub>2</sub>  $\nu_2$  population can be described by a thermal equilibrium between  $n$  bending modes:



The rate of transition between the ground and first excited state,  $k_f$ , combined with the Boltzmann distribution of states, determines the remaining rates.  $k_f$  is ultimately determined by the probability of stochastically gaining a quantum of energy in a CO<sub>2</sub> bending mode, most likely through collisions in the local environment. The first backwards rate,  $k_r$ , is directly analogous to the off-equilibrium vibrational energy relaxation rate.

A model that combines probabilistic fluctuations in bending mode population based on

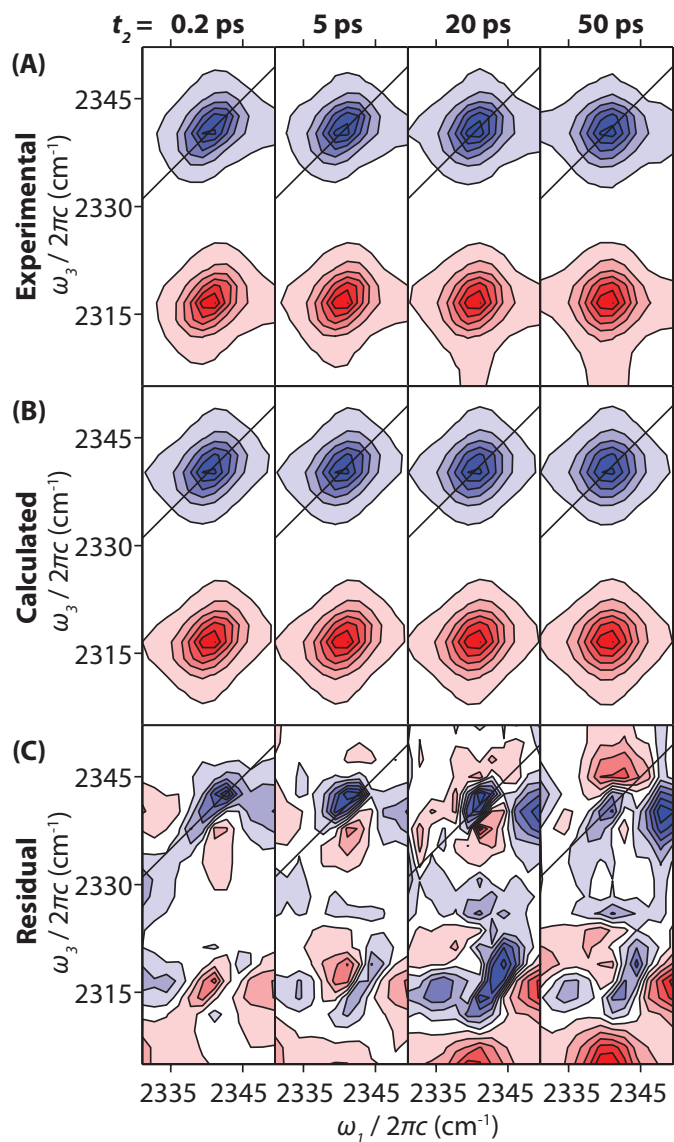


Figure 10: Example of global fitting of spectra. (A) Experimental 2D-IR spectra of CO<sub>2</sub> in [Im<sub>4,1</sub>][TFA]. (B) Calculated 2D-IR for  $\nu_3$  based on a third-order response function, which is fitted to A by optimizing the correlation time  $\tau_c$ , frequency range  $\Delta$ , and dephasing time  $T_2$ , in addition to several other parameters (see text). (C) Residual between the experimental and calculated spectra.

Table 7: Best fit correlation function parameters, by ionic liquid.  $\tau_c$  indicates the timescale of structural relaxation around CO<sub>2</sub>, the inhomogeneous linewidth  $\Delta$  reflects the range of frequencies experienced by CO<sub>2</sub> in different local environments of each ionic liquid, and the homogeneous dephasing time  $T_2$  arises from fast motions, such as librations, of CO<sub>2</sub>.

Anion	$\Delta$ (cm <sup>-1</sup> )	$\tau_c$ (ps)	$T_2$ (ps)
[PF <sub>6</sub> ] <sup>-</sup>	2.0 ± 0.1	104 ± 10	3.3 ± 0.1
[Tf <sub>2</sub> N] <sup>-</sup>	1.6 ± 0.1	26 ± 5	2.8 ± 0.1
[TfO] <sup>-</sup>	1.7 ± 0.1	25 ± 5	2.7 ± 0.1
[TFA] <sup>-</sup>	1.8 ± 0.1	40 ± 5	2.6 ± 0.1
[DCA] <sup>-</sup>	1.6 ± 0.1	13 ± 3	3.2 ± 0.2
[SCN] <sup>-</sup>	1.8 ± 0.1	16 ± 3	3.4 ± 0.1

Equation 2.13 with standard response function treatment is able to capture the essential physics required for such stochastic hot bands and their cross-peaks. The frequency of the  $\nu_3$  mode has two sources of variation: (1) the classical bath of intermolecular modes usually encountered in solvation dynamics and (2) the quantum bath of intramolecular vibrations to which the  $\nu_3$  band is coupled. The frequency of the  $\nu_3$  mode,  $\omega(t)$  fluctuates as

$$\omega(t) = \langle \omega \rangle + \delta\omega_{\text{inter}}(t) + \delta\omega_{\text{intra}}(t). \quad (2.14)$$

This stochastic frequency can be used as an input to standard nonlinear response function formalism. For example, the first order response function

$$R^{(1)}(t) \propto \exp(-i \langle \omega \rangle t) \times \left\langle \exp \left( -i \int_0^t dt' \delta\omega_{\text{inter}}(t') + \delta\omega_{\text{intra}}(t') \right) \right\rangle, \quad (2.15)$$

can be separated into the two parts, assuming the intra- and inter-molecular modes are uncorrelated,

$$R^{(1)}(t) \propto \exp(-i \langle \omega \rangle t) \times \left\langle \exp \left( -i \int_0^t dt' \delta\omega_{\text{inter}}(t') \right) \right\rangle \times \left\langle \exp \left( -i \int_0^t dt'' \delta\omega_{\text{intra}}(t'') \right) \right\rangle. \quad (2.16)$$

The intermolecular component can be treated with the cumulant expansion truncated at second order

$$R^{(1)}(t) \propto \exp(-i \langle \omega \rangle t - g(t)) \times \left\langle \exp \left( -i \int_0^t dt'' \delta\omega_{\text{intra}}(t'') \right) \right\rangle. \quad (2.17)$$

where  $g(t)$  is the lineshape function. The extension to third-order response functions is straightforward.

We performed a stochastic simulation in which an ensemble of trajectories was generated by allowing probabilistic instantaneous transitions between the ground and excited states of the  $\nu_2$ , with upward and downward rates consistent with the equilibrium populations. These transitions were allowed to happen at any point in the simulation, including during the coherence times.

The simulated lineshape is sensitive to the rate of thermal fluctuations in  $\nu_2$ . By tuning the rate constant  $k_f$  of stochastic fluctuation from the ground state into the first excited state and enforcing detailed balance to preserve equilibrium, we can control (1) whether or not diagonal shoulders will appear, and (2) the kinetics of cross-peak formation. In the limit of fast fluctuations, there are no clearly observed shoulders or cross-peaks, as all three peaks coalesce into a single band. In the limit of slow fluctuations, there are clearly defined shoulders, which persist throughout the experimental timescale, and cross-peaks do not grow into the spectrum. In an intermediate regime, we are able to reproduce both the lineshape and kinetics (Figures 11 and 7) seen in the relative intensities of the first shoulder and the cross-peaks between it and the main peak as functions of population time.

The microscopic rate constant ( $k_{\text{up}}$ ) for  $\text{CO}_2$  bending mode fluctuations in  $[\text{Im}_{4,1}][\text{TfO}]$ , from the Monte Carlo simulations is estimated to be  $k_{\text{up}} = 2.5 \pm 0.1 \text{ ns}^{-1}$ . The corresponding down rate, which is analogous to the vibrational relaxation rate in off-equilibrium pump-probe experiments, is estimated to be  $k_{\text{down}} = 30 \pm 12 \text{ ns}^{-1}$ . The kinetic Monte Carlo simulations include the non-equilibrium pumping that was explained by Fayer et al.<sup>119</sup> Their pump-probe measurements showed an increase in the ground state bleach band due to differences in transition dipole moment between the ground state  $|000\rangle$  and the ‘hot’ band  $|01^10\rangle$  transition dipole moments. The 2D measurement resolves the excitation frequency, so those dynamics are visible as a separate cross-peaks, where the pump-probe measurement

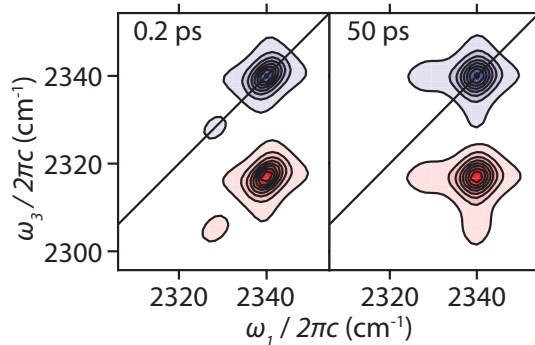


Figure 11: Calculated spectrum of CO<sub>2</sub> in [Im<sub>4,1</sub>][TFA] at  $t_2 = 0.2$  and 50 ps. Spectra are based on a stochastic simulation that allows the bending mode of a particular oscillator to fluctuate over the course of the experiment (while preserving a Boltzmann distribution). The lineshape as a function of  $t_2$  closely approximates the lineshape of the main peak (1a and 2a), first diagonal shoulder (1b and 2b), and cross-peaks (1b,a; 2a,b and 2b,a) in the experimental spectra (Figure 6).

observes the sum of the two peaks (the projection of the spectrum onto the  $\omega_3$ -axis). The main ground state bleach decays with bending mode exchange and the cross-peak grows. If the dipoles of the two states are the same these two terms cancel in the pump-probe measurement; however, the differences in dipole lead to effective rises in the pump-probe ground state bleach signal.

## 2.5 MOLECULAR INTERPRETATION

The resulting molecular picture is that the slower dynamics of the ionic liquid solvent *gate* CO<sub>2</sub>'s dynamics in solution. The slow timescale ( $\tau_c$ ) arises from structural relaxation of the solvent around CO<sub>2</sub>, and corresponds to the breakup of local ion shells. Until this liberating event, CO<sub>2</sub> is caged in a relatively well-ordered local environment. Librational and other fast motions only sample a narrow range of instantaneous frequencies, but the variation in

instantaneous frequency between different local environments gives rise to the inhomogeneous linewidth.

We assign the inhomogeneous width,  $\Delta$ , to the interactions of the  $\text{CO}_2$  with its local ion cage. Because charge transfer drives the distortion of the  $\text{CO}_2$  geometry, and the geometry determines the  $\nu_3$  frequency,  $\Delta$  reports the range of local structural motifs that the  $\text{CO}_2$  can sample. This range varies from  $2 \text{ cm}^{-1}$  in  $[\text{PF}_6]$  to  $1.6 \text{ cm}^{-1}$  in  $[\text{DCA}]$  (a 20% decrease).  $\Delta$  is also not strongly correlated to the average frequency shift or other structural parameters, and so we attribute it to the range of structures in the condensed phase.

The inhomogeneous linewidth of  $\text{CO}_2$  is the narrowest of the IR probes in recent 2D-IR experiments. Thiocyanate,  $\text{SCN}^-$ , has a total inhomogeneous linewidth  $\Delta_{\text{total}} \sim 8 \text{ cm}^{-1}$ ;<sup>77</sup> heavy water,  $\text{HOD}$ ,  $\Delta_{\text{total}} \sim 5 \text{ cm}^{-1}$ ;<sup>71</sup> and  $\text{CO}_2$   $\Delta_{\text{total}} \sim 2 \text{ cm}^{-1}$ . This trend reflects the strength of the coupling of the vibrational chromophore to its environment. The  $\text{SCN}^-$  anion is directly integrated into the ion network and hydrogen-bonded to the imidazolium cation through the 2-position.  $\text{HOD}$  interacts more weakly with the ionic liquid. It associates primarily with the anion, but it is sensitive to the electric field projected on the OH (or OD) bond axis. Because  $\text{HOD}$  is dipolar, it still experiences relatively large frequency fluctuations.  $\text{CO}_2$  is even more weakly still coupled to its environment.  $\text{CO}_2$ , which has a quadrupole moment and no dipole moment, is even less influenced by the local electric fields and is sensitive to the more chemical nature of the  $\text{CO}_2$ -anion-cation interaction.

Similarly, we assign the spectral diffusion time,  $\tau_c$ , to a local ion cage's lifetime around  $\text{CO}_2$ . The observed timescale reflects the time for the ion cage around  $\text{CO}_2$  to break up and permit  $\text{CO}_2$  to move to a novel local environment. This interpretation is consistent with previous computational work which indicate that the ionic liquid solvent reorients spontaneously to accommodate  $\text{CO}_2$  in well-defined locations in the ionic liquid,<sup>56</sup> and with NMR studies showing a well-defined angular distribution of  $\text{CO}_2$  around the cation.<sup>74</sup>

The bulk viscosity,  $\eta$ , serves as a proxy for this rate of diffusion which we can compare across the anions. Of course, small, neutral molecules like  $\text{CO}_2$  experience less friction from the solvent than the viscosity implies.<sup>120,121</sup> Nevertheless, the linear correlation of bulk viscosity and  $\tau_c$  ( $R^2 = 0.82$ ) supports our assignment (Figure 12). This correlation further suggests that the motion of the smaller, more mobile  $\text{CO}_2$  through the ionic liquid is gated

by the motion of the solvent ions.

CO<sub>2</sub> has a Lewis acid-base interaction with the anion of the ionic liquid. Isolated water has possible hydrogen bonds with both the anion and the cation. SCN<sup>-</sup> has a strong electrostatic attraction to the cation. Thus, all three chromophores have different specific interactions with their ionic liquid solvent. Nevertheless, the correlation times seen in 2D-IR studies of [SCN]<sup>-</sup> in [Im<sub>4,1</sub>][Tf<sub>2</sub>N]<sup>77</sup> and D<sub>2</sub>O/HOD in [Im<sub>4,1</sub>][PF<sub>6</sub>]<sup>71</sup> follow the same viscosity trend, when you account for the expected  $\sim 25\%$  decrease in the viscosity of [Im<sub>4,1</sub>][PF<sub>6</sub>] with  $\chi_{\text{H}_2\text{O}} \approx 0.04$ .<sup>122</sup> This fact suggests that each of these chromophores is reporting on the local diffusive motion of the surrounding solvent, despite the fact that SCN<sup>-</sup>, water, and CO<sub>2</sub> have different specific interactions with the ions in their solvation shells.

Furthermore, recent MD simulations suggest a direct relationship between ion cage lifetime and bulk transport properties such as self-diffusivity and conductivity.<sup>126</sup> While diffusivity of CO<sub>2</sub> and self-diffusivity of an ionic liquid are not identical parameters, the dynamics reported by CO<sub>2</sub> and [SCN]<sup>-77</sup> in [Im<sub>4,1</sub>][Tf<sub>2</sub>N] are nearly identical, and both have been related to ion cage lifetime. Thus, it is reasonable to speculate that CO<sub>2</sub> mass transport in these ionic liquids may also depend on ion cage lifetime. In this case, the correlation times reported could provide an avenue to directly address the molecular mechanism of CO<sub>2</sub> mass transport in ionic liquids, and might even give insight into other transport properties such as self-diffusivity and conductivity.

The homogeneous dephasing time,  $T_2$ , depends on both the timescale of fast motions for CO<sub>2</sub>,  $\tau_H$ , and on the frequency range experienced during those motions,  $\Delta_H$  ( $T_2 = (\Delta_H^2 \tau_H)^{-1}$ ). Experimentally, it is impossible to disentangle these two contributions, but the dephasing time ( $\sim 3$  ps) is significantly longer than that seen for either HOD ( $\sim 1$  ps) or SCN<sup>-</sup> ( $\sim 1.4$  ps) in ionic liquids. Thus, either CO<sub>2</sub> samples a relatively narrow frequency range during its homogeneous motions, or the those motions are particularly fast. Based on the small inhomogeneous linewidth,  $\Delta$ , and the fact that CO<sub>2</sub> is a small molecule whose moment of inertia (and consequently, the timescale for fast motions such as librations) is unchanged between ionic liquids and molecular solvents, it seems reasonable that the long dephasing time results from a narrow frequency range in a well-defined local environment, rather than a decrease in solvent interactions that slow the molecule.



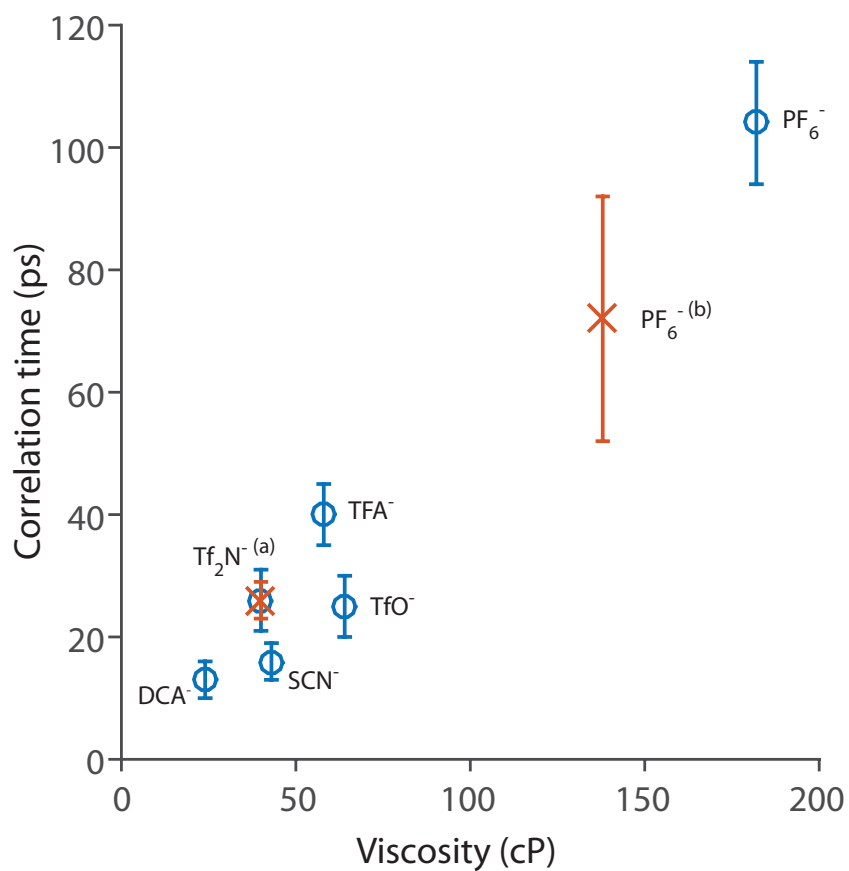


Figure 12: Correlation time,  $\tau_c$ , compared with viscosity for the 1-butyl-3-methylimidazolium ionic liquid solvent. Data points marked with an ‘O’ are from CO<sub>2</sub> in this work, while those marked ‘X’ are correlation times from other works (a) SCN<sup>-</sup> in [Im<sub>4,1</sub>][Tf<sub>2</sub>N],<sup>77</sup> (b) HOD in [Im<sub>4,1</sub>][PF<sub>6</sub>],<sup>71</sup> with ionic liquid viscosity scaled to account for water content.<sup>122</sup> Literature values for viscosity were used for all ionic liquids.<sup>123–125</sup>

Furthermore, the computational results show that, in gas phase clusters, CO<sub>2</sub> bends and adopts a distorted equilibrium geometry. In condensed phase dynamics, transitioning from a linear to a bent geometry could be one method of populating an excited state bending mode, and could have a direct impact on  $k_f$ , the rate of bending mode transitions. In principle, the spectral diffusion rate should influence the rate of transition between equilibrium geometries, and thus drive excitation and de-excitation of bending modes. Since we are able to model  $k_f$  based on equilibrium measurements, a systematic study of ionic liquids with different spectral diffusion rates could experimentally elucidate to how the motions of the solvent around a small molecule influence stochastic fluctuations in bending mode population.

These molecular mechanisms can be both tested and clarified by the comparison of these results to simulation. Frequency mapping techniques, combined with classical molecular dynamics simulations can be used to calculate the IR absorption spectrum and the spectral diffusion of modes of interest for small molecules such as water,<sup>127-129</sup> nitriles and thiocyanate,<sup>130,131</sup> and azides<sup>132,133</sup> in molecular solvents. More recently these methods have been expanded to explore isolated water in imidazolium ionic liquid solvents,<sup>134</sup> with good agreement with experiment.<sup>71</sup> Similar approaches for CO<sub>2</sub> could verify the molecular mechanism of CO<sub>2</sub> solvation in ionic liquids. It is likely, given the calculated dependence of  $\nu_3$  on CO<sub>2</sub> geometry, that any molecular dynamics simulation would need to account for the geometrical distortion of the CO<sub>2</sub> due to charge transfer, either directly via on-the-fly QM calculations, or with a classical proxy that indirectly accounts for these effects.

These initial studies utilized imidazolium-based ionic liquids because they are commercially available, are “archetypal,” and are relatively well-characterized; thus, they provide a good initial platform on which to develop spectroscopic methods. Many ionic liquids of interest, however, involve either novel classes of anions and cations or chemical modification of existing ionic liquids. The changes in solvent structure and dynamics that result from such modifications are generally *not* well-understood. This type of spectroscopy can provide valuable molecular insights into how and why chemical modification of ionic liquids determines bulk properties of interest, especially CO<sub>2</sub> uptake and local ion mobility (which is related to viscosity and conductivity). Chemisorbing ionic liquids are also of interest, especially for carbon capture applications, and many (1) have an initial physisorption step, and (2) have

an equilibrium between physisorbed and chemisorbed (reacted) CO<sub>2</sub>. These types of studies could help to elucidate the molecular mechanisms of the equilibrium between CO<sub>2</sub> in its free and bound forms.

## 2.6 CONCLUSIONS

We have demonstrated that the CO<sub>2</sub>  $\nu_3$  mode can act as a probe of local structure and dynamics in imidazolium ionic liquids. This method has potential application to the analysis of structure and dynamics in ionic liquids being developed for CO<sub>2</sub> capture.

The  $\nu_3$  frequency is sensitive to the timescale of local structural relaxation in ionic liquids. The timescale of this relaxation,  $\tau_c$  is determined by spectral modelling using a third-order response function formalism, with a Bloch-Kubo lineshape. For the imidazolium ionic liquids studied,  $\tau_c$  varies by as much as an order of magnitude between solvents, and correlates with the viscosity of the ionic liquids. The molecular mechanism posited for this timescale is the breakup of local ion cages around the CO<sub>2</sub>.

Computational studies aid understanding the origin of the  $\nu_3$  center frequency shifts in different imidazolium ionic liquids and suggest that geometrical distortion of the CO<sub>2</sub>, driven by charge transfer from the anion into virtual orbitals of CO<sub>2</sub> and from occupied orbitals of CO<sub>2</sub> into virtual orbitals of the cation. A simple one exciton Hamiltonian is able to reproduce the scaled harmonic frequencies with an RMS error of 1.2 cm<sup>-1</sup> by accounting for dependence of average frequency,  $\alpha$ , on total bond length,  $L$ , and the coupling constant,  $\beta$ , on magnitude of the angular distortion of CO<sub>2</sub>,  $\theta$ .

Anharmonic coupling of  $\nu_2$  and  $\nu_3$  allows thermal fluctuations of  $\nu_2$  population to stochastically shift the CO<sub>2</sub>  $\nu_3$  by units of  $-12$  cm<sup>-1</sup> (the coupling constant), and cause the appearance of a shoulder and dynamic cross-peaks on the 2D spectrum. Modelling of the stochastic bending mode population over the timescale of the experiment gives an estimate of the rate of excitation and de-excitation of bending mode population in the condensed phase at thermal equilibrium, estimated to be  $k_f = 2.5 \pm 1.0$  ns<sup>-1</sup> and  $k_r = 30 \pm 12$  ns<sup>-1</sup>. Additional spectral features (the “second shoulder” and apparent “cross-peak” 1c) arise from fifth-order signal.

The molecular picture that arises from this work is one in which imidazolium ionic liquids solvate CO<sub>2</sub> in well-defined local environments. The interactions of CO<sub>2</sub> with its solvent are dominated by local interactions with its nearest neighbor anion and cation. The picosecond dynamics of CO<sub>2</sub> are gated by the slower local diffusive motions of the anion and cation, whose translations and rotations are hindered due to electrostatic friction from surrounding ions, and potentially due to dispersive interactions between nanosegregated alkyl chains.

The methods and analysis developed in this work describe CO<sub>2</sub> in imidazolium ionic liquids. We expect that they will be transferable, however, to broad classes of materials, such as polymers or metal-organic frameworks, as well as to other ionic liquids.

## 2.7 SUPPORTING INFORMATION

### 2.7.1 Comparison of Global Fitting with Center Line Slope

For the center line slope method, we fit the signal size as a function of final frequency ( $\omega_3$ ) with two Gaussians with opposite signs for each initial frequency ( $\omega_1$ ) data point. The resolved positions of the Gaussians of the 0 to 1 transition peak are considered the center points. The center line slope is determined by fitting the center points linearly as a function of  $\omega_1$ . Estimated errors are propagated accordingly.

The resulting center line slope is fitted to a biexponential decay as a function of  $t_2$

$$c_2 = \sum_{i=1}^2 a_i \exp(-t_2/\tau_i) \tag{2.18}$$

with the resulting parameters for CO<sub>2</sub>  $\nu_3$  in [Im<sub>4,1</sub>][TFA]:  $a = 0.09 \pm 0.04$ ,  $\tau_1 = 2.1 \pm 2.1$ ps,  $b = 0.35 \pm 0.04$ ,  $\tau_2 = 35 \pm 5$ ps.

### 2.7.2 2D IR Spectra

### 2.7.3 Computational Results

In tables 8 and 9, ' $\nu_1$ ' is the frequency of the symmetric stretching mode of CO<sub>2</sub>, and ' $\nu_3$ ' is the frequency of the antisymmetric stretching mode of CO<sub>2</sub>.  $\alpha$  is the average of the two

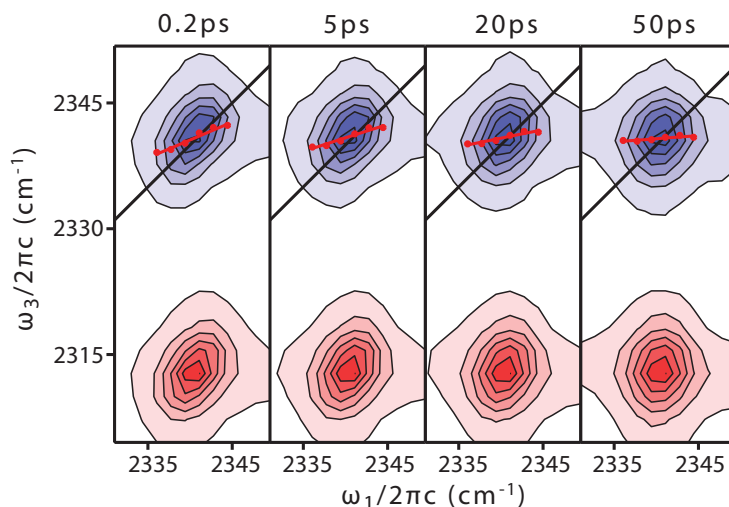


Figure 13: CLS overlaid on the 2D spectrum of CO<sub>2</sub> in [Im<sub>4,1</sub>][TFA]

normal mode frequencies,  $\alpha = (\omega_s + \omega_a)/2$ .  $\beta$  is the coupling constant, or the difference of the two local mode frequencies,  $\beta = (\omega_s - \omega_a)/2$ .

‘CT: CO<sub>2</sub> to IL’ is the amount of charge transferred from the CO<sub>2</sub> into the ionic liquid components, ‘CT: IL to CO<sub>2</sub>’ is the amount of charge transferred from the ionic liquid components into the CO<sub>2</sub>, and ‘CT: net’ is the net charge transferred into the CO<sub>2</sub>.

‘geom: angle’ is the CO<sub>2</sub> O–C–O angle, and ‘geom:  $\theta$ ’ is the deviation of the angle from 180°. ‘geom:  $l_1$ ’ and ‘geom:  $l_2$ ’ are the bond lengths of the two C–O bonds, ‘geom:  $l_2 - l_1$ ’ is the difference between the two bonds lengths, and ‘geom:  $L$ ’ is the sum of the two bond lengths. ‘geom: O<sub>12</sub>’ is the through-space oxygen-oxygen distance in CO<sub>2</sub>.

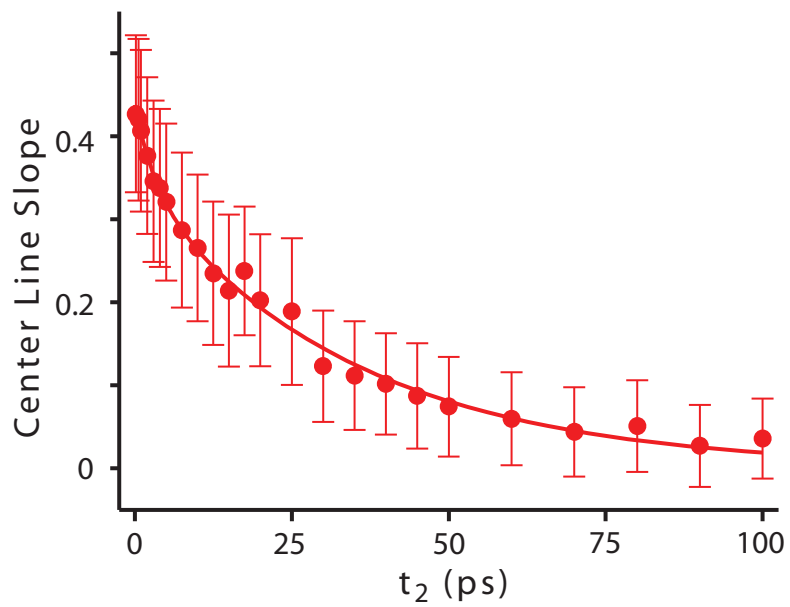


Figure 14: CLS fitting of  $\text{CO}_2 \nu_3$  2D spectrum from  $[\text{Im}_{4,1}][\text{TFA}]$

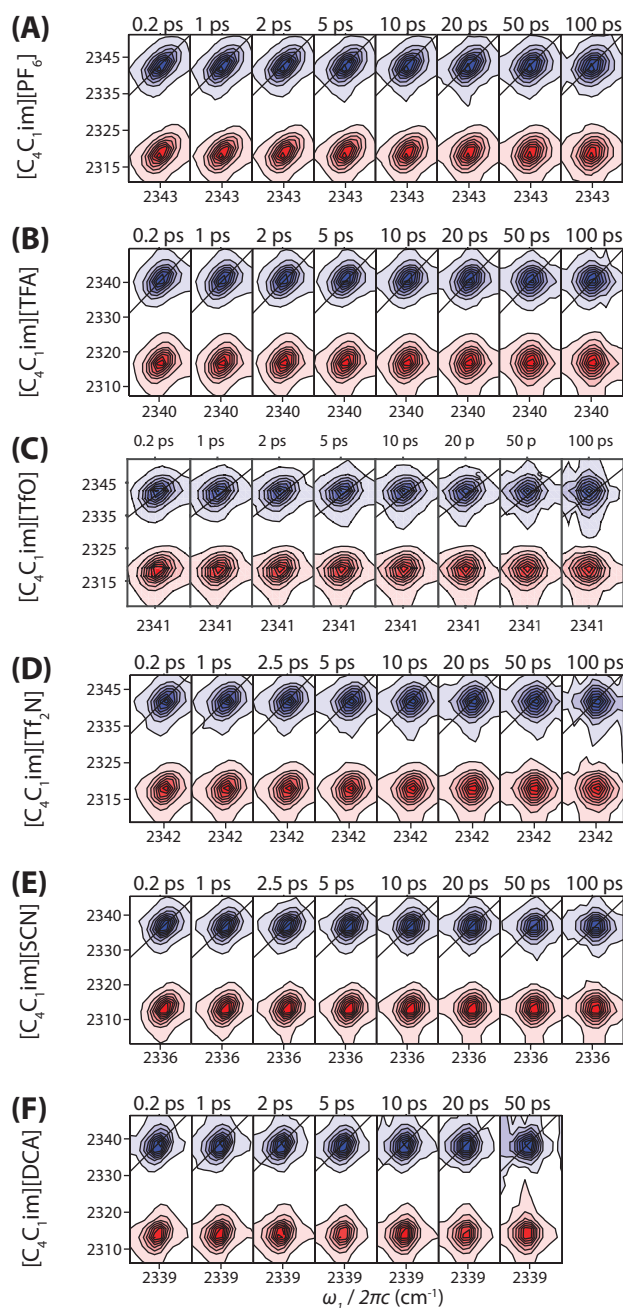


Figure 15: 2D-IR spectra of  $\text{CO}_2$   $\nu_3$  in  $[\text{Im}_{4,1}]^+$  (A)  $[\text{PF}_6]^-$ , (B)  $[\text{TFA}]^-$ , (C)  $[\text{TfO}]^-$ , (D)  $[\text{Tf}_2\text{N}]^-$ , (E)  $[\text{SCN}]^-$ , and (F)  $[\text{DCA}]^-$ , showing the range of timescales for spectral diffusion in  $\nu_3$ .

Table 8: CO<sub>2</sub>-IL geometries optimized allowing charge transfer. “Free CO<sub>2</sub>” is CO<sub>2</sub> not in the presence of the IL, “cation” is CO<sub>2</sub> only in the presence of the IL cation (no anion), and all other columns specify the IL anion identity. The cation is [Im<sub>1,1</sub>]<sup>+</sup> or [C<sub>1</sub>C<sub>1</sub>im]<sup>+</sup>.

	Free CO <sub>2</sub>	Cation	[BF <sub>4</sub> ] <sup>-</sup>	[DCA] <sup>-</sup>	[PF <sub>6</sub> ] <sup>-</sup>	[SCN] <sup>-</sup>	[TFA] <sup>-</sup>	[Tf <sub>2</sub> N] <sup>-</sup>	[TfO] <sup>-</sup>
$\nu_1$ (cm <sup>-1</sup> )	1371.93	1372.12	1372.96	1370.71	1373.95	1370.35	1371.68	1372.88	1371.14
$\nu_3$ (cm <sup>-1</sup> )	2436.1	2443.07	2434.71	2430.85	2437.48	2430.27	2429.81	2437.74	2433.89
$\beta$ (cm <sup>-1</sup> )	-532.085	-535.475	-530.875	-530.07	-531.765	-529.96	-529.065	-532.43	-531.375
$\alpha$ (cm <sup>-1</sup> )	1904.015	1907.595	1903.835	1900.7	1905.715	1900.31	1900.745	1905.31	1902.515
CT: CO <sub>2</sub> to IL (me <sup>-</sup> )	0	2.251	1.264	2.259	1.012	1.323	1.618	1.603	2.339
CT: IL to CO <sub>2</sub> (me <sup>-</sup> )	0	0.079	4.558	3.376	3.336	3.317	5.131	2.493	3.009
CT: net (me <sup>-</sup> )	0	-2.172	3.294	1.117	2.324	1.994	3.513	0.89	0.67
geom: angle (°)	179.96	179.93	175.49	175.01	176.21	175.65	174.78	177.36	175.99
geom: $\theta$ (°)	0.042	0.07	4.51	4.99	3.79	4.35	5.22	2.64	4.01
geom: $l_2 - l_1$ (Å)	$1.9 \times 10^{-7}$	0.015	0.0087	0.011	0.0071	0.0089	0.0076	0.0076	0.011
geom: $l_1$ (Å)	1.169	1.176	1.174	1.175	1.173	1.165	1.173	1.165	1.175
geom: $l_2$ (Å)	1.169	1.161	1.165	1.164	1.165	1.174	1.166	1.173	1.164
geom: O <sub>12</sub> (Å)	2.338	2.337	2.337	2.337	2.337	2.338	2.337	2.337	2.337
geom: $L$ (Å)	2.338	2.337	2.338	2.339	2.338	2.339	2.339	2.338	2.339



Table 9: CO<sub>2</sub>-IL geometries optimized without allowing charge transfer. “Free CO<sub>2</sub>” is CO<sub>2</sub> not in the presence of the IL, “cation” is CO<sub>2</sub> only in the presence of the IL cation (no anion), and all other columns specify the IL anion identity. The cation is [Im<sub>1,1</sub>]<sup>+</sup> or [C<sub>1</sub>C<sub>1</sub>im]<sup>+</sup>.

	Free CO <sub>2</sub>	Cation	[BF <sub>4</sub> ] <sup>-</sup>	[DCA] <sup>-</sup>	[PF <sub>6</sub> ] <sup>-</sup>	[SCN] <sup>-</sup>	[TFA] <sup>-</sup>	[Tf <sub>2</sub> N] <sup>-</sup>	[TfO] <sup>-</sup>
$\nu_1$ (cm <sup>-1</sup> )	1371.93	1371.47	1373.31	1373.44	1373.33	1373.14	1373.65	1373.06	1373.45
$\nu_3$ (cm <sup>-1</sup> )	2436.1	2439.86	2438.56	2439.1	2438.47	2438.78	2438.39	2438.767	2439.37
$\beta$ (cm <sup>-1</sup> )	-532.085	-534.195	-532.625	-532.83	-532.57	-532.82	-532.37	-532.85	-532.96
$\alpha$ (cm <sup>-1</sup> )	1904.015	1905.665	1905.935	1906.27	1905.9	1905.96	1906.027	1905.91	1906.41
CT: CO <sub>2</sub> to IL (me <sup>-</sup> )	0	1.743	1.038	1.836	1.154	1.302	1.023	1.246	1.941
CT: IL to CO <sub>2</sub> (me <sup>-</sup> )	0	0.039	3.053	2.266	2.221	1.64	2.916	1.274	2.048
CT: net (me <sup>-</sup> )	0	-1.704	2.015	0.43	1.067	0.338	1.893	0.028	0.107
geom: angle (°)	179.96	179.97	177.30	177.29	177.58	177.78	178.30	177.19	177.39
geom: $\theta$ (°)	0.04	0.03	2.70	2.70	2.42	2.22	1.70	2.81	2.61
geom: $l_2 - l_1$ (Å)	$1.9 \times 10^{-7}$	0.012	0.008	0.008	0.007	0.008	0.007	0.007	0.009
geom: $l_1$ (Å)	1.169	1.175	1.173	1.173	1.172	1.165	1.166	1.172	1.173
geom: $l_2$ (Å)	1.169	1.163	1.165	1.165	1.166	1.173	1.172	1.166	1.165
geom: O <sub>12</sub> (Å)	2.338	2.338	2.337	2.337	2.337	2.337	2.338	2.337	2.337
geom: $L$ (Å)	2.338	2.338	2.338	2.338	2.338	2.338	2.338	2.338	2.338

### 3.0 MODELING CARBON DIOXIDE VIBRATIONAL FREQUENCIES IN IONIC LIQUIDS: I. *AB INITIO* CALCULATIONS

The text in this chapter has been adapted from Berquist, E. J.; Daly, C. A.; Brinzer, T.; Bullard, K. K.; Campbell, Z. M.; Corcelli, S. A.; Garrett-Roe, S.; Lambrecht, D. S. Modeling Carbon Dioxide Vibrational Frequencies in Ionic Liquids: I. Ab Initio Calculations. *J. Phys. Chem. B* **2017**, *121*, 208–220, DOI: [10.1021/acs.jpcc.6b09489](https://doi.org/10.1021/acs.jpcc.6b09489), and is copyright the American Chemical Society. The author’s contribution to the work included performing all quantum chemical calculations and analyses (excepting those for DVR), designing the charge transfer mechanism and counterpoise correction analyses, writing those respective parts of the manuscript, and editing/revising the remainder.

#### 3.1 SUMMARY

This work elucidates the molecular binding mechanism of CO<sub>2</sub> in [C<sub>4</sub>C<sub>1</sub>im][PF<sub>6</sub>] ionic liquid (IL) and its interplay with the CO<sub>2</sub> asymmetric stretch frequency  $\nu_3$ , and establishes computational protocols for the reliable construction of spectroscopic maps for simulating ultrafast 2D-IR data of CO<sub>2</sub> solvated in ILs. While charge transfer drives the static frequency shift between *different* ionic liquids [27], we find here that electrostatic and Pauli repulsion effects dominate the dynamical frequency shift between different geometries sampled from the finite-temperature dynamics within a single ionic liquid. This finding is also surprising because dispersion interactions dominate the CO<sub>2</sub>–IL interaction energies, but are comparably constant across different geometries. An important practical consequence of this finding is that density functional theory is expected to be sufficiently accurate for constructing potential

energy surfaces for CO<sub>2</sub> in [C<sub>4</sub>C<sub>1</sub>im][PF<sub>6</sub>], as needed for accurate anharmonic calculations to construct a reliable spectroscopic map. Similarly, we established appropriate computational and chemical models for treating the extended solvent environment. We found that a QM/MM treatment including at least 2 cation-ion pairs at the QM level and at least 32 pairs at the MM level is necessary to converge vibrational frequencies to within 1 cm<sup>-1</sup>. Using these insights, this work identifies a computational protocol as well as a chemical model necessary to construct accurate spectroscopic maps from first principles.

### 3.2 INTRODUCTION

Capturing anthropogenic CO<sub>2</sub> before its release into the atmosphere is a pressing need, and most methods will require the development of novel materials, such as metal-organic frameworks, polymers, or ionic liquids (ILs).<sup>31–33,36,41,136</sup> Understanding how to control the interactions between CO<sub>2</sub> and its condensed-phase environment is a key to achieving efficient carbon capture and sequestration<sup>41,137–139</sup> and to developing routes toward potentially transforming CO<sub>2</sub> into value-added chemicals including fuels.<sup>140–142</sup> To rationally develop such technologies, a molecular level understanding of the CO<sub>2</sub>-sorbent interactions, structures and dynamics is necessary.

In a previous paper,<sup>27</sup> some of us established that CO<sub>2</sub>'s asymmetric vibrational stretch mode ( $\nu_3$ ) can be used to effectively probe the structure and dynamics of CO<sub>2</sub> dissolved in ILs. Using a combination of ultrafast two-dimensional infrared (2D-IR) spectroscopy and computational modeling, we determined structural candidates for the immediate CO<sub>2</sub> solvent environments whose vibrational solvation shifts were consistent between experiment and theory. Experimentally, the timescales of structural reorganization relative to the CO<sub>2</sub> molecule were correlated to the bulk viscosity. This experimental correlation between molecular property (solvation timescales) and bulk property (viscosity) suggests that CO<sub>2</sub> motions are gated by the same motions that lead to bulk diffusion. Molecular models are needed, however, to establish physical explanations for the correlation.

CO<sub>2</sub> solvation in ionic liquids gives rise to several interesting and challenging effects that

are worth exploring thoroughly, taking both computational methodology and the chemical picture into account. Our previous work suggested that charge transfer (CT) between CO<sub>2</sub> and the IL dominates the differentiation of the calculated vibrational signatures between different ionic liquids. Understanding the strength and nature of intermolecular interactions between CO<sub>2</sub> and its IL solvent fundamentally shapes our model of the CO<sub>2</sub>-IL interaction. Investigating CO<sub>2</sub>-IL interactions will thus deepen our understanding of the mechanism of CO<sub>2</sub> solubility in ionic liquids, and even of (de)activation of CO<sub>2</sub> for catalytic reductions. Interestingly, the asymmetric stretch frequency of CO<sub>2</sub> encodes the strength of intermolecular interactions as they are manifested in the molecular geometry of CO<sub>2</sub> interacting with the surrounding IL. In other words, one can determine the correct vibrational frequency shift for CO<sub>2</sub> from the distortion of the CO<sub>2</sub> geometry, but correctly determining the CO<sub>2</sub> geometry requires an understanding of intermolecular interactions between CO<sub>2</sub> and the IL solvent. These effects must be considered in the development of more reliable force fields that describe CO<sub>2</sub>-IL solvation, and of empirical structure-spectra maps used for comparing MD results with results from ultrafast spectroscopies.

This publication is first in a series aiming to unravel the physics underlying CO<sub>2</sub>-IL interactions as probed via vibrational spectroscopy and to develop a spectroscopic map to facilitate simulation of these spectra. *The central point of the present publication is to establish a more refined picture of intermolecular interactions and their correlations with vibrational shifts by improving both the computational approach and the chemical model as described below.* Some of the results from this publication inform the method choices for the subsequent paper in the series, where we develop and validate a spectroscopic map enabling one to reliably predict both the position and the width of the CO<sub>2</sub> asymmetric vibrational peak within classical MD simulations.<sup>143</sup>

In this manuscript, we address the critical challenges needed to generalize our previous work. The key issues are: treating the condensed phase environment, testing the dependence on the electronic structure theory method and basis set, and addressing anharmonic effects.

First, we extend our previous results by analyzing both the convergence and stability of the calculations with respect to the computational approach. Our absolutely localized molecular orbital (ALMO) calculations<sup>100,102,103</sup> evaluated the impact of CT and other chem-

ically intuitive components in the calculation of spectral signatures by using the modest B3LYP/6-31G(d,p) level of theory. Like most decomposition approaches, the ALMO results are expected to show some dependence on the underlying density functional approximation; in particular, it is known that the predicted amount of charge transfer depends on the amount of self-interaction error (SIE) present and the resulting (spurious) delocalization.<sup>104</sup> As a result, non-SIE corrected functionals are expected to overestimate the amount of CT, whereas the SIE-free Hartree–Fock approach can be used to estimate a lower bound for CT effects. Likewise, one can expect the CT contribution to depend on the basis set diffuseness, since the ALMO definition of CT is closely linked to the penetration of basis functions between different fragments. Thus, one goal of the present study is to quantify the method and basis set dependence of the calculated vibrational shifts and their effect on CT.

Although our previous density functional theory (DFT) calculations provided excellent agreement of predicted vibrational solvation shifts compared to experiment, these employed a minimalistic gas-phase cluster model consisting of only one CO<sub>2</sub> molecule together with a single cation/anion pair. Here we investigate the convergence of results with the size of the surrounding solvent shell both using a hybrid quantum mechanics molecular mechanics (QM/MM) approach with up to 195 atoms (6 molecular ion pairs) in the QM region and up to 8192 solvent atoms (256 molecular ion pairs) using classical point charges. In addition, we aim to assess the impact of solvent disorder by investigating different solvent geometries based on 85 statistically uncorrelated ( $R^2 = 0.0004$  for DVR-based vibrational frequency of snapshot  $N$  to vibrational frequency of snapshot  $N + 1$ ) snapshots sampled from a classical MD simulation (see paper II<sup>143</sup> for details). We use 1-butyl-3-methylimidazolium hexafluorophosphate ([C<sub>4</sub>C<sub>1</sub>im][PF<sub>6</sub>]) as a model IL in our calculations.

We will further analyze the impact of anharmonic effects, which we neglected in our previous (harmonic) calculations and which are necessary for a fair comparison to experiment. Here, anharmonic vibrational frequencies are obtained using the discrete variable representation (DVR) approach. The DVR method numerically solves the vibrational Schrödinger equation using a calculation of the CO<sub>2</sub> stretch potential energy surface, resulting in vibrational frequencies that include anharmonicity to all orders. An additional advantage of DVR is that it can be applied rigorously to systems at non-equilibrium geometries, which allows us

to include the disorder introduced by temperature for comparison with experiments carried out at temperatures above absolute zero.

The energy decomposition approaches we use add new and valuable physical insight into the origin of the vibrational frequency shifts. Many groups have used QM/MM approaches to estimate the condensed phase effects of a solvent on the transition energies of a chromophore, which provide a rich interpretation of the experiments.<sup>131,132,134,144,145</sup> Nevertheless, the interpretation for why the environment shifts the transition frequencies is difficult because many effects simultaneously determine the transition frequencies. The ALMO and symmetry-adapted perturbation theory (SAPT) energy decomposition approaches can separate the interactions of the vibrational chromophore with its environments into meaningful components allowing us to develop an empirical spectroscopic map and also understand its physical origins.

The paper is organized as follows. In section 3.3, we give a detailed account of the computational approaches used, including DFT calculations and DVR vibrational frequency calculations. In section 3.4, we analyze the dependence of the calculated CO<sub>2</sub> vibrational frequencies on the computational approach and the chemical model. For the computational approach, we quantify both the impacts of the electronic structure method (density functional and basis set) as well as the inclusion of anharmonic effects via the DVR approach. We analyze the dependence on the chemical model by comparing our previous gas-phase cluster calculations with results obtained for extended solvent boxes both at the classical point charge and the hybrid quantum mechanics/molecular mechanics (QM/MM) level, where structures are sampled from extensive molecular dynamics (MD) simulations (details of the MD simulations will be discussed in a follow-up paper<sup>143</sup>). This way, we aim to include the most important effects of solvent electrostatics, exchange-repulsion, and solvent disorder. In section 3.5 we present concluding remarks and an outlook for future research.

### 3.3 COMPUTATIONAL METHODS

#### 3.3.1 Methods and Basis Sets

Similar to our previous work,<sup>27</sup> we choose gas-phase clusters consisting of one  $[\text{C}_4\text{C}_1\text{im}][\text{PF}_6]$  ion pair with  $\text{CO}_2$  to quantify the effects of quantum chemical method and basis set on the quantum mechanically calculated harmonic frequencies. For methods, we employ the BLYP<sup>146,147</sup> generalized gradient approximation (GGA) and B3LYP<sup>94,95</sup> hybrid GGA density functionals, along with Hartree–Fock (HF) theory. This choice of methods allows us to test the dependence of the results on the percentage of exact (HF) exchange, as these methods have 0%, 20%, and 100% HF exchange, respectively. For basis sets, we choose 6-31G(d,p)<sup>96,97</sup> and 6-311++G(d,p)<sup>148–150</sup> to represent the commonly-used Pople-style basis sets (abbreviated as “small Pople” [SP] and “large Pople” [LP] in the following), along with Dunning’s correlation-consistent basis sets from double- to quadruple- $\zeta$  quality (cc-pVXZ, where X = D,T,Q, abbreviated as VXZ).<sup>151,152</sup> Initial geometries were constructed using Avogadro<sup>153,154</sup> before optimization to local minima with or without charge transfer allowed between molecules (*vide infra*), followed by harmonic frequency calculations.

To examine solvent effects, we investigate the converge of QM-calculated harmonic frequencies as a function of the solvent box size, where we vary both the number of ionic liquid pairs treated explicitly (“QM pairs”) and treated as point charges (“MM pairs”). A given combination is abbreviated as ( $n$  QM/ $m$  MM), where  $n$  and  $m$  are the number of QM and MM pairs, respectively. An ionic liquid pair, or briefly “ion pair”, is defined as one cation ( $[\text{C}_4\text{C}_1\text{im}]^+$ ) plus one anion ( $[\text{PF}_6]^-$ ). Ion pairs are included in the QM region based on the closest atom distance between individual cations and anions and the  $\text{CO}_2$ . The geometries for these calculations are taken from MD snapshots (see paper II<sup>143</sup> for details). Point charges for MM pairs are also extracted from the MD simulations.

To identify the effects of intermolecular interactions such as charge transfer (CT), we employ two types of calculations: 1. standard self-consistent field (SCF) calculations, and 2. “molecular interaction” calculations (SCF-MI) within the absolutely localized orbital (ALMO) framework.<sup>100,102,103</sup> ALMOs are constructed to utilize only atomic orbitals localized

to individual fragments. This approach is in contrast to canonical molecular orbitals, which may be significantly delocalized over different fragments. One can therefore employ the ALMO results to define intermolecular charge transfer contributions and in the following we will denote ALMO/SCF-MI results with “CT off” and conventional SCF results with “CT on”. We note that this definition of charge transfer is by no means unique, and it has been pointed out recently that constrained density functional theory (cDFT) predicts more reliable numbers for CT.<sup>155</sup> However, if applied consistently, we expect ALMO to provide qualitatively consistent trends across different systems, and in the present case it is beneficial to employ ALMO to allow comparison with our previous results. The decomposition results depend on the choice of interacting fragments. For ALMO-based calculations on the above gas-phase clusters, each individual molecule is chosen to be a separate fragment, whereas for all other calculations we chose two fragments — CO<sub>2</sub> as the first and all IL molecules as the second. These choices were made on the one hand to allow comparability with our previous results, and on the other hand to allow comparison with SAPT.

For further decomposition of interaction energies between fragments, in particular to identify the dispersion contribution ( $E_{\text{disp}}$ ), we use symmetry-adapted perturbation theory<sup>156</sup> (SAPT0<sup>157,158</sup>) as implemented in Psi4.<sup>159</sup> We employ the 6-31G(d,p) basis set to allow comparison to our DFT results within the same basis set, as well as jun-cc-pVTZ<sup>160</sup> for a more accurate comparison. Both primary basis sets use the jun-cc-pVTZ density fitting basis set during the SCF and SAPT iterations. CO<sub>2</sub> is treated as the first fragment, and two ionic liquid pairs are treated together as the second fragment, with no point charges included.

All other calculations employ a development version of the Q-Chem quantum chemistry program package.<sup>93</sup> Our DFT calculations use a numerical integration grid of (99,302) quality or higher throughout. Numerical tests suggest that vibrational frequencies are converged to within 0.2 cm<sup>-1</sup> with this grid. All ALMO calculations use the Gianinetti projector<sup>161</sup> to ensure suppression of charge transfer. Harmonic frequencies with CT turned off calculate the Hessian by numerical differentiation of analytical gradients to avoid solving the coupled perturbed self-consistent field equations within the ALMO formalism. Calculations in the VQZ basis set also employ numerical Hessians due to restrictions in the high-angular momentum



derivative code. The reported harmonic frequencies are unscaled.

In order to reduce the number of the costliest calculations (harmonic frequencies and SAPT0 energies), we use a sampling and weighting scheme as follows. From the 1000 statistically independent MD snapshots, a distribution of B3LYP/6-31G(d,p) harmonic frequencies is calculated on the 0 QM/0 MM substructures. The snapshots are placed into five bins centered around the mean harmonic  $\nu_3$  frequencies, each with a width corresponding to the standard deviation of the population of harmonic  $\nu_3$  frequencies. The weight for each bin is calculated by dividing the count for each bin by the total number of values in the histogram so the weights sum to 1. Five snapshots are chosen randomly from each of the five bins, giving the 25 snapshots used for calculating the dependence of harmonic frequencies on MD box size. From this subset, the first three snapshots are chosen for the interaction energy breakdown using ALMO-EDA and SAPT. The reported harmonic frequencies are unweighted unless explicitly stated.

### 3.3.2 Anharmonic Vibrational Frequency Calculations

We calculated anharmonic vibrational frequencies for the asymmetric stretch of CO<sub>2</sub> in each of 1000 statistically independent snapshots sampled from the dilute CO<sub>2</sub>/[C<sub>4</sub>C<sub>1</sub>im][PF<sub>6</sub>] MD simulations using an approach developed previously for CD<sub>2</sub> and PO<sub>2</sub> groups.<sup>162,163</sup> In this procedure, one numerically solves for the eigenvalues of the two-dimensional Schrödinger equation with the Hamiltonian,

$$H = \frac{p_1^2}{2\mu} + \frac{p_2^2}{2\mu} + \frac{p_1 p_2 \cos(\theta)}{m_C} + V(r_1, r_2), \quad (3.1)$$

using the DVR method.<sup>164,165</sup> In Eq. (3.1),  $r_1$  and  $r_2$  are the CO bond lengths,  $p_1$  and  $p_2$  are their conjugate momenta,  $\mu$  is the reduced mass of the CO bond,  $\theta$  is the OCO bond angle, and  $m_C$  is the mass of the carbon atom.

Our DVR analysis does not mix the stretching and bending vibrations. The anharmonic coupling between the stretches and bend is included classically; the quantum mechanically calculated stretch frequencies depend parametrically on the classical  $\theta$  coordinate. This is a reasonable approach to model infrared absorption experiments because the asymmetric

stretch is far from the overtones of the bending mode. Our approach would not be appropriate to simulate Raman spectra of CO<sub>2</sub>, where there is a Fermi resonance between the *symmetric* stretch and bend overtone. The separation of the stretch and bend could, in principle, be generalized by extending the dimensionality of the DVR potential energy surface to include the bend coordinates. Because the bend is doubly degenerate, the potential would become four-dimensional, however, and the computational cost to generate the potential would be infeasible.

The two-dimensional potential energy surface,  $V(r_1, r_2)$ , was obtained from density functional theory calculations performed as  $r_1$  and  $r_2$  were incremented from 0.955 to 1.45 Å in 0.045 Å steps, which corresponds to a  $12 \times 12$  grid (Figure 16). All production calculations were performed at the B3LYP/LP level of theory. The DVR calculation provides the vibrational energy levels,  $\{\varepsilon_n\}$ . The ground state has energy  $\varepsilon_0$ , the first excited state (the symmetric stretch) has energy  $\varepsilon_1$ , and the second excited state (the asymmetric stretch) has energy  $\varepsilon_2$ . The transition frequency for the asymmetric stretch frequency is then

$$\tilde{\nu}_{AS} = \frac{\varepsilon_2 - \varepsilon_0}{hc}. \quad (3.2)$$

For CO<sub>2</sub> isolated in the gas phase, the calculated anharmonic asymmetric stretch vibrational frequency was calculated to be 2383.7 cm<sup>-1</sup>, but is 2349.1 cm<sup>-1</sup> experimentally, giving a ratio of 0.9855 for calculated to experimental frequencies, which was used as a scaling factor to correct the vibrational frequencies for these calculations where noted below.

The DVR calculation also returns vibrational wave functions calculated on the same grid of points as the potential energy surface. This information was used to find the expectation value of the bond lengths with respect to the ground-state vibrational wave function,

$$\langle r_{1,2} \rangle = \sum_{i=1}^N \sum_{j=1}^N r_{1,2}^{ij} |\psi_{ij}|^2 \quad (3.3)$$

where  $r_1^{ij}$  or  $r_2^{ij}$  is the bond length at grid point  $(i, j)$ ,  $N = 12$  is the number of grid points along each coordinate, and  $\psi_{ij}$  is the value of the ground-state wave function at grid point  $(i, j)$ . Due to the anharmonicity of the potential energy surface, the expectation value of the bond lengths is longer than the bond length obtained from geometry optimizations.

For instance, the optimized gas-phase bond length is 1.1608 Å, and the gas-phase DVR expectation value is 1.1647 Å. In the IL environment the vibrationally averaged CO<sub>2</sub> bond lengths vary from snapshot to snapshot. On average, the vibrationally averaged CO<sub>2</sub> bond length is 1.1648 Å in the IL. The average bond length from the classical MD simulation snapshots is 1.1610 Å, which is almost identical to the harmonic equilibrium bond length in the force field (1.1600 Å).

### 3.4 SENSITIVITY OF THE CALCULATED VIBRATIONAL SIGNATURES TO THE UNDERLYING COMPUTATIONAL AND CHEMICAL MODEL

In a previous publication<sup>27</sup> we identified the predominant role of CT for the asymmetric stretch frequency of CO<sub>2</sub> in different ionic liquids. Here we aim to analyze whether our previous findings also hold when more sophisticated computational and chemical models are used. To this end, we investigate the impact of method, basis set, anharmonicity, electrostatics of the surrounding condensed phase, and solvent disorder on absolute and relative trends in the CO<sub>2</sub> asymmetric stretch frequency.

#### 3.4.1 Method and Basis Set Dependence

The simplest system for examining the quantum chemical method and basis set dependence of geometries is CO<sub>2</sub> in the gas phase. We first examine the sensitivity of the optimized geometry (Tab. 10). Increasing the fraction of HF exchange present in a given density functional leads to a decrease in bond lengths. This bond shortening can be attributed to the lack of dynamic correlation in HF, which leads to a tendency to underestimate bond lengths. An increase in basis set size also results in decreased bond lengths, with SP bond lengths slightly longer than those calculated using VDZ, and LP bond lengths similar to those calculated using VTZ.

Bond lengths for CO<sub>2</sub> combined with a single ion pair are shown in the lower half of

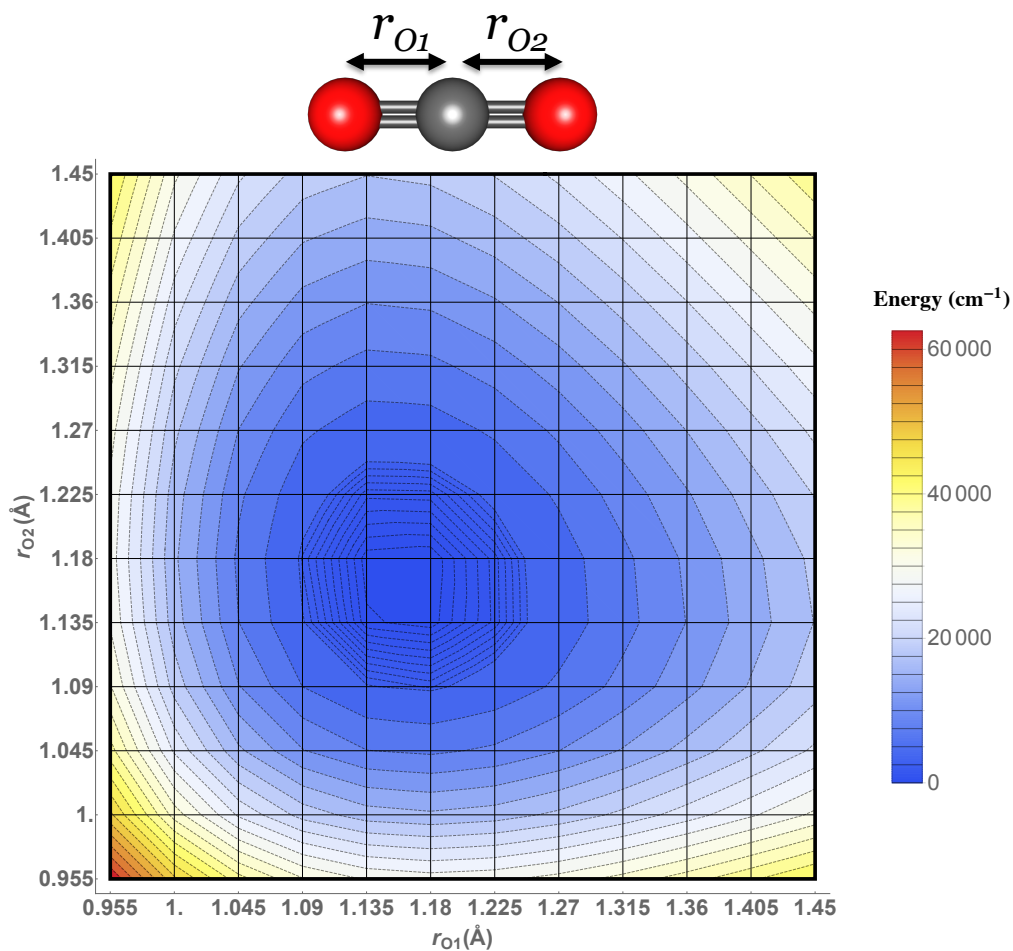


Figure 16: Contour plot of a discrete variable representation (DVR) of the Born–Oppenheimer potential energy surface of gas phase  $\text{CO}_2$ . Density functional theory single point energy calculations with the B3LYP functional and the LP basis set were performed on carbon dioxide, incrementing each CO bond’s length in steps of  $0.045 \text{ \AA}$ , from  $0.955$  to  $1.45 \text{ \AA}$ . Below a relative energy of  $2500 \text{ cm}^{-1}$ , contour lines are spaced by  $100 \text{ cm}^{-1}$ , above they are spaced by  $2500 \text{ cm}^{-1}$ . Mesh intersections indicate individual single point energy calculations. In order to calculate vibrational frequencies, this potential energy surface is incorporated into a discretized version of the Hamiltonian for the stretching modes of  $\text{CO}_2$  (Eq. 3.1), which is then numerically diagonalized. The asymmetric stretch frequency is obtained from the differences between the energy levels (Eq. 3.2).

Table 10: Dependence of  $r_{O1}$  and  $r_{O1} + r_{O2}$  bond lengths on functional (BLYP, B3LYP, and HF) and basis set (SP, LP, VDZ, VTZ, and VQZ) for a gas-phase cluster consisting of  $CO_2$  with a cation/anion pair. All values are in Å.

CO <sub>2</sub> (free)	$r_{O1}$			$r_{O1} + r_{O2}$		
	BLYP	B3LYP	HF	BLYP	B3LYP	HF
SP	1.1828	1.1692	1.1433	2.3656	2.3383	2.2865
LP	1.1744	1.1608	1.1357	2.3487	2.3216	2.2715
VDZ	1.1815	1.1674	1.1406	2.3630	2.3348	2.2811
VTZ	1.1736	1.1604	1.1362	2.3472	2.3208	2.2724
VQZ	1.1720	1.1588	1.1345	2.3441	2.3176	2.2690

CO <sub>2</sub> /[BMIM][PF <sub>6</sub> ]	$r_{O1}$			$r_{O1} + r_{O2}$		
	BLYP	B3LYP	HF	BLYP	B3LYP	HF
SP	1.1854	1.1723	1.1473	2.3649	2.3380	2.2871
LP	1.1775	1.1646	1.1402	2.3485	2.3217	2.2725
VDZ	1.1845	1.1710	1.1450	2.3622	2.3345	2.2819
VTZ	1.1766	1.1639	1.1402	2.3464	2.3204	2.2727
VQZ	1.1749	1.1622	1.1385	2.3434	2.3173	2.2694

Table 11: Dependence of  $\nu_3$  harmonic frequencies on functional (BLYP, B3LYP, and HF) and basis set (SP, LP, VDZ, VTZ, and VQZ) for an optimized gas-phase cluster consisting of CO<sub>2</sub> and one cation/anion pair. All frequencies reported in cm<sup>-1</sup>. Reported frequencies are unscaled.

Basis	Method		
	BLYP	B3LYP	HF
SP	2348.88	2437.23	2582.87
LP	2329.31	2419.80	2571.33
VDZ	2332.11	2422.71	2577.29
VTZ	2330.23	2417.28	2562.28
VQZ	2322.26	2409.34	2553.85

Table 10. Adding the ion pair leads to coordination of the CO<sub>2</sub> to both the cation and anion, leading to asymmetry in the CO<sub>2</sub> bond lengths. The effect is small, ranging from 0.006 Å at the BLYP/SP level to 0.008 Å at the HF/VDZ level. Trends in individual CO<sub>2</sub> bond lengths with varying method and basis set agree for both ion pair-coordinated CO<sub>2</sub> and free CO<sub>2</sub>, even with the asymmetry. Based on these results, we estimate that the CO bond lengths presented here are converged to within 0.01 Å regarding basis set effects.

Next we investigate the impact of method and basis set on the harmonic vibrational frequencies for the  $\nu_3$  mode in a CO<sub>2</sub>-IL complex (Tab. 11). As expected, the absolute values of harmonic frequencies depend significantly on the method, overall the frequency can vary by more than 200 cm<sup>-1</sup> depending on the choice of HF exchange percentage. We find a linear and positive correlation ( $R^2 = 0.96$ ) with the fraction of HF exchange present in the method for all basis sets (SI). This correlation is consistent with the tendency of HF theory to overbind. Increasing the basis set size, on the other hand, results in decreasing the harmonic CO<sub>2</sub>  $\nu_3$  frequency. The convergence of frequencies with basis size is rather slow, and even from VTZ to VQZ we still observe a change of 8 to 9 cm<sup>-1</sup>. Overall, the sensitivity

of the absolute vibrational frequencies to the method and basis set choice is large compared to the accuracy required to quantitatively describe the frequency shifts for CO<sub>2</sub> solvated in different ionic liquids, which is on the order of 10 cm<sup>-1</sup>.

It is therefore imperative to investigate how sensitive the prediction of relative trends is with respect to the computational approach. To this end, we consider snapshots from MD simulations (see Ref. [143] for details), which allows us to test how well different computational approaches can predict *trends* in dependence of the local coordination environment around the CO<sub>2</sub> and the bulk solvent structure. Fig. 21 shows the CO<sub>2</sub>  $\nu_3$  harmonic frequencies calculated for 1000 statistically uncorrelated MD snapshots (0 QM/256 MM) using various SCF-type approaches, as compared to Møller–Plesset perturbation theory to second order (MP2) as the least expensive wave function-based method that incorporates dispersion effects.<sup>166</sup> The predicted harmonic frequencies in Fig. 21 are parallel to each other for most of the frequency range, independent of method and basis set choice. These results for relative trends in vibrational frequencies are highly encouraging. Aside from a multiplicative scaling factor, any of the common quantum chemical methods investigated here can qualitatively reproduce the distribution of harmonic frequencies.

Identifying the role of different intermolecular interactions in determining the vibrational signature of solvated CO<sub>2</sub> is an important aspect of our previous and ongoing work.<sup>27,104,167</sup> In our previous publication,<sup>27</sup> we found that the CT contribution is decisive for discriminating between the vibrational signatures of CO<sub>2</sub> solvated in different ionic liquids. We therefore end this section by investigating the method and basis set dependence of the CT contributions to relative shifts in the CO<sub>2</sub> asymmetric stretch frequency due to solvation. As discussed in our previous publication,<sup>27</sup> CT can enter the frequency shift at two stages: (i) during the geometry optimization (i.e., influencing the geometries sampled by the solvated CO<sub>2</sub>), and (ii) during the frequency calculation (i.e., by modifying the curvature of the potential energy surface at the point where the frequency is calculated).

To quantify the sensitivity of both mechanisms to the computational approach, we investigated the CT contributions to the calculated frequencies for both mechanisms (Table 12). To assess the “geometry mechanism” (i), we calculated the frequency shift between geometries optimized using standard “CT on” and ALMO/SCF-MI “CT off” calculations, respectively.

During the frequency calculation, we used the default “CT on” potential energy surface.

Our results show that the frequency shift varies by up to  $0.9\text{ cm}^{-1}$  depending on the method and by up to  $1.53\text{ cm}^{-1}$  depending on the basis set (Tab. 12, top). For the “curvature mechanism” (ii), we calculated the shift between standard “CT on” and “CT off” frequency calculations at the same, conventionally (“CT on”) optimized geometries (Tab. 12, bottom). Here we find variations with the method of up to  $0.49\text{ cm}^{-1}$  and up to  $1.86\text{ cm}^{-1}$  with the basis set. Compared to the magnitude of the total CT frequency shifts, the basis set dependence is not negligible.

This finding is not surprising, because the definition of CT used within the ALMO approach is intimately linked to the locality of the basis set. However, we note that *all* methods and basis sets tested here provide qualitatively similar predictions, namely a negative frequency shift between  $-1.19$  to  $-3.50\text{ cm}^{-1}$  for the “geometry mechanism” and between  $-0.94$  and  $-2.97\text{ cm}^{-1}$  for the “curvature mechanism”. For future work, it will be useful to consider alternative definitions of CT that are less dependent on basis set locality (see e.g. Ref. [155]).

The relative magnitudes of the “geometry” versus “curvature” mechanism results warrant some discussion, given that we found previously that the geometry contribution dominates the differentiation between different ionic liquids. According to the present results, the impact of CT via the geometry is typically bigger than the curvature effect by  $\sim 15$  to  $35\%$  (with exception of the LP results, where the geometry effect is smaller than the curvature effect), but on the other hand this means the curvature effect still makes up a significant portion of the CT frequency shift. This finding may be surprising at first, but it is important to note that our previous discussion was not about the absolute magnitude of the shifts, but about the differentiation between different ionic liquids. In fact, the current results of  $-3.50$  and  $-2.86\text{ cm}^{-1}$  are in good agreement with those from our previous publication,  $-3.29\text{ cm}^{-1}$  and  $-2.29\text{ cm}^{-1}$ , respectively. While the absolute values of the CT shifts are largely caused by the impact of CT on the force constants during the frequency calculation, the modification of the geometry determines the *differentiation* between different ILs.



Table 12: Dependence of CT contributions to the  $\nu_3$  frequency based on functional (BLYP, B3LYP, and HF) and basis set (SP, LP, VDZ, VTZ, and VQZ). Calculations are on an optimized gas-phase cluster of CO<sub>2</sub> with a single cation/anion pair. All frequencies reported in cm<sup>-1</sup>. We distinguish two mechanisms by which CT can enter the frequency: (i) a “geometry mechanism” where CT determines the optimized geometry of the cluster, calculated as the difference between standard harmonic frequency results at the optimized standard (“CT on”) and ALMO (“CT off”) geometries. (ii) a “curvature mechanism” where CT enters by modifying the force constant at the optimized geometry, calculated as the difference between the “CT on” and “CT off” frequencies at the optimized standard (“CT on”) geometry.

	(i) “Geometry Mechanism”			(ii) “Curvature Mechanism”		
	method			method		
Basis	BLYP	B3LYP	HF	BLYP	B3LYP	HF
SP	-3.16	-3.50	-3.33	-2.72	-2.86	-2.80
LP	-2.80	-3.38	-3.72	-1.31	-1.25	-1.55
VDZ	-2.46	-3.11	-2.93	-2.92	-2.97	-2.53
VTZ	-1.90	-2.35	-1.19	-1.43	-1.33	-0.94
VQZ	-2.44	N/A	N/A	-1.18	N/A	N/A

### 3.4.2 Anharmonicity

We employed a grid-based anharmonic (DVR) method to assess the effects of anharmonicity on the CO<sub>2</sub>  $\nu_3$  frequency. That is, we numerically solved for the vibrational wave function on the discretized potential energy surface spanned by two degrees of freedom for the CO<sub>2</sub> molecule (along the OCO axis), constraining the degrees of freedom involved in CO<sub>2</sub> bending modes, as discussed earlier. Solving for the fully anharmonic vibrational wave function is important for connecting to experiment for the following reasons. First, an accurate comparison to experiment is only possible via anharmonic calculations because experiment probes vibrational transitions that take place on the full (anharmonic) potential energy surface. Second, harmonic frequencies change their physical interpretation as vibrational energy levels when calculated away from the minimum where they pick up contributions from non-zero forces and from higher-order derivatives. For these reasons, DVR calculations are essential for being able to calculate vibrational frequencies for different geometries (MD snapshots) that are consistent with experimental conditions sampling dynamical structures away from the equilibrium geometries. Consequently, the DVR approach is instrumental for the construction of a spectroscopic map as presented in Ref. [143], as it requires sampling various (non-equilibrium) geometries from MD simulations. In this section, we investigate the impact of anharmonicity on the frequency and geometry (i.e., position expectation values) of gas-phase CO<sub>2</sub>.

Our goal is to separate the impact of anharmonicity from the numerical errors arising from using a discretized and reduced-dimensional (2D) potential energy surface. To this end, we performed standard harmonic calculations (using analytical derivatives on the fully-dimensional surface), harmonic calculations on the discretized reduced-dimensional grid, and DVR calculations on the same grid (Tab. 13). All potential energy surfaces were calculated at the B3LYP/LP level. The standard (analytical) harmonic calculation yielded a vibrational frequency of 2420 cm<sup>-1</sup> and an equilibrium bond length of 1.161 Å.

For the harmonic grid-based calculations, we fit an accurate analytical potential to the grid-based potential and calculated analytical second (harmonic) derivatives. The analytical potential used an expansion in a local mode basis and included quadratic, cubic, and quartic

Table 13: Comparison of analytical harmonic (AH), grid-based harmonic (GH), and grid-based anharmonic (DVR) absolute  $\nu_3$  frequencies for CO<sub>2</sub> in the gas phase (B3LYP/LP potential energy surface). Reported frequencies are unscaled.

	$\nu_3$ (cm <sup>-1</sup> )	$r_{\text{OC}}$ (Å)
analytical harmonic (AH)	2420	1.161
grid-based harmonic (GH)	2414	1.159
grid-based anharmonic (DVR)	2384	1.165

diagonal terms, but not quartic terms in the off-diagonal (see the [Supporting Information](#)). This approach resulted in an excellent agreement with the discrete surface ( $R^2 = 0.9997$ ) and yielded a harmonic frequency of 2414 cm<sup>-1</sup> and an equilibrium CO bond length of 1.159 Å. This data suggests that the errors due to discretization and reduction of dimensionality are on the order of 6 cm<sup>-1</sup> and 0.002 Å, respectively. These results support that the grid-based approach largely captures the correct physics. In comparison, the effects of anharmonicity are much larger than the discretization errors: DVR predicts a vibrational frequency of 2384 cm<sup>-1</sup> (a 36 cm<sup>-1</sup> red shift) and an equilibrium bond length of 1.165 Å (a 0.004 Å bond lengthening). The red shift due to anharmonicity is significantly larger than the solvation shifts observed in our previous publication.<sup>27</sup>

As another point supporting the necessity for DVR calculations, we observe that the standard deviation of the calculated harmonic vibrational frequency distribution of  $\sim 100$  cm<sup>-1</sup> (Tab. 17) is about an order of magnitude larger than the experimental distribution of  $\sim 6$  to 10 cm<sup>-1</sup>. This large error arises from the fact that harmonic calculations pick up contaminations in the frequency that grow with the displacement from the minima. As we will show in the follow-up paper,<sup>143</sup> the anharmonic DVR calculations result in frequency distributions that are consistent with experiment. In summary, we conclude that the inclusion of anharmonic effects is imperative for constructing a spectroscopic map that allows meaningful comparison to experiment.

Table 14: Summary of DVR data averaged from 85 MD snapshots (SP = small Pople = 6-31G(d,p), LP = large Pople = 6-311++G(d,p)) at the QM/MM level, treating CO<sub>2</sub> plus 2 ion pairs quantum mechanically. Reported frequencies are unscaled.

PES	$\nu_3$ (cm <sup>-1</sup> )	$r_{O1}$ (Å)	$r_{O2}$ (Å)	$r_{O1} + r_{O2}$ (Å)
B3LYP/SP, CT off	2399.4	1.1724	1.1729	2.3454
B3LYP/SP, CT on	2389.5	1.1728	1.1734	2.3462
B3LYP/LP, CT on	2373.9	1.1645	1.1650	2.3295

Since our previous results emphasized that CT is a decisive factor for the CO<sub>2</sub> solvation shift between different ionic liquids, it is important to investigate the significance of CT in the context of anharmonic calculations. The fact that DVR calculations are agnostic toward the level of theory used to generate the discretized potential energy surface allows us to perform DVR calculations on standard (“CT on”) and ALMO/SCF-MI (“CT off”) surfaces (Tab. 14). Excluding CT, we obtain a frequency of 2399.4 cm<sup>-1</sup> and bond lengths of 1.172 Å and 1.173 Å, respectively. The slight bond length asymmetry arises from the fact that we averaged over 85 snapshots taken from classical MD simulations, resulting in slightly non-symmetric solvation environments (see Ref. [143] for details). Upon turning CT on, the vibrational frequency decreases to 2389.5 cm<sup>-1</sup>, whereas bond lengths remain almost unchanged at 1.173 and 1.173 Å, respectively. This 9.9 cm<sup>-1</sup> red-shift is comparable in size to experimentally observed solvation shifts for CO<sub>2</sub> in ionic liquids,<sup>27,168</sup> which means that the anharmonic results are consistent with our earlier conclusion that CT is important to understand the vibrational signature of solvated CO<sub>2</sub>. We plan to quantify the *relative* effects of CT when studying ILs other than [C<sub>4</sub>C<sub>1</sub>im][PF<sub>6</sub>] in future work.

### 3.4.3 Molecular Mechanism of CO<sub>2</sub>-IL Interactions

We also studied the interplay between intermolecular interaction energies and the CO<sub>2</sub>  $\nu_3$  frequency in order to (1) further elucidate the molecular mechanism governing CO<sub>2</sub> solvation

in ILs and to (2) inform the selection of computational approaches used to construct the CO<sub>2</sub>-IL spectroscopic map.<sup>143</sup>

**3.4.3.1 Role of Charge Transfer and Basis Set Superposition Error** We previously concluded<sup>27</sup> that the CT term, as defined in the ALMO approach, plays an important role in determining the relative frequency shifts of  $\nu_3$  in different ILs, primarily by modifying the CO<sub>2</sub>-IL cluster optimized geometry rather than by modifying the frequency calculated at a given geometry. To quantitatively analyze this conclusion, we compared predicted frequencies calculated with and without the ALMO approach (i.e., with CT on and off, respectively), at geometries again calculated with ALMO turned off and on, respectively (Fig. 17). At the SCF geometry (i.e., geometry optimized with CT on, Fig. 17a), the predicted frequencies for both ALMO (CT off) and standard DFT (CT on) are linearly correlated ( $R^2 = 0.978$ ). At the ALMO geometry (i.e., geometry optimized with CT off, Fig. 17b), however, neither standard DFT nor ALMO (CT off) frequency calculations correlate with the standard frequency/geometry result ( $R^2 = 0.116$  and  $0.011$  respectively).

These findings quantitatively support our earlier conclusions that (a) CT during the frequency calculation is not significant for differentiating  $\nu_3$  experiencing different IL solvation environments and that (b) CT is crucial for determining the correct geometry, because removing CT during the geometry optimization eliminates this frequency differentiation.

However, ALMO simultaneously corrects for CT as well as the basis set superposition error (BSSE). We therefore performed counterpoise (CP) corrected geometry optimizations and frequency calculations to isolate artificial BSSE from physical CT effects (Fig. 18).<sup>169,170</sup> The frequencies calculated at CP-corrected geometries and CP-*uncorrected* geometries correlate well ( $R^2 = 0.888$ , Fig. 18a). This finding suggests that the impact of BSSE on the geometry is small, at least as far as it is probed by the vibrational frequency. Consequently, we infer that indeed CT causes the geometry change leading to differentiation of  $\nu_3$  between different IL solvation environments (and not a BSSE artifact).

The CP-corrected frequency calculations show a somewhat smaller correlation with the CP-*uncorrected* frequencies ( $R^2 = 0.629$ , Fig. 18b). We believe that, for the given combination of systems and computational methodologies, the CP-*uncorrected* frequencies are

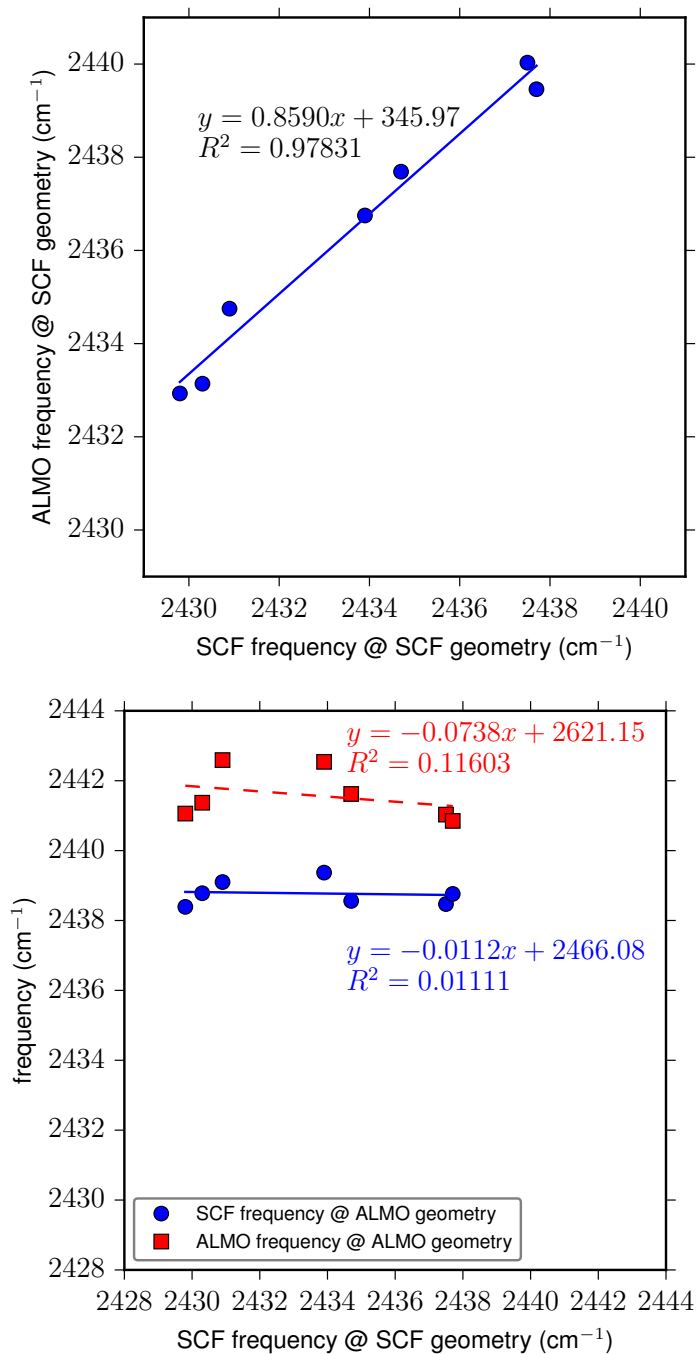


Figure 17: Comparison of B3LYP/SP frequencies calculated (a, top) within the ALMO (CT off) approach versus standard SCF (CT on) at the standard SCF-optimized geometry and (b, bottom) at the ALMO-optimized geometry versus standard SCF-optimized geometry with Hessian within the ALMO approach (red squares) versus Hessian using standard DFT (blue triangles). The frequencies are unscaled.

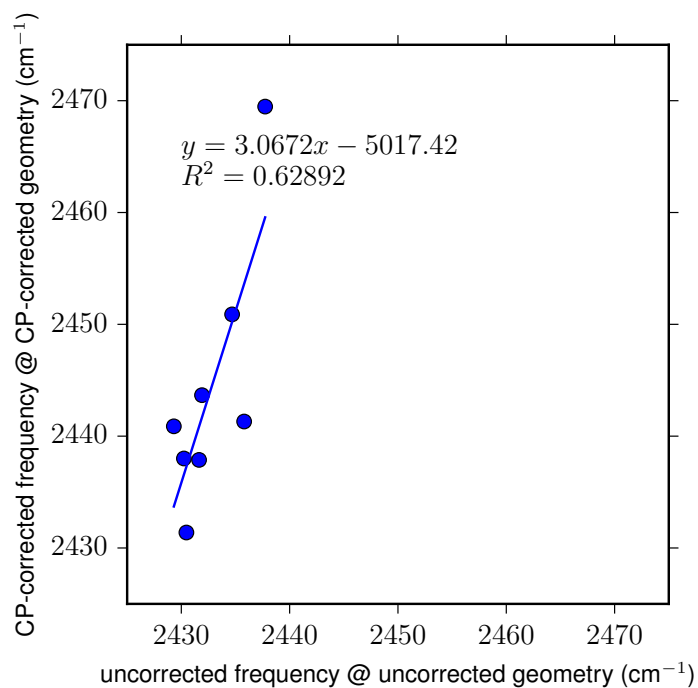
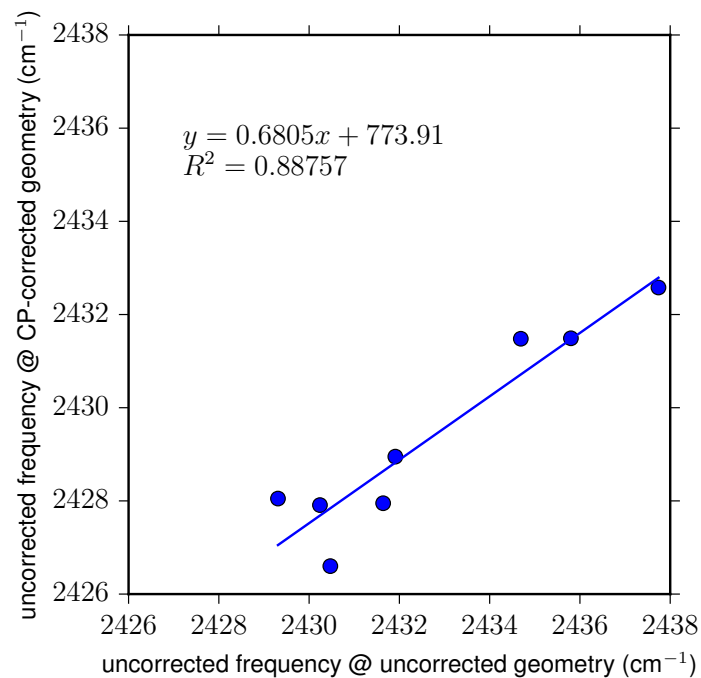


Figure 18: Comparison of B3LYP/SP frequencies (a, top) calculated at counterpoise (CP) corrected geometries versus uncorrected geometries and (b, bottom) calculated using CP-corrected Hessians at CP-corrected geometries versus uncorrected Hessians at uncorrected geometries. The frequencies are unscaled.

actually more accurate than the CP-corrected frequencies for a number of reasons: (1) The trends in CP-corrected frequencies contradict our ALMO-frequency results, where ALMO predicts a *negligible* impact of CT+BSSE in the frequency calculation step, whereas the CP-corrected frequency calculations suggest a significant impact of BSSE on the relative frequencies. DFT integration grid superposition errors can be excluded here because we did not find significant dependence on the grid size. (2) We suspect that the CP-correction overestimates the impact of BSSE for the given systems and computational approaches, because a detailed analysis of intermolecular interaction energies shows that, with the CP correction applied, one obtains unphysical, repulsive CT energies (see the SI): ALMO-CT should always be non-repulsive, because a system would not undergo CT if it were energetically unfavorable. However, after applying a CP correction to the ALMO delocalization energy, one obtains repulsive CT energies for several MD snapshots. We conclude that the CP correction overestimates the amount of BSSE for the given system. Similar conclusions have been drawn before, for example arguing<sup>171</sup> that the CP overcorrects for BSSE, or that BSSE is actually beneficial because it reduces basis set incompleteness.<sup>172</sup>

These fundamental problems in the CP-corrected frequencies manifest themselves in two findings (Table 21): (1) overestimation of the solvatochromic shifts for  $\nu_3$  in different ILs, and (2) incorrect ordering of the CP-corrected frequencies compared with experiment. The CP-corrected frequencies overestimate the range of solvatochromic shifts for  $\nu_3$  by a factor of  $5.9\times$  when compared with experiment ( $38.1\text{ cm}^{-1}$  versus  $6.5\text{ cm}^{-1}$ ), whereas the CP-uncorrected shifts agree much more closely ( $8.4\text{ cm}^{-1}$ ). Furthermore, the ordering of the CP-corrected frequencies is incorrect. For example, the lowest experimental  $\nu_3$  frequency is found in  $[\text{C}_4\text{C}_1\text{im}][\text{SCN}]$  ( $2336\text{ cm}^{-1}$ ), while the CP-corrected frequencies predict the lowest result in  $[\text{C}_4\text{C}_1\text{im}][\text{DCA}]$ . The CP-uncorrected frequencies correctly predict  $[\text{SCN}]^-$  as causing the largest shift. Furthermore, anions that cause identical experimental frequencies ( $[\text{Tf}_2\text{N}]^-$  and  $[\text{BF}_4]^-$ :  $2341.7\text{ cm}^{-1}$ ) differ by  $9.6\text{ cm}^{-1}$  in the CP-corrected prediction (and only by  $1.1\text{ cm}^{-1}$  in the uncorrected calculation).



### 3.4.4 Energetics of CO<sub>2</sub>-IL Interactions and Dependence on Decomposition Approach

To understand how the choice of decomposition method affects the predicted solute-solvent interactions, we decomposed the CO<sub>2</sub>-IL interaction energies of 15 representative MD snapshots using both ALMO and symmetry adapted perturbation theory (SAPT) calculations (Tab. 15). Because of computational cost, the SAPT calculations were performed within the uncorrelated monomer approximation (SAPT0). While individual terms in the ALMO and SAPT0 energies do not have direct correspondence, we combined contributions into roughly comparable terms describing (a) frozen-fragment interactions (electrostatics plus Pauli repulsion,  $E_{\text{frz}}$ ), (b) polarization plus Pauli repulsion ( $E_{\text{pol}}$ ), (c) charge transfer plus Pauli repulsion ( $E_{\text{CT}}$ ), (d) dispersion plus Pauli repulsion ( $E_{\text{disp}}$ ), and (e) the total interaction energy ( $E_{\text{int}}$ ).

Our results show that  $E_{\text{frz}}$  between CO<sub>2</sub> and the IL is, on average, repulsive. ALMO predicts repulsion between +0.08 and 0.50 kcal mol<sup>-1</sup>, whereas SAPT0 predicts somewhat larger repulsion between 0.59 and 1.47 kcal mol<sup>-1</sup>. The attractive interaction of -2.45 kcal mol<sup>-1</sup> at the SAPT0/SP MCBS level is likely due to an underestimation of inter-monomer overlap effects, which would result in a smaller contribution from Pauli repulsion that is recovered at the SAPT0/SP DCBS level. The finding of an overall repulsive  $E_{\text{frz}}$  is not surprising given that CO<sub>2</sub> is a neutral, non-polar molecule, which should have little electrostatic attraction to the surrounding ions when polarization or charge delocalization are excluded.

The polarization energies contribute, as expected, to the binding. ALMO predicts polarization energies between -1.24 and -1.30 kcal mol<sup>-1</sup>, whereas SAPT0 predicts somewhat more attractive energies between -1.91 and -2.47 kcal mol<sup>-1</sup>.

The largest qualitative differences between predictions from ALMO and SAPT0 are found for charge transfer and dispersion energies. The ALMO-HF CT (-0.10 kcal mol<sup>-1</sup>) is comparable in magnitude to the SAPT0 CT (-0.28 kcal mol<sup>-1</sup>), whereas the ALMO-DFT CT is significantly larger (-1.18 kcal mol<sup>-1</sup>). Overestimation of CT effects is expected for DFT-based ALMO because of self-interaction error, leading to spurious delocalization of electrons. The dispersion component does not exist in the SCF-based ALMO approach used here and

Table 15: Comparison of ALMO- versus SAPT0-decomposed interaction energies averaged over 15 representative molecular dynamics snapshots. ALMO calculations were performed within the SP basis to allow comparison to results from Ref. [27]. For SAPT0, we report both monomer-centered basis set (MCBS) and dimer-centered basis set (DCBS) results. Energies are reported in kcal mol<sup>-1</sup>.

component	ALMO-DFT/SP	ALMO-HF/SP	SAPT0/SP		SAPT0/jun-cc-pVTZ	
			MCBS	DCBS	MCBS	DCBS
$E_{\text{el}}^{(10)}$	—	—	-5.295	-7.018	-6.097	-6.154
$E_{\text{exch}}^{(10)}$	—	—	2.845	7.611	7.106	7.622
$\mathbf{E}_{\text{frz}}$	<b>0.500</b>	<b>0.082</b>	<b>-2.450</b>	<b>0.593</b>	<b>1.009</b>	<b>1.468</b>
$E_{\text{ind}}^{(20)}$	—	—	-1.218	-3.636	-2.593	-3.916
$E_{\text{ind-exch}}^{(20)}$	—	—	0.009	2.147	0.921	2.041
$\delta_{\text{HF}}$	—	—	-0.839	-0.421	-0.801	-0.507
$\mathbf{E}_{\text{pol}}$	<b>-1.241</b>	<b>-1.303</b>	<b>-2.049</b>	<b>-1.910</b>	<b>-2.473</b>	<b>-2.382</b>
$\mathbf{E}_{\text{CT}}$	<b>-1.177</b>	<b>-0.102</b>	<b>-0.280</b>		<b>-0.203</b>	
$E_{\text{disp}}^{(20)}$	—	—	-4.193	-5.197	-7.323	-8.084
$E_{\text{disp-exch}}^{(20)}$	—	—	0.058	0.483	0.393	0.671
$\mathbf{E}_{\text{disp}}$	—	—	<b>-4.135</b>	<b>-4.714</b>	<b>-6.930</b>	<b>-7.413</b>
$\mathbf{E}_{\text{int}}$	<b>-1.918</b>	<b>-1.323</b>	<b>-8.634</b>	<b>-6.031</b>	<b>-8.394</b>	<b>-8.326</b>

can therefore not be compared.

We note, importantly, that SAPT0 predicts the dispersion contribution of  $-4.14$  to  $-7.41$  kcal mol $^{-1}$  to be the major binding contribution for the CO $_2$ -IL interactions. In contrast, ALMO-DFT predicts polarization and CT to be the dominant and equally important contributions to the overall binding, and ALMO-HF predicts polarization to be the largest contribution. Overall, the ALMO total binding energies range from  $-1.32$  to  $-1.92$  kcal mol $^{-1}$ , whereas SAPT0 binding energies within the MCBS and DCBS are  $-6.03$  and  $-8.63$  kcal mol $^{-1}$ , respectively. Given the magnitude of the dispersion interaction, it is not surprising that SAPT0 predicts a much more attractive total binding energy.

The three approaches used here disagree on the dominating mechanism of binding: ALMO-HF favors polarization, ALMO-DFT polarization in combination with CT, and SAPT0 predicts dispersion as the most important binding component. Interestingly, though, individual energy contributions that exist both in ALMO and SAPT0 agree qualitatively. We expect that the SAPT0/jun-cc-pVTZ DCBS numbers are the most accurate of the results reported here because these results are free of SIE (unlike ALMO-DFT) and explicitly account for inter-monomer correlation (dispersion) effects.

The conclusion changes when comparing *relative trends* in binding energies between different MD snapshots. Comparing the correlation coefficients for linear regression analysis between individual energy components and the total energies (Tab. 16a), *all* of the methods predict that the frozen monomer electrostatics plus Pauli repulsion component ( $E_{\text{frz}}$ ) is the single most important contribution to trends in the total energy ( $R^2 = 0.870 - -0.944$ ). This finding implies that, although ALMO and SAPT0 disagree on the identity of the largest absolute energy component, they agree that the  $E_{\text{frz}}$  is most important to capture differential effects between different MD snapshots. This finding is somewhat surprising, given that  $E_{\text{frz}}$  is a rather small energy contribution on an absolute scale, whereas dispersion makes up the largest contribution to the total binding energy. However, this trend shows that dispersion varies less than the frozen-fragment interaction across different points on the potential energy surface.

These findings have at least two important implications. Since ALMO-DFT, ALMO-HF and SAPT0 qualitatively agreed on the  $E_{\text{frz}}$  component, all of these methods should correctly

Table 16: Correlation coefficients ( $R^2$ ) between different solute-solvent interaction energy components as calculated at the ALMO-DFT (B3LYP), ALMO-HF, and SAPT0 level within the SP basis. (a) Comparison between individual components and the total interaction energy ( $E_{\text{tot}}$ ) calculated for each methodology. (b) Correlation coefficients between contributions to the CO<sub>2</sub>-IL interaction energy calculated via ALMO as compared to the corresponding SAPT0 energy terms. Reported SAPT0 results are within the dimer-centered basis set (DCBS), except for the charge transfer component, which is calculated as the difference between monomer- and dimer-centered basis set results for the induction energy.

(a) Component	ALMO-DFT	ALMO-HF	SAPT0
$E_{\text{frz}}$	0.887	0.944	0.870 <sup>1</sup>
$E_{\text{pol}}$	0.170	0.087	0.073
$E_{\text{CT}}$	0.225	0.270	0.093
$E_{\text{disp}}$	—	—	0.150

<sup>1</sup>  $E_{\text{el}}$ : 0.602,  $E_{\text{exch}}$ : 0.073

(b) Component (ALMO — SAPT0)	ALMO-DFT	ALMO-HF
$E_{\text{frz}} - E_{\text{el}} + E_{\text{exch}}$	0.989	0.993
$E_{\text{pol}} - E_{\text{ind}} + E_{\text{ind-exch}} + \delta_{\text{HF}}$	0.887	0.927
$E_{\text{CT}} - E_{\text{CT}}$	0.720	0.383
$E_{\text{tot}} - E_{\text{tot}}$	0.913	0.927

describe relative trends in the energetics across different MD snapshots. That is, despite self-interaction error and the lack of dispersion interactions, DFT is expected to correctly capture energetic trends across different snapshots along the potential energy surface. Furthermore, any spectroscopic map (*i.e.* a map between CO<sub>2</sub> geometry and vibrational frequency) *must* contain an electrostatic term to capture these leading-order relative effects. ALMO and SAPT approaches predict the CT and dispersion terms to be the second most important for relative frequency trends, respectively ( $R^2 = 0.225 - -0.270$  for ALMO-CT and  $R^2 = 0.150$  for SAPT-dispersion), which are expected to show a different geometry dependence. Consequently, adding a Lennard-Jones type potential to the spectroscopic map may allow implicit collection of these non-electrostatic terms.

The good correlation between ALMO and SAPT0 electrostatic plus Pauli repulsion terms is also demonstrated by a term-by-term comparison (Tab. 16b): ALMO-DFT has a correlation coefficient of 0.989 with SAPT0, and ALMO-HF has one of 0.993. Polarization energies also correlate well ( $R^2 = 0.887$  and  $0.927$ , respectively). The lowest agreement is found for the charge transfer term ( $R^2 = 0.720$  and  $0.383$ , respectively). This low degree of correlation is somewhat surprising, since the formal definitions of charge transfer are very similar between ALMO and SAPT0 — both involve the difference between energetics calculated for a monomer-centered and a dimer-centered basis set. Furthermore, the absolute magnitude of the ALMO-HF CT term agrees better with SAPT0, but it shows less correlation across different MD snapshots. The ALMO-DFT CT term, on the other hand, is much larger than in SAPT0, but seems to correlate better with trends across MD snapshots, although one would expect the ALMO-DFT results to be less reliable due to self-interaction error. This large difference between ALMO-DFT and SAPT0 CT energies remains even with the use of range-separated density functionals, but we note that adding a dispersion correction leads to a good agreement between DFT-D and SAPT0 (Table 19).

Interestingly, the ALMO total binding energies correlate well with SAPT0 binding energies ( $R^2 = 0.913$  and  $0.927$ , respectively), even though ALMO binding energies are generally much smaller than SAPT0 energies. This good correlation can be understood by considering that the major discrepancy between ALMO and SAPT0 binding energies comes from the dispersion term, which, despite its large magnitude, is relatively constant across MD snapshots

and therefore does not correlate strongly with the total binding energy.

In summary, despite shortcomings in capturing the nuanced physics of intermolecular CO<sub>2</sub>-IL binding, DFT- and even HF-based approaches are expected to produce qualitatively correct trends, if the goal is to compare binding energies across different MD snapshots. Consequently, we expect that DFT will capture qualitative trends in the potential energy landscape, and can be safely used to generate DVR potential energy surfaces for DVR used to construct and validate a spectroscopic map for CO<sub>2</sub> in ionic liquids.

### 3.4.5 Electrostatic Interactions with the Extended Solvent Phase

Our previous publication<sup>27</sup> used a rather crude gas-phase cluster model for the solvated CO<sub>2</sub>. Here we investigate the effect of adding a more extended solvation environment by varying the MM region (Fig. 19). Calculations were performed at the B3LYP/SP level of theory, with the ionic liquid molecules represented as point charges. Each data point is an average from 25 snapshots, which were selected using the sampling scheme previously described.

As expected, the largest frequency shifts occur when the first few solvent ion pairs are added to the solvated CO<sub>2</sub>. The frequency is converged to within 1 cm<sup>-1</sup> at a solvent layer size of ~32 ion pairs, which corresponds to a solvent droplet of radius ~14 Å. This relatively fast convergence (from the perspective that the solvent molecules are charged) is attributed to electrostatic screening due to the solvent.

We observe qualitatively the same trends and convergence patterns independent of whether a pure point charge embedding is used or whether one or two ion pairs are included at the QM level. This finding indicates that solvation boxes including ~32 or more ion pairs are sufficient to converge electrostatic effects on the vibrational frequency independent of how many solvent molecules are treated at the QM level. However, we notice that there is a significant shift in frequencies depending on the size of the QM region. The fact that increasing the solvent box (treated as classical point charges) does not cause the differently sized QM regions to reconcile indicates that the underlying effects are quantum mechanical (as opposed to purely electrostatic) in nature. It is therefore indicated to further investigate the dependence of the frequency on the size of the QM region.

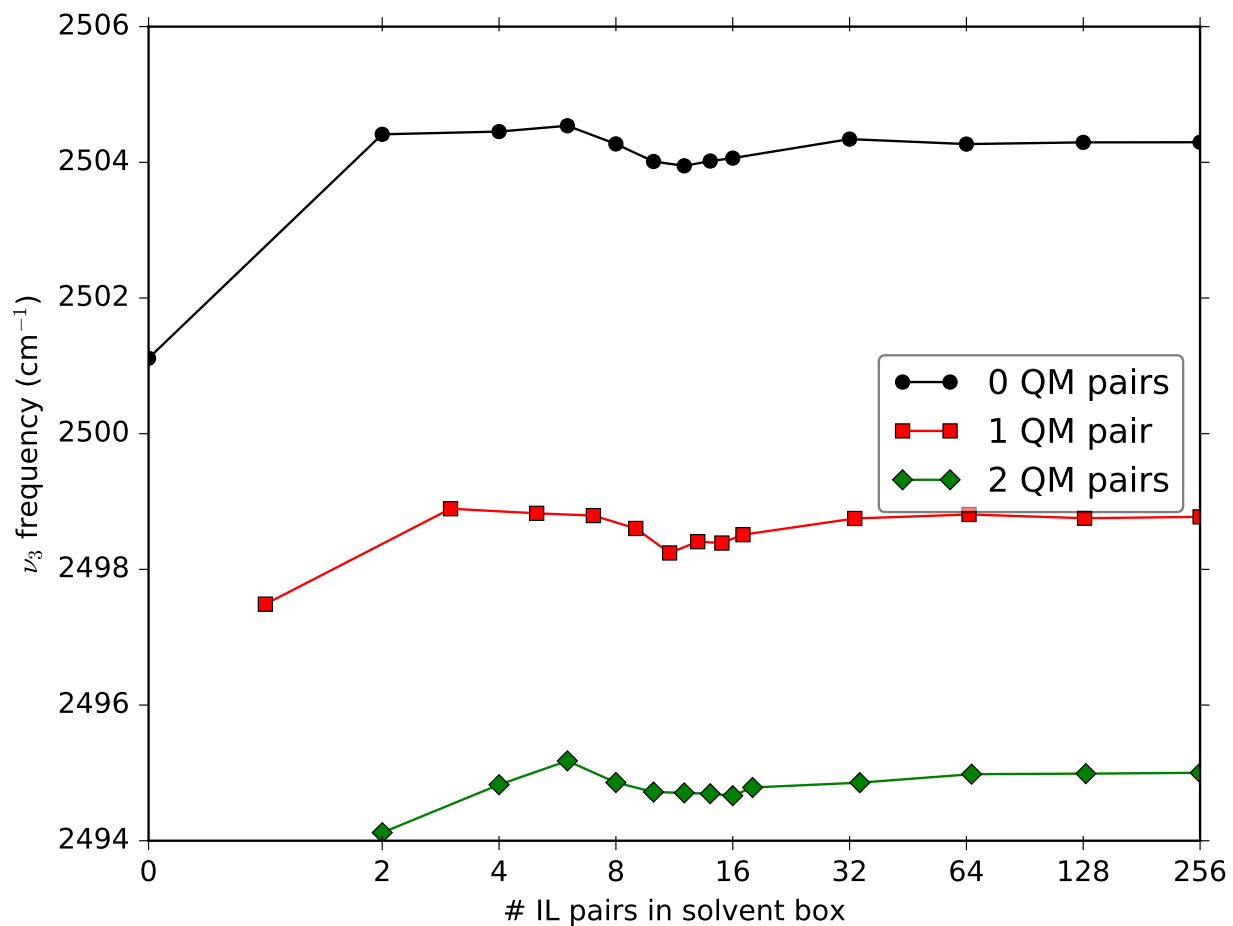


Figure 19: Dependence of harmonic  $\nu_3$  frequencies (unscaled) on the solvent box size (B3LYP/SP potential energy surface).

### 3.4.6 Quantum Mechanical Interactions with the Solvation Shell

To test for convergence of the asymmetric frequency with respect to the size of the QM region, we selected a subset of 10 snapshots from the MD trajectories and carried out QM/MM harmonic frequency calculations with increasing numbers of anion–cation pairs treated quantum mechanically (Fig. 20). These snapshots were carefully selected to sample the entire range of local vibrational frequencies encountered in the MD trajectories and were averaged by weighting with the appropriate probabilities for each frequency bin. We therefore expect the following results to be representative of large parts of the entire potential energy surface sampled. We tested several functions to fit the mean frequency  $\bar{\nu}$  (i.e., the weighted average over all sampled frequencies) versus the QM region size, and found that an exponential decay function yields the best overall fit,

$$\bar{\nu}(n) = a \exp(-kn) + c$$

where  $n$  is the size of the QM region (measured in number of ion pairs) and  $a$ ,  $k$ , and  $c$  are constants. The quality of the resulting fit was excellent ( $R^2 = 0.99$ ), which allowed us to extrapolate the average frequency to infinite QM region size with high confidence. This extrapolated frequency then allows us to assess the accuracy (convergence) of QM/MM calculations with differently sized QM regions.

We determine the converged frequency to be  $c = 2388.8 \pm 1.5 \text{ cm}^{-1}$  (Fig. 20). This data indicates that the 6 QM calculations are nearly numerically converged ( $\bar{\nu}(6) - \bar{\nu}(\infty) = 0.3 \text{ cm}^{-1}$ ). To test the qualitative accuracy of the calculations using smaller QM regions, we examined the correlation between frequencies calculated with  $n = 0 - 5$  QM pairs and 6 QM pairs for each snapshot (for details, see Supporting Information). We find that the  $n = 0 - 1$  results do not correlate well with the  $n = 6$  benchmark ( $R^2 = 0.41$  and  $0.66$ , respectively). The  $n = 2$  calculations yield an acceptable correlation coefficient of  $R^2 = 0.82$  at a  $\sim 9$ -fold reduced computational cost compared to  $n = 6$  ( $\sim 0.8$  versus  $\sim 7$  CPU hours for a single point calculation). The  $n = 3 - 4$  calculations show even better correlations with  $R^2 = 0.87 - -0.96$ , respectively, but the computational cost increase to  $\sim 3$  CPU hours does not appear justified. In summary, the  $n = 2$  calculations reproduce trends in the  $\text{CO}_2$



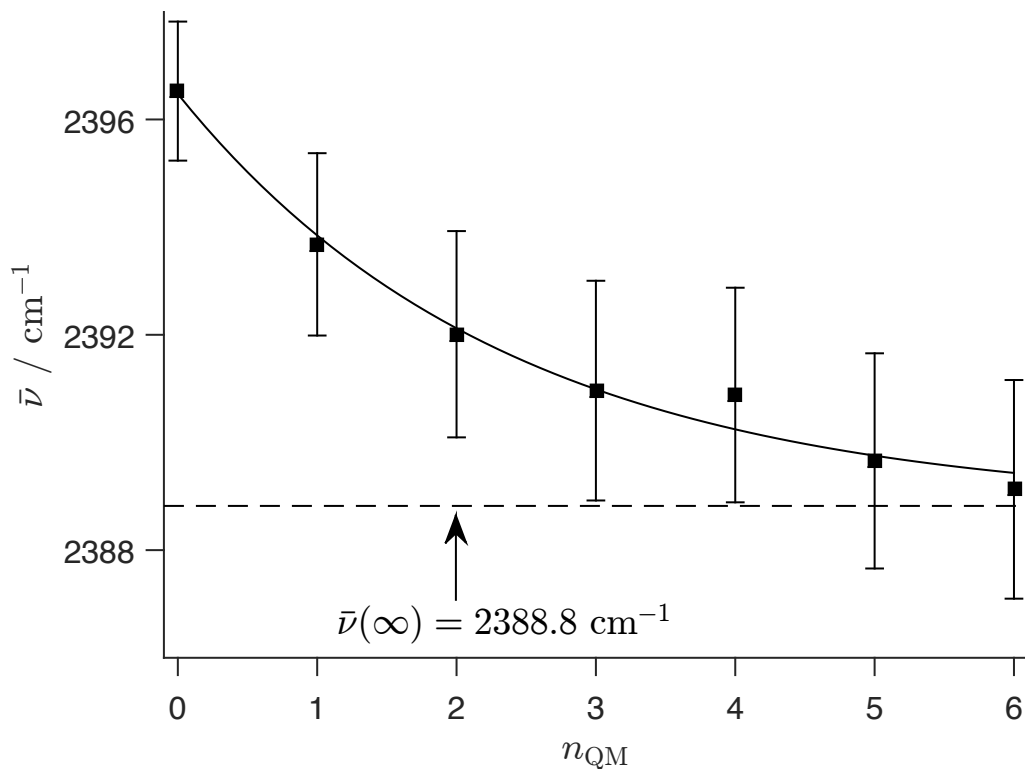


Figure 20: Dependence of DVR  $\nu_3$  mean frequency ( $\bar{\nu}$ ) on the number of solvent ion pairs treated quantum mechanically ( $n_{QM}$ ). Mean frequencies were calculated from  $N = 10$  representative MD snapshots using a B3LYP/SP potential energy surface. Error bars represent the standard deviation of the mean ( $SD_{\bar{x}} = \sigma_x / \sqrt{N}$ ). The data were fitted to an exponential decay ( $\bar{\nu}_{QM} = A \exp(-kn_{QM}) + \bar{\nu}_{converged}$ ), with weights of  $1/SD_{\bar{x}}^2$  ( $R^2 = 0.99$ ). The resulting estimate for the fully converged frequency at this level of theory is  $2388.8 \pm 1.5 \text{ cm}^{-1}$ .

asymmetric stretch frequency reasonably well and offer a good balance between accuracy and cost for our DVR calculations.

### 3.4.7 Predicted Solvation Shift

A central objective of this publication is to establish a computational approach that will allow us to develop a reliable spectroscopic map for the prediction and interpretation of 2D-IR spectroscopic signatures for CO<sub>2</sub> solvated in ionic liquids. While the spectroscopic map and its validation are the subject of the second publication in this series, we can use the predicted solvation shift (i.e., the frequency shift between gas-phase CO<sub>2</sub> and dissolved in [C<sub>4</sub>C<sub>1</sub>im][PF<sub>6</sub>]) as a first indicator for the quality of the computational and chemical model. We note that absolute frequencies are very hard to reproduce with all but the most sophisticated quantum chemical approaches, which are prohibitively expensive for the present application. We therefore scale the calculated frequencies by a factor of 0.9855 so the DVR/B3LYP/LP gas-phase result of 2384 cm<sup>-1</sup> (Tab. 13) agrees with the experiment (2349 cm<sup>-1</sup>). At the proposed DVR/B3LYP/LP level (within a QM/MM approach with  $n = 2$  ion pairs in the QM region and sampling from 85 MD snapshots for the solvated CO<sub>2</sub>), the scaled frequency of the solvated CO<sub>2</sub> is 2339 cm<sup>-1</sup> (Tab. 14). The predicted solvation frequency shift is then -10 cm<sup>-1</sup>, which is in good agreement with the experimental solvation shift of -6.8 cm<sup>-1</sup>. We note that these results were calculated from only 85 MD snapshots and are therefore probably not statistically converged. We will present an in-depth validation of convergence with respect to number of MD steps in the subsequent paper. The already rather nice agreement makes us optimistic that an accurate spectroscopic map can be developed using the computational approach proposed here.

## 3.5 CONCLUSIONS

This work provides fundamental insights into the molecular origins of the vibrational frequency shifts of CO<sub>2</sub> in ionic liquids. First, we validated the computational methodology

used for predicting vibrational frequencies — as defined by basis set, wave function or density functional method, size of the QM region, and the role of anharmonicity. Second, we analyzed the physical origins of CO<sub>2</sub>-IL binding and its interplay with the vibrational frequency shifts using ALMO and SAPT energy decomposition schemes. Our calculations provide insights that are perhaps quite surprising, but definitely nuanced. In a previous publication we concluded that, unlike for other vibrational chromophores, electrostatics alone poorly predict the vibrational frequency shifts for CO<sub>2</sub> in different ILs. From ALMO-EDA we concluded that this frequency shift between different ionic liquids is actually not driven by the electrostatics, but by charge transfer from the anion. This work confirms that our previous results are qualitatively independent from the choices of basis set, ab initio method, and treatment of anharmonicity, confirming the validity of our previous results within the ALMO framework. For different points on the potential energy surface of a *single* ionic liquid (here: [C<sub>4</sub>C<sub>1</sub>im][PF<sub>6</sub>]), however, the energetics are dominated by electrostatics plus Pauli repulsion. This frequency shift mechanism is surprising, firstly because it is fundamentally different from the mechanism that drives the CO<sub>2</sub> frequency shift between different ILs, and secondly because electrostatics plus Pauli repulsion is a relatively small energetic contribution compared to, e.g., dispersion interactions, which one would therefore expect to dominate the frequency shifts. An important practical consequence of this finding is that density functional theory is expected to be sufficiently accurate for constructing potential energy surfaces for CO<sub>2</sub> in [C<sub>4</sub>C<sub>1</sub>im][PF<sub>6</sub>], as needed for the DVR calculations to construct a reliable spectroscopic map.

Similarly, we established appropriate computational and chemical models for treating the extended solvent environment. Our calculations show that a QM/MM treatment with CO<sub>2</sub> plus 2 cation-anion pairs treated quantum mechanically yields vibrational frequencies that are sufficiently close to the converged QM results. Furthermore, adding around 32 ion pairs to the MM solvent box leads to vibrational frequencies converged to within 1 cm<sup>-1</sup>.

In summary, this work elucidated the molecular binding mechanism of CO<sub>2</sub> in the [C<sub>4</sub>C<sub>1</sub>im][PF<sub>6</sub>] ionic liquid and its interplay with the CO<sub>2</sub> asymmetric stretch frequency  $\nu_3$ , and established computational protocols for the reliable construction of spectroscopic maps for simulating ultrafast 2D-IR data of CO<sub>2</sub> solvated in ILs. For future publications, it

will be interesting to employ similar energy decomposition schemes to analyze the energetics and spectral signatures of other chromophores such as  $\text{SCN}^-$ ,  $\text{N}_3^-$ , amides, and phosphates.

### 3.6 ACKNOWLEDGEMENTS

SAC is grateful for financial support from the National Science Foundation (CHE-1565471), the American Chemical Society Petroleum Research Fund (52648-ND6), and the Sustainable Energy Initiative at the University of Notre Dame. This material is based upon work supported by the National Science Foundation under Grant No. (CHE-1454105). TAB acknowledges support of the Pittsburgh Quantum Institute. DSL thanks the University of Pittsburgh for start-up funding. EJB thanks Dr. Mary Sherman for performing initial SAPT calculations and cclib<sup>173,174</sup> for the analysis framework. The authors thank Prof. Kenneth D. Jordan for helpful discussions. SAC and CAD are also thankful for high-performance computing resources and support from the Center for Research Computing at the University of Notre Dame. EJB and DSL thank the Center for Simulation and Modeling at the University of Pittsburgh for providing computational resources.

### 3.7 SUPPORTING INFORMATION

Statistics on calculated harmonic frequencies across different MD snapshots, bin sizes and weights for 15-25 representative structures sampled from MD snapshots, method and basis set dependence, QM/MM box size convergence, quartic fit of the DVR potential energy surface, and effects of BSSE and CP on harmonic frequencies of  $\text{CO}_2$  asymmetric stretch across different ionic liquids.

Table 17: Statistics for dependence of CO<sub>2</sub> asymmetric stretch frequencies on quantum chemical method and basis set dependence on 1000 MD snapshots, 0 QM/256 MM. All frequencies in cm<sup>-1</sup>.

method	basis set	min	max	range	mean	median	standard deviation (population)	standard deviation (sample)
BLYP	6-31G(d,p)	2179.26	2915.23	735.97	2551.84	2552.16	112.92	112.86
TPSS	6-31G(d,p)	2247.16	2967.72	720.56	2611.68	2612.06	110.39	110.33
B3LYP	6-31G(d,p)	2141.31	2892.22	750.91	2522.55	2522.96	115.09	115.03
$\omega$ B97X-D	6-31G(d,p)	2138.86	2894.15	755.29	2523.38	2523.93	114.14	114.08
HF	6-31G(d,p)	1990.89	2814.31	823.42	2415.50	2416.41	125.46	125.40
RI-MP2	6-31G(d,p)	2264.73	2964.82	700.09	2615.07	2615.27	107.68	107.62
BLYP	cc-pVTZ	2083.99	2802.55	718.56	2448.66	2449.17	110.15	110.10
TPSS	cc-pVTZ	2164.75	2871.70	706.95	2523.22	2523.74	108.27	108.22
B3LYP	cc-pVTZ	2047.30	2781.22	733.92	2420.89	2421.51	112.37	112.31
$\omega$ B97X-D	cc-pVTZ	2048.35	2787.24	738.89	2425.06	2425.52	111.67	111.61
HF	cc-pVTZ	1908.93	2714.98	806.05	2325.54	2326.49	122.60	122.54
RI-MP2	cc-pVTZ	2165.68	2842.63	676.95	2504.51	2504.63	104.39	104.34

Table 18: Weights and bin counts for each quantum chemical method and basis set, 0 QM/256 MM. Weights from B3LYP/6-31(d,p) are used for selecting the structures for box size dependence and SAPT calculations.

method	basis set	bin edges	histogram	weights	sum (histogram)
BLYP	6-31G(d,p)	[2269.54 2382.46 2495.38 2608.30 2721.23 2834.13]	[59 245 398 224 64]	[0.05959596 0.24747475 0.40202020 0.22626263 0.06464646]	990
TPSS	6-31G(d,p)	[2335.71 2446.10 2556.49 2666.88 2777.27 2887n.66]	[59 244 400 223 64]	[0.05959596 0.24646465 0.40404040 0.22525253 0.06464646]	990
B3LYP	6-31G(d,p)	[2234.82 2349.91 2465.00 2580.09 2695.18 2810.27]	[59 245 398 224 64]	[0.05959596 0.24747475 0.40202020 0.22626263 0.06464646]	990
$\omega$ B97X-D	6-31G(d,p)	[2238.04 2352.18 2466.31 2580.45 2694.59 2808.72]	[57 241 401 225 63]	[0.05775076 0.24417427 0.40628166 0.22796353 0.06382979]	987
HF	6-31G(d,p)	[2101.85 2227.31 2352.77 2478.23 2603.69 2729.15]	[61 240 399 228 63]	[0.06155399 0.24217962 0.40262361 0.23007064 0.06357215]	991
RI-MP2	6-31G(d,p)	[2345.88 2453.56 2561.23 2668.91 2776.58 2884.26]	[59 246 398 223 64]	[0.05959596 0.24848485 0.40202020 0.22525253 0.06464646]	990
BLYP	cc-pVTZ	[2173.29 2283.44 2393.59 2503.74 2613.89 2724.04]	[59 245 398 224 64]	[0.05959596 0.24747475 0.40202020 0.22626263 0.06464646]	990
TPSS	cc-pVTZ	[2252.54 2360.81 2469.09 2577.36 2685.64 2793.91]	[59 245 396 226 64]	[0.05959596 0.24747475 0.40000000 0.22828283 0.06464646]	990
B3LYP	cc-pVTZ	[2139.97 2252.34 2364.70 2477.07 2589.44 2701.80]	[59 244 399 224 64]	[0.05959596 0.24646465 0.40303030 0.22626263 0.06464646]	990
$\omega$ B97X-D	cc-pVTZ	[2145.89 2257.56 2369.22 2480.89 2592.56 2704.22]	[57 243 399 225 63]	[0.05775076 0.24620061 0.40425532 0.22796353 0.06382979]	987
HF	cc-pVTZ	[2019.04 2141.64 2264.24 2386.84 2509.44 2632.04]	[61 240 401 226 63]	[0.06155399 0.24217962 0.40464178 0.22805247 0.06357215]	991
RI-MP2	cc-pVTZ	[2243.53 2347.92 2452.32 2556.71 2661.11 2765.50]	[59 247 394 226 64]	[0.05959596 0.24949495 0.39797980 0.22828283 0.06464646]	990

### 3.7.1 Method and Basis Set Dependence of Harmonic Frequencies

It is imperative to investigate how sensitive the prediction of relative trends is with respect to the computational approach. To this end, we consider snapshots from MD simulations (see Ref. [143] for details), which allow us to test how well different computational approaches can predict *trends* in dependence of the local coordination environment around the CO<sub>2</sub> and the bulk solvent structure. Fig. 21 shows the CO<sub>2</sub>  $\nu_3$  harmonic frequencies calculated for 1000 statistically uncorrelated MD snapshots (0 QM/256 MM) using various SCF-type approaches, as compared to Møller-Plesset perturbation theory to second order (MP2) as the least expensive wave function-based method that incorporates dispersion effects.<sup>166</sup> The predicted harmonic frequencies in Fig. 21 are parallel to each other for most of the frequency range, independent of method and basis set choice. In fact, there is almost complete overlap between the B3LYP/SP and  $\omega$ B97X-D/SP distributions, and to a lesser degree for B3LYP/VTZ and  $\omega$ B97X-D/VTZ. There is minor crossing of curves in two instances, 1. TPSS/VTZ with  $\omega$ B97X-D/SP, and 2. HF/SP with both B3LYP/VTZ and  $\omega$ B97X-D/VTZ, but nevertheless the conservation of qualitative trends seems excellent. The ordering of calculated frequencies is, from smallest to largest, HF, B3LYP,  $\omega$ B97X-D, BLYP, TPSS, and MP2, using the SP basis set. Within the VTZ basis, the ordering of MP2 and TPSS is interchanged. As expected, the MP2 results are most sensitive to the basis set size, as wave-function based correlation methods require larger basis sets for convergence compared to self-consistent field approaches (HF and DFT). The VTZ basis set predicts frequency distributions that are red-shifted compared to the SP basis set, in agreement with the trends observed from Table 11.

These results for relative trends in vibrational frequencies are highly encouraging. Aside from a multiplicative scaling factor, any of the common quantum chemical methods investigated here can qualitatively reproduce the distribution of harmonic frequencies. The similarity in performance between B3LYP and  $\omega$ B97X-D is somewhat surprising, as the former is a global hybrid with a fixed fraction of HF exchange over all interelectronic distances, and  $\omega$ B97X-D is a range-separated functional with a variable fraction of HF exchange. An explanation of this observation will be provided together with the SAPT results.

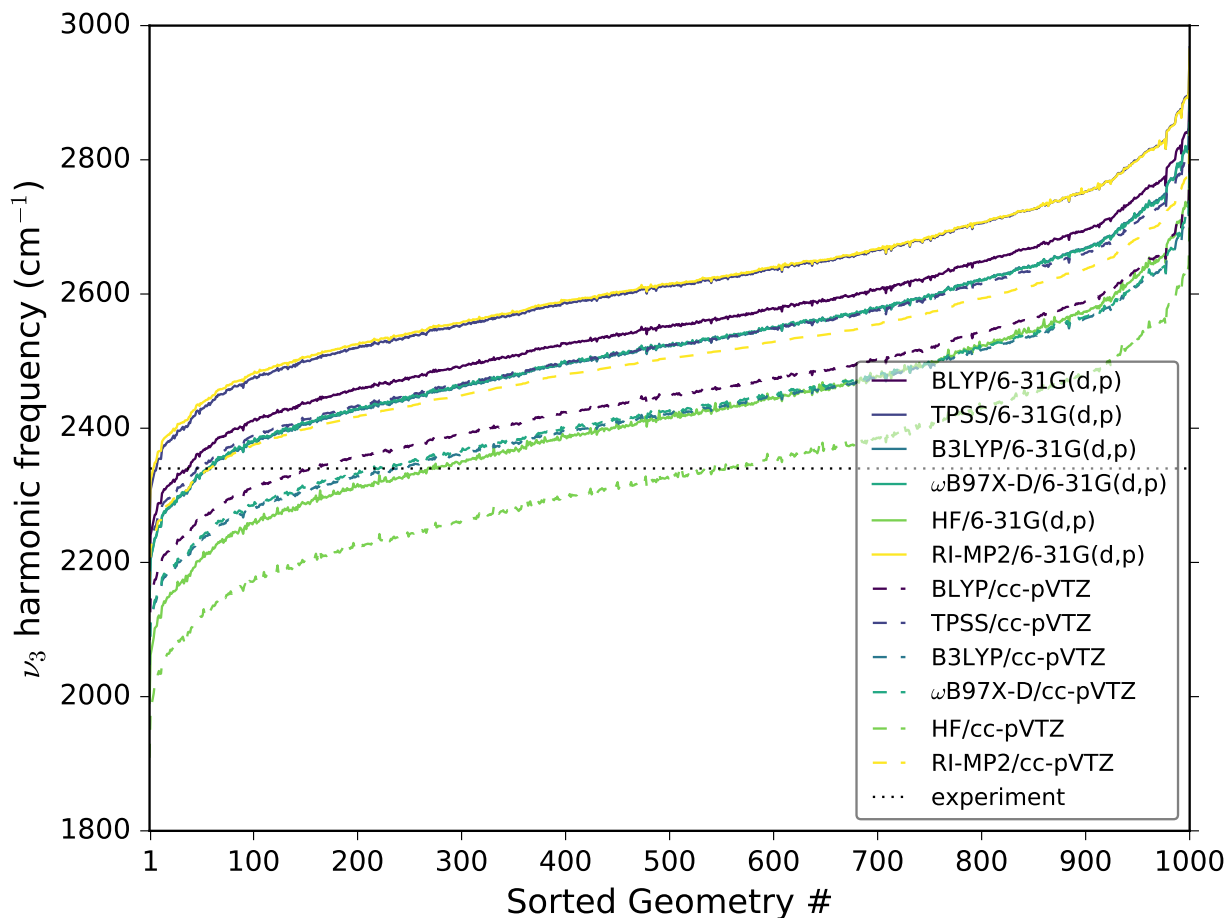


Figure 21: Harmonic  $\nu_3$  frequencies (unscaled) for different methods and basis sets, ordered from lowest to highest frequency according to MP2/VTZ (0 QM/256 MM). The TPSS meta-GGA<sup>175</sup> and the  $\omega$ B97X-D<sup>176</sup> hybrid GGA density functionals are included as representatives of more recent functionals.  $\omega$ B97X-D is both a range-separated functional, with a minimum 22% HF exchange at  $r_e = 0$ , increasing smoothly to 100% as  $r_e \rightarrow \infty$ , and it includes an empirical dispersion correction. To aid in recognizing trends, we have ordered the structures by increasing MP2/VTZ frequency. MP2 calculations use the resolution of the identity (RI) approximation<sup>177–179</sup> and the cc-pVQZ-RI fitting basis set.<sup>180</sup>



To confirm that the good agreement in relative trends is not an artifact of the reordering, CO<sub>2</sub>  $\nu_3$  harmonic frequencies for the first 50 of 1000 MD snapshots are shown in Fig. 22 (0 QM/256 MM, VTZ basis set). Although there are large absolute jumps between snapshots due to the 50 ps time step between them, the ordering between quantum chemical methods does not change, and the gaps between each method are constant.

### 3.7.2 Potential Energy Surface Fitting

We fitted the discrete variable representation (DVR) of the Born-Oppenheimer potential energy surface of gas phase CO<sub>2</sub> to a vibrational potential expanded in the local modes basis using a nonlinear least squares method. The potential energy expansion was truncated at fourth order, and off-diagonal fourth order terms (e.g.  $c_3[(x - x_0)^2(y - y_0)^2]$ ) were omitted, using the following form:

$$\begin{aligned}
 V(x, y) = & a_1(x - x_0)^2 + a_2(y - y_0)^2 + a_3(x - x_0)(y - y_0) \\
 & + b_1(x - x_0)^3 + b_2(y - y_0)^3 + b_3[(x - x_0)^2(y - y_0) + (x - x_0)(y - y_0)^2] \\
 & + c_1(x - x_0)^4 + c_2(y - y_0)^4
 \end{aligned}$$

The resulting fit to the data was excellent ( $R^2 = 0.9997$ ). The parameters are given in Table 24.

The force constants and internal coordinates were calculated (the F matrix) and multiplied by the appropriate mass weighting (G matrix). The product was diagonalized to obtain eigenvalues and eigenvectors, which give the normal modes and frequencies of the symmetric stretch ( $\nu_1 = 1377 \text{ cm}^{-1}$ ) and antisymmetric stretch ( $\nu_3 = 2414 \text{ cm}^{-1}$ ).

### 3.7.3 $\nu_3$ Frequency Convergence with Increasing QM Size

To test the qualitative accuracy of the calculations using smaller QM regions, we examined the correlation between frequencies calculated with  $n = 0 - 5$  QM pairs and 6 QM pairs for each snapshot. The data were fitted using a linear regression analysis.

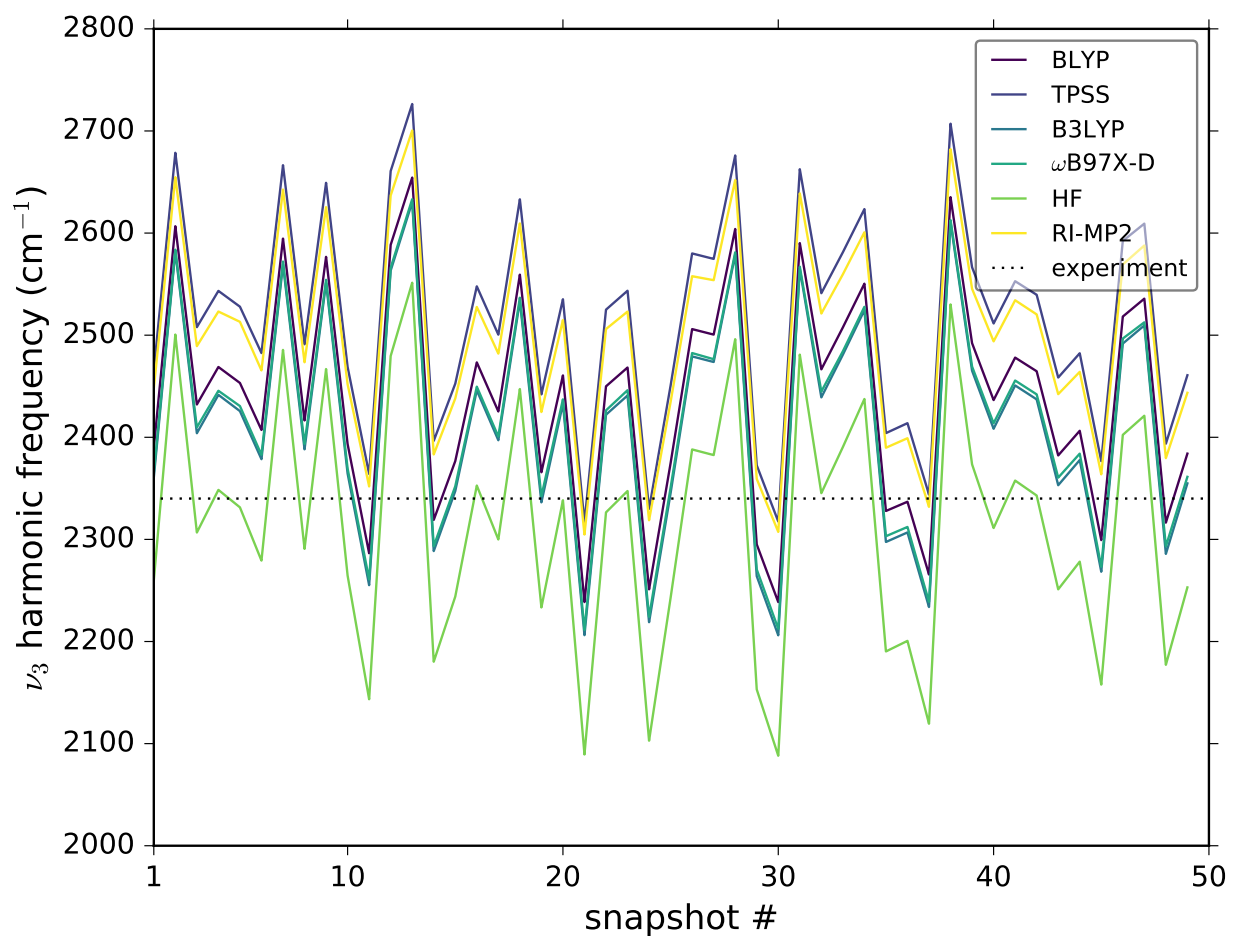


Figure 22: Harmonic  $\nu_3$  frequencies (unscaled) for the first 50 (out of 1000) snapshots calculated with different methods and the cc-pVTZ basis set (0 QM/256 MM).

Table 19: Effect of empirical and self-consistent dispersion corrections on ALMO-decomposed interaction energies with comparison to SAPT0. Values are weighted averages over the same 15 snapshots as in Table 15. Energies are reported in kcal mol<sup>-1</sup>.

method	$E_{\text{frz}}$	$E_{\text{pol}}$	$E_{\text{CT}}$	$E_{\text{disp}}$	$E_{\text{tot}}$
HF/6-31G(d,p)	0.082	-1.303	-0.102	—	-1.323
B3LYP/6-31G(d,p)	0.500	-1.241	-1.177	—	-1.918
B3LYP-D2/6-31G(d,p)	0.500	-1.241	-1.177	-6.269	-8.187
B3LYP-D3/6-31G(d,p)	0.500	-1.241	-1.177	-6.464	-8.382
$\omega$ B97X-D/6-31G(d,p)	-3.343	-1.245	-1.168	—	-5.757
$\omega$ B97X-D/cc-pVTZ	-3.114	-1.617	-1.209	—	-5.941
$\omega$ B97M-V/6-31G(d,p)	-5.494	-1.256	-0.523	—	-7.274
$\omega$ B97M-V/cc-pVTZ	-5.149	-1.649	-0.889	—	-7.687
SAPT0/6-31G(d,p)/MCBS	-2.450	-2.049		-4.135	-8.634
SAPT0/6-31G(d,p)/DCBS	0.593	-1.910	-0.280	-4.714	-6.031
SAPT0/jun-cc-pVTZ/MCBS	1.009	-2.473		-6.930	-8.394
SAPT0/jun-cc-pVTZ/DCBS	1.469	-2.382	-0.203	-7.413	-8.326

Table 20: Effect of allowing or disallowing charge transfer using the ALMO approximation when applied during geometry optimization and/or the harmonic frequency calculation. on the CO<sub>2</sub>  $\nu_3$  harmonic frequency. Clusters are with 1 CO<sub>2</sub>, 1 MMIM cation, and 1 anion. All frequencies are in cm<sup>-1</sup> and unscaled. All calculations were performed using B3LYP/SP with a (100,302) grid.

Anion	full	ALMO geom SCF Hessian	ALMO geom ALMO Hessian	SCF geom ALMO Hessian
[BF <sub>4</sub> ] <sup>-</sup>	2434.70	2438.56	2441.62	2437.69
[DCA] <sup>-</sup>	2430.90	2439.10	2442.59	2434.75
[PF <sub>6</sub> ] <sup>-</sup>	2437.50	2438.47	2441.03	2440.03
[SCN] <sup>-</sup>	2430.30	2438.78	2441.37	2433.14
[TFA] <sup>-</sup>	2429.80	2438.39	2441.06	2432.93
[Tf <sub>2</sub> N] <sup>-</sup>	2437.70	2438.76	2440.85	2439.46
[TfO] <sup>-</sup>	2433.90	2439.37	2442.54	2436.75

Table 21: Effect of counterpoise correction on CO<sub>2</sub>  $\nu_3$  harmonic frequency when applied during geometry optimization and/or the harmonic frequency calculation. Clusters are with 1 CO<sub>2</sub>, 1 MMIM cation, and 1 anion. All frequencies are in cm<sup>-1</sup> and unscaled. The first column is from our previous paper. CP-corrected calculations were performed using Cuby as a driver for Turbomole 6.6 at the B3LYP/SP level with numerical integration grid 7.

	CP-correction for geometry? / CP-correction for frequencies?			
anion	no/no	yes/no	no/yes	yes/yes
TFA	2429.31	2428.05	2457.238	2440.879
SCN (S-coordinated)	2430.24	2427.91	2445.642	2438.003
DCA	2430.47	2426.60	2436.113	2431.383
SCN (N-coordinated)	2431.64	2427.95	2432.691	2437.876
TfO	2431.91	2428.95	2441.921	2443.665
BF <sub>4</sub>	2434.69	2431.48	2454.767	2450.890
Tf <sub>2</sub> N	2435.80	2431.49	2451.851	2441.309
PF <sub>6</sub>	2437.74	2432.58	2479.806	2469.473

Table 22: Counterpoise correction analysis for  $[\text{C}_1\text{C}_1\text{im}][\text{BF}_4]$ . All calculations use B3LYP/SP. The CP geometry is from Cuby driving Turbomole. The no CP geometry is from Q-Chem. (100,302) is the grid used in all calculations for the 1st paper.<sup>27</sup>

Geometry?	Hessian?	Program	XC grid	Frequency ( $\text{cm}^{-1}$ )
CP	CP	Cuby (Turbomole)	7	2450.89
CP	CP	Cuby (Turbomole)	m5	2450.871
CP	no CP	Cuby (Turbomole)	7	2435.634
CP	no CP	Turbomole	7	2434.72 <sup>1</sup>
CP	no CP	Q-Chem	(75,302)	2431.55
CP	no CP	Q-Chem	(100,302)	2431.48 <sup>2</sup>
CP	no CP	Q-Chem	(99,590)	2431.77 <sup>3</sup>
no CP	no CP	Q-Chem	(100,302)	2434.69
no CP	CP	Cuby (Turbomole)	7	2454.767

<sup>1</sup> Difference Hessian no CP/CP:  $2450.89 - 2434.72 = 16.17$

<sup>2</sup>  $(100, 302) - (75, 302) = -0.07$

<sup>3</sup>  $(99, 590) - (75, 302) = 0.22$

Table 23: Counterpoise correction analysis for  $[\text{C}_1\text{C}_1\text{im}][\text{PF}_6]$ . All calculations use B3LYP/SP. The CP geometry is from Cuby driving Turbomole. The no CP geometry is from Q-Chem. (100,302) is the grid used in all calculations for the 1st paper.<sup>27</sup>

Geometry?	Hessian?	Program	XC grid	Frequency ( $\text{cm}^{-1}$ )
CP	CP	Cuby (Turbomole)	7	2469.473
CP	no CP	Cuby (Turbomole)	7	2436.846
CP	no CP	Turbomole	7	2435.93 <sup>1</sup>
CP	no CP	Q-Chem	(75,302)	2432.81
CP	no CP	Q-Chem	(100,302)	2432.58 <sup>2</sup>
CP	no CP	Q-Chem	(99,590)	2432.99 <sup>3</sup>
no CP	no CP	Q-Chem	(100,302)	2437.74
no CP	CP	Cuby (Turbomole)	7	2479.806

<sup>1</sup> Difference Hessian no CP/CP:  $2469.473 - 2435.93 = 33.543$

<sup>2</sup>  $(100, 302) - (75, 302) = -0.23$

<sup>3</sup>  $(99, 590) - (75, 302) = 0.18$

Table 24: Best fit parameters for DVR representation of the gas phase CO<sub>2</sub> potential energy surface.

coefficient name	value (with 95% confidence bound)
$a_1$	$1.8834 \pm 0.0178 E_h \text{ \AA}^{-2}$
$a_2$	$1.8834 \pm 0.0178 E_h \text{ \AA}^{-2}$
$a_3$	$0.3305 \pm 0.0087 E_h \text{ \AA}^{-2}$
$b_1$	$-5.0013 \pm 0.1124 E_h \text{ \AA}^{-3}$
$b_2$	$-5.0005 \pm 0.1123 E_h \text{ \AA}^{-3}$
$b_3$	$-0.2205 \pm 0.0330 E_h \text{ \AA}^{-3}$
$c_1$	$6.3508 \pm 0.4031 E_h \text{ \AA}^{-4}$
$c_2$	$6.3485 \pm 0.4031 E_h \text{ \AA}^{-4}$
$x_0$	$1.1594 \pm 0.0009 \text{ \AA}$
$y_0$	$1.1594 \pm 0.0009 \text{ \AA}$



Table 25: Fitting results for correlation of QM/MM harmonic frequencies of 10 MD snapshots, with increasing numbers of anion-cation pairs ( $n_{\text{QM}}$ ) treated quantum mechanically.

$n_{\text{QM}}(1)$	$n_{\text{QM}}(2)$	fitting results		
0	6		Estimate	Standard Error
		Intercept	-20.915	1024.4
		Slope	1.0056	0.42745
		Root Mean Squared Error: 5.52		
		R-squared: 0.409, Adjusted R-Squared: 0.335		
1	6		Estimate	Standard Error
		Intercept	60.512	593.34
		Slope	0.97282	0.24788
		Root Mean Squared Error: 4.2		
		R-squared: 0.658, Adjusted R-Squared: 0.615		
2	6		Estimate	Standard Error
		Intercept	98.181	384
		Slope	0.95775	0.16053
		Root Mean Squared Error: 3.08		
		R-squared: 0.816, Adjusted R-Squared: 0.794		
3	6		Estimate	Standard Error
		Intercept	173.5	306.89
		Slope	0.92667	0.12835
		Root Mean Squared Error: 2.62		
		R-squared: 0.867, Adjusted R-Squared: 0.85		
4	6		Estimate	Standard Error
		Intercept	2.1463	172.16
		Slope	0.99837	0.072008
		Root Mean Squared Error: 1.44		
		R-squared: 0.96, Adjusted R-Squared: 0.955		
5	6		Estimate	Standard Error
		Intercept	-26.98	85.681
		Slope	1.0111	0.035855
		Root Mean Squared Error: 0.717		
		R-squared: 0.99, Adjusted R-Squared: 0.989		

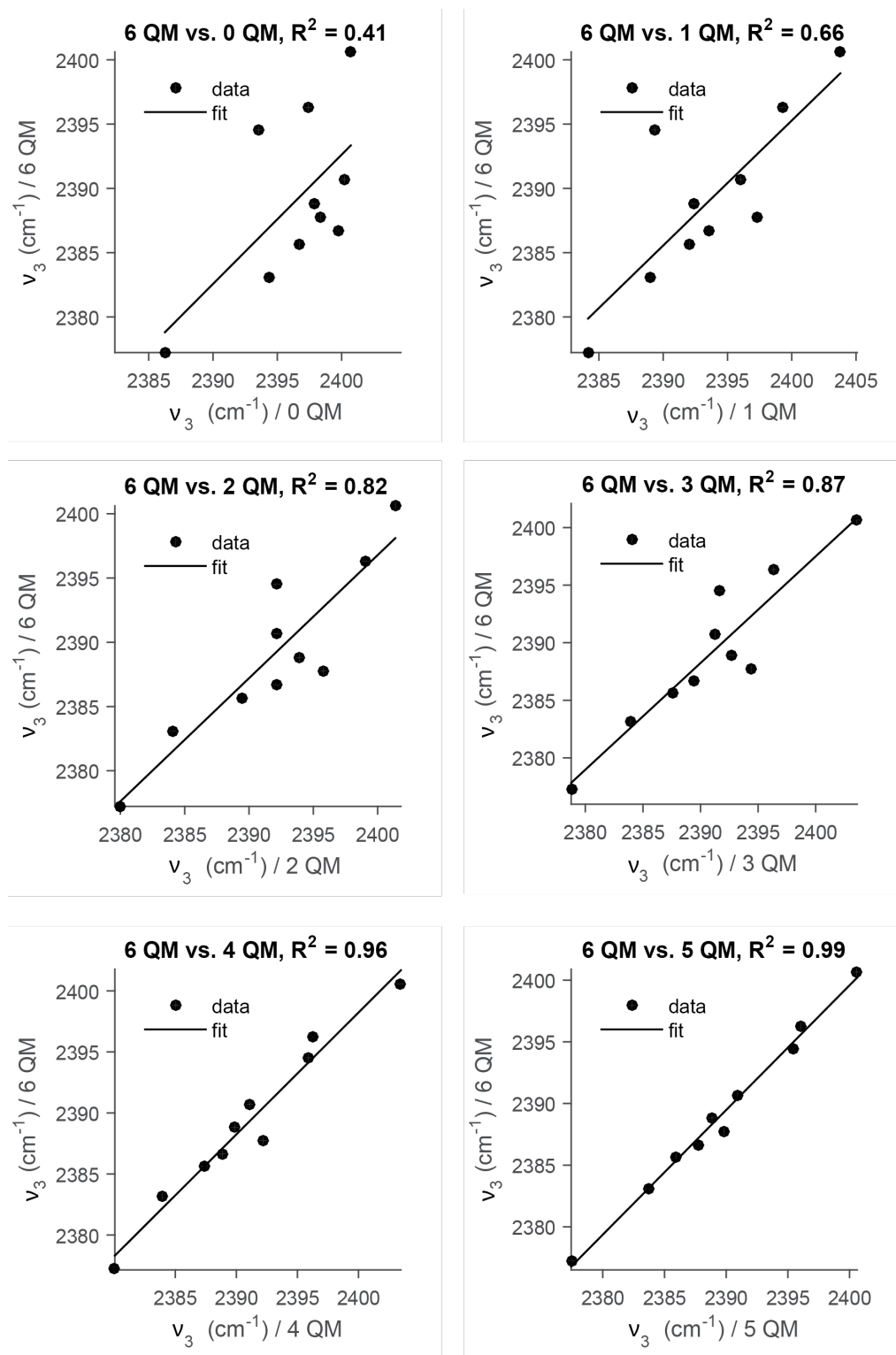


Figure 23: Correlation plots for QM/MM harmonic frequencies of 10 MD snapshots, with increasing numbers of anion-cation pairs treated quantum mechanically.

## 4.0 MODELING CARBON DIOXIDE VIBRATIONAL FREQUENCIES IN IONIC LIQUIDS: II. SPECTROSCOPIC MAP

The text in this chapter has been adapted from Daly, C. A.; Berquist, E. J.; Brinzer, T.; Garrett-Roe, S.; Lambrecht, D. S.; Corcelli, S. A. Modeling Carbon Dioxide Vibrational Frequencies in Ionic Liquids: II. Spectroscopic Map. *J. Phys. Chem. B* **2016**, *120*, 12633–12642, DOI: [10.1021/acs.jpcc.6b09509](https://doi.org/10.1021/acs.jpcc.6b09509), and is copyright the American Chemical Society. The author’s contribution to the work included performing the symmetry-adapted perturbation theory (SAPT) calculations and decomposing each empirical spectroscopic map term in terms of quantum mechanical phenomena.

### 4.1 SUMMARY

The primary challenge for connecting molecular dynamics (MD) simulations to linear and two-dimensional infrared (2D-IR) measurements is the calculation of the vibrational frequency for the chromophore of interest. Computing the vibrational frequency at each time step of the simulation with a quantum mechanical method like density functional theory (DFT) is generally prohibitively expensive. One approach to circumnavigate this problem is the use of spectroscopic maps. Spectroscopic maps are empirical relationships that correlate the frequency of interest to properties of the surrounding solvent that are readily accessible in the MD simulation. Here, we develop a spectroscopic map for the asymmetric stretch of CO<sub>2</sub> in the 1-butyl-3-methylimidazolium hexafluorophosphate ([C<sub>4</sub>C<sub>1</sub>im][PF<sub>6</sub>]) ionic liquid (IL). DFT is used to compute the vibrational frequency of 500 statistically independent CO<sub>2</sub>-[C<sub>4</sub>C<sub>1</sub>im][PF<sub>6</sub>] clusters extracted from an MD simulation. When the map was tested on a

500 different  $\text{CO}_2\text{-[C}_4\text{C}_1\text{im][PF}_6\text{]}$  clusters, the correlation coefficient between the benchmark frequencies and the predicted frequencies was  $R = 0.94$  and the root mean squared error was  $2.7\text{ cm}^{-1}$ . The calculated distribution of frequencies also agrees well with experiment. The spectroscopic map required information about the  $\text{CO}_2$  angle, the electrostatics of the surrounding solvent, and the Lennard-Jones interaction between the  $\text{CO}_2$  and the IL. The contribution of each term in the map was investigated with symmetry-adapted perturbation theory (SAPT) calculations.

## 4.2 INTRODUCTION

Ionic liquids (ILs) have attracted tremendous attention because of their properties as environmentally friendly alternatives to volatile organic solvents, and their applications involving the production, storage, and efficient utilization of energy.<sup>35,36,181–183</sup> ILs exhibit unique physical properties relative to conventional liquids in terms of vapor pressure, viscosity, electrical and thermal conductivity, solubility of polar and nonpolar molecules, and melting point.<sup>36,184–187</sup> Moreover, these properties can be tuned to specific applications by chemically modifying the molecules that comprise the liquid. For example, by functionalizing the molecules of an IL to react with  $\text{CO}_2$ , improved design for preferentially separating  $\text{CO}_2$  from gas mixtures was achieved.<sup>39–41,187,188</sup> Thus, ILs offer a promising new direction for the removal of environmentally harmful  $\text{CO}_2$  from postcombustion flue gas.

It is essential that the fundamental structure and dynamics of ILs be understood to aid in the design of new ILs for unique applications. Unlike conventional solvents, ILs exhibit heterogeneous structure and dynamics that have profound implications for their physical properties. Two-dimensional infrared (2D-IR) spectroscopy offers several unique advantages for interrogating the structure and dynamics of liquids because of its exquisite time and spatial resolution.<sup>67,69,77,189</sup> The spatial resolution results from the size of suitably chosen vibrational chromophores. The vibrational frequencies of these reporters depend sensitively on their local environment.<sup>77,127,144,162,168,190</sup> As that local environment evolves, so too will the vibrational frequency of the probe — a process called spectral diffusion. 2D-IR spectroscopy

measures these frequency fluctuation dynamics, which relate back to the intrinsic dynamics of the surroundings of the vibrational chromophore.

Recently, Brinzer *et al.* have demonstrated that the asymmetric stretch of CO<sub>2</sub> ( $\nu_3$  is an excellent vibrational reporter of its local environment in ILs.<sup>27</sup> In particular, these experiments have established (1) that the asymmetric stretch of CO<sub>2</sub> exhibits a significant solvatochromic shift with respect to the choice of anion in a series of imidazolium-based ILs, (2) that the CO<sub>2</sub> vibrational population lifetime is sufficiently long to measure 2D-IR spectra on a 100 ps timescale, and (3) that the longest spectral diffusion timescale correlates empirically with the viscosity of the IL.<sup>27</sup> Fayer and coworkers have also studied CO<sub>2</sub> in ILs with 2D-IR spectroscopy, including detailed measurements and modeling of the rotational dynamics of CO<sub>2</sub> and how this motion results in reorientational-induced spectral diffusion (RISD). Through analysis of polarization-selective 2D-IR measurements, the RISD contribution to the overall spectral diffusion process was quantified.<sup>119,191</sup> The RISD analysis assumed that shifts in the CO<sub>2</sub> vibrational frequency were governed by a second-order Stark effect.

Among the great successes of multidimensional vibrational spectroscopy is revealing the dynamics of hydrogen-bond network rearrangements in liquid water.<sup>128,192–204</sup> However, these profound insights from experiment were only possible in conjunction with a robust theoretical effort.<sup>176,190,194–196,205–223</sup> Much of that theoretical effort focused on the development and application of empirical relationships connecting the instantaneous vibrational frequency of interest to structural properties — usually the electrostatics — of the surrounding condensed-phase environment.<sup>127,133,206</sup> Such relationships have come to be known as “spectroscopic maps.” With a spectroscopic map in hand, quantities such as the linear IR absorption spectrum, 2D-IR spectra, and the frequency fluctuation correlation function that quantifies spectral diffusion, can be readily calculated in a conventional molecular dynamics (MD) simulation.<sup>133,134,211</sup> With the emergence of 2D-IR measurements on CO<sub>2</sub> in ILs, there is ample motivation to develop a spectroscopic map for the asymmetric stretch of CO<sub>2</sub> in an IL.

In paper I,<sup>135</sup> we developed and validated a robust quantum mechanics/molecular mechanics (QM/MM) protocol for calculating anharmonic CO<sub>2</sub> vibrational frequencies in the 1-butyl-3-methylimidazolium hexafluorophosphate ([C<sub>4</sub>C<sub>1</sub>im][PF<sub>6</sub>]) IL. Here, we have used

the protocol to calculate the asymmetric stretch vibrational frequency of CO<sub>2</sub> in 1000 statistically independent snapshots extracted from an MD simulation. For each frequency calculation, the CO<sub>2</sub> molecule and two pairs of IL molecules are treated quantum mechanically with density functional theory (DFT). The rest of the solvent is included in the calculation as point charges that polarize the quantum mechanical region. The two-dimensional potential energy surface for the CO<sub>2</sub> stretches is constructed on a 12 × 12 grid and the resulting vibrational Schrödinger equation is solved using a discrete variable representation (DVR) method. Once the vibrational frequencies were calculated, 500 of these snapshots were used to parameterize the spectroscopic map and the other 500 snapshots were used to quantify the accuracy of the spectroscopic map.

Previous spectroscopic maps have primarily been based on electrostatics,<sup>168,207,224–228</sup> but our initial quantum chemistry investigations<sup>27,135</sup> indicate that the antisymmetric stretch of CO<sub>2</sub> is sensitive to other physical effects, including charge transfer, dispersion, exchange repulsion, and electrostatics. Accordingly, we found that a suitably accurate spectroscopic map could not be constructed using only electrostatic properties of the IL environment. Instead, we had to include both electrostatic and Lennard-Jones (LJ) terms in the map. Błasiak and Cho previously found that including dispersion interactions resulted in an improved spectroscopic map for the amide I vibration of *N*-methylacetamide.<sup>229</sup> In addition, since the CO<sub>2</sub> molecule was modeled as flexible in solution, the map also has a dependence on the CO<sub>2</sub> bend angle whose contribution was investigated in detail.

Spectroscopic maps are inherently empirical and can, in principle, utilize any variable that is sufficiently correlated with the vibrational frequencies, even if that variable is not the cause of the vibrational frequency shifts. Therefore, the dual goals of this work are to develop and validate a spectroscopic map, and to understand how the causal variables manifest themselves in the map. To achieve the first goal, the average frequency and distribution of vibrational frequencies were compared to inhomogeneous vibrational spectra extracted from 2D-IR measurements. To achieve the second goal a selection of snapshots were analyzed with symmetry adapted perturbation theory (SAPT)<sup>156–158</sup> calculations.

In addition to the intermolecular interactions, CO<sub>2</sub> has an important intramolecular degree of freedom, the bending mode. Our previous work<sup>27</sup> has implicated the bending mode

in the experimentally observed solvatochromic shifts. At room temperature, the bending mode has an energy of approximately  $3k_B T$ , placing it in an intermediate regime where it is not clear if a flexible (classical) or a rigid (quantum) model should be more appropriate. To better understand the role of CO<sub>2</sub> flexibility in the spectroscopic map, we calculated histograms of vibrational frequencies for a rigid (bond angle = 180°) and a flexible model of CO<sub>2</sub> in the [C<sub>4</sub>C<sub>1</sub>im][PF<sub>6</sub>] IL. We also examined a third possibility where the CO<sub>2</sub> is modeled as flexible in the MD simulation, but the bend angle is relaxed prior to applying the spectroscopic map.

The paper is organized as follows. In Section 4.3 the details of the MD simulations and the anharmonic vibrational frequency calculations are described. In Section 4.4, the spectroscopic map is constructed. In Section 4.5, the spectroscopic map is validated by comparison to experiment. In Section 4.6, the contributions of the electrostatic, exchange repulsion, and dispersive interactions in the spectroscopic map are analyzed with ALMO and SAPT calculations. Finally, in Section 4.7 we provide some concluding remarks.

### 4.3 COMPUTATIONAL METHODS

Molecular dynamics (MD) simulations were performed using the large-scale atomic/molecular massively parallel simulator (LAMMPS)<sup>230</sup> with a time step of 2 fs. 256 ion pairs of [C<sub>4</sub>C<sub>1</sub>im][PF<sub>6</sub>] and one molecule of CO<sub>2</sub> were simulated at 300 K in a cubic box with periodic boundary conditions. Previous studies have confirmed that 256 ion pairs is a sufficiently large simulation box to mitigate finite-size effects.<sup>231</sup> The original atomic coordinates and box size (45 Å) were generated from a previous study of [C<sub>4</sub>C<sub>1</sub>im][PF<sub>6</sub>] containing a single water solute, which had been subjected to a rigorous equilibration protocol.<sup>134</sup> The water was replaced with a CO<sub>2</sub> solute, and was subjected to the following equilibration procedure: (1) 1 ns in the NVT ensemble at 300 K, (2) heating to 600 K over 1 ns, (3) cooling to 300 K over 1 ns, (4) 1 ns in the NVT ensemble at 300 K, and (5) 1 ns in the NVE ensemble. Production run trajectories were collected in the NVE ensemble. Energy conservation was excellent, with fits to the energy and temperature over 10 ns revealing slopes of  $3.3 \times 10^{-5}$  kcal mol<sup>-1</sup> ps<sup>-1</sup>

and  $9.8 \times 10^{-6} \text{ K ps}^{-1}$ , respectively. All molecules were modeled as fully flexible except for bonds containing hydrogen, which were held fixed at their equilibrium lengths using the SHAKE algorithm.<sup>232,233</sup> Also, in certain cases (see below), the  $\text{CO}_2$  bond lengths and angle were held fixed at their equilibrium values using the LAMMPS rigid integrator.<sup>230</sup> The force fields for  $[\text{C}_4\text{C}_1\text{im}][\text{PF}_6]$  were the same as in our previous simulation studies involving this IL.<sup>134</sup> Briefly, the bends, bonds, dihedrals, and Lennard-Jones parameters for  $[\text{C}_4\text{C}_1\text{im}]^+$  are from the generalized Amber force field (GAFF),<sup>234,235</sup> and partial charges were obtained from DFT calculations.<sup>236</sup> The  $[\text{PF}_6]^-$  force field parameters were from the work of Liu *et al.*<sup>237</sup> Charges on the ions were scaled by 0.84 to empirically account for charge transfer and polarization effects in the IL.<sup>238,239</sup>  $\text{CO}_2$  was modeled using the TraPPE force field, with additional terms developed by Perez-Blanco and Maginn for flexible bond lengths and angle.<sup>240,241</sup> Lennard-Jones interactions were truncated at  $15 \text{ \AA}$  and the long-ranged electrostatics were computed using particle-mesh Ewald summation with a  $15 \text{ \AA}$  real space cutoff.<sup>242</sup>

In order to create a spectroscopic map, 1000 statistically independent snapshots separated by 50 ps were collected from a pair of 50 ns simulations, one with a fully flexible  $\text{CO}_2$  and a second with a fully rigid  $\text{CO}_2$ . For each snapshot, the Born-Oppenheimer potential energy surface (PES) for  $\text{CO}_2$  stretching modes was obtained from single point energy calculations performed as the CO bond lengths were stretched from  $0.955$  to  $1.45 \text{ \AA}$  in  $0.045 \text{ \AA}$  steps. During these calculations, the nearest two pairs of ions by center of mass were included quantum mechanically, and the remaining ions within  $20 \text{ \AA}$  were included as their point charges from the MD force field. The resulting PES was included in a discretized construction of the Hamiltonian for CO stretches, which was then diagonalized, producing the asymmetric stretch frequency. More details about this method can be found in paper 1 of this series. Least squares multiple linear regression was used to empirically fit the electric field due to the anions and cations along the CO bonds and the Lennard-Jones potential energy on the  $\text{CO}_2$  carbon and oxygens to the asymmetric stretch of  $\text{CO}_2$  for 500 of the flexible snapshots, and the accuracy of the resulting fit was tested using the remaining 500 snapshots. 500 of the rigid snapshots were used as a secondary test set. This is described in more detail in section 4.5. In certain cases, the  $\text{CO}_2$  angle from the flexible simulation was relaxed holding all other degrees-of-freedom and the  $\text{CO}_2$  center of mass fixed prior to vibrational frequency



calculations for further analysis. This is discussed further in Section 4.6.2.

#### 4.4 SPECTROSCOPIC MAP FOR CO<sub>2</sub> VIBRATIONS

Empirical spectroscopic maps relate the instantaneous vibrational frequency of an IR reporter to properties of its surroundings that can be readily accessed in MD simulations.<sup>134,211,243</sup> Once a spectroscopic map has been parameterized, it can be used to calculate IR absorption spectra, 2D-IR spectra, and frequency fluctuation time correlation functions from a MD simulation. For the asymmetric stretch of CO<sub>2</sub> in [C<sub>4</sub>C<sub>1</sub>im][PF<sub>6</sub>], we were unable to obtain a suitably accurate spectroscopic map from electrostatics alone. Instead, we needed to include information about the CO<sub>2</sub> bend angle, as well as the Lennard-Jones (LJ) interactions between CO<sub>2</sub> and the surrounding IL. The spectroscopic map has the following form

$$\omega_a = \omega_g + \Delta\omega_\theta + \Delta\omega_{\text{solvent}} \quad (4.1)$$

where  $\omega_a$  is the predicted CO<sub>2</sub> asymmetric stretch vibrational frequency,  $\omega_g$  is the experimental gas phase frequency (2349.1 cm<sup>-1</sup>),  $\Delta\omega_\theta$  is the dependence of the frequency on the OCO bend angle, and  $\Delta\omega_{\text{solvent}}$  captures the change in the vibrational frequency due to interactions with the IL solvent. Figure 24 shows the dependence of the CO<sub>2</sub> asymmetric stretch vibrational frequency on the OCO angle,  $\theta$ , calculated for CO<sub>2</sub> isolated in the gas-phase. The calculated data are fit exquisitely well ( $R^2 = 0.999$ ) by the single-parameter function

$$\Delta\omega_\theta = a(1 + \cos\theta) \quad (4.2)$$

where  $a = -1160.9 \text{ cm}^{-1}$ .

Figure 24 also shows the vibrational frequency of 500 statistically independent CO<sub>2</sub> in [C<sub>4</sub>C<sub>1</sub>im][PF<sub>6</sub>] snapshots. The vibrational frequencies were calculated using the DVR approach described in paper 1 in this series. In these calculations, the CO<sub>2</sub> and the closest two pairs of [C<sub>4</sub>C<sub>1</sub>im][PF<sub>6</sub>] molecules — determined using the distance between the center-of-mass of the IL molecule and the CO<sub>2</sub> carbon atom — were treated quantum mechanically at the B3LYP/6-311++G(d,p) level of theory. Any IL molecule whose center-of-mass was

within 20 Å was modeled using its molecular mechanics partial atomic charges, which then polarize the quantum mechanical region. IL molecules were added to the molecular mechanics region in pairs to maintain charge neutrality. The overall trend in the vibrational frequencies roughly follows the angle dependence in the gas phase, but there is significant scatter due to interactions with the IL.

A map for the solvent effects on the asymmetric CO<sub>2</sub> vibrational frequency was constructed assuming the following form,

$$\Delta\omega_{\text{solvent}} = b_1 E_{\text{O}}^{\text{Cation}} + b_2 E_{\text{O}}^{\text{Anion}} + c_1 U_{\text{O}} + c_2 U_{\text{C}} \quad (4.3)$$

where  $E$  and  $U$  represent contributions from the electric field and Lennard-Jones (LJ) interactions with the solvent, respectively. The subscript, C or O, indicates whether the interaction is computed at the location of the CO<sub>2</sub> central carbon or at the oxygen atoms. For  $E_{\text{O}}$  and  $U_{\text{O}}$ , the value used in Eq. (4.3) is the average for the two CO<sub>2</sub> oxygen sites. The LJ interaction is computed using the expression,

$$U = \sum_j \varepsilon_j \left[ \left( \frac{\sigma_j}{r_j} \right)^{12} - \left( \frac{\sigma_j}{r_j} \right)^6 \right] \quad (4.4)$$

where the sum is over all atoms in the surrounding liquid,  $\varepsilon_j$  and  $\sigma_j$  are the LJ parameters for the atom, and  $r_j$  is the distance to the atom. The electric fields are calculated with respect to the oxygen atoms of CO<sub>2</sub> and are projected along the relevant CO bond,

$$E = \hat{r}_{\text{CO}} \cdot \sum_j \frac{q_j \hat{r}_j}{r_j^2} \quad (4.5)$$

where the sum is over all relevant atoms in the surrounding liquid (i.e. those associated with the cations for  $E_{\text{O}}^{\text{Cation}}$  and those associated with the anions for  $E_{\text{O}}^{\text{Anion}}$ ),  $q_j$  is the partial atomic charge,  $r_j$  is the distance to the charge,  $\hat{r}_j$  is a unit-vector directed toward the site of the charge, and  $\hat{r}_{\text{CO}}$  is a unit vector from the carbon atom of CO<sub>2</sub> to the relevant oxygen atom. Long range electrostatics are corrected using the damped shifted force method.<sup>244</sup>

The four parameters,  $b_1$ ,  $b_2$ ,  $c_1$ , and  $c_2$ , in Eq. (4.3) were determined empirically by applying multiple linear regression using the 500 calculated frequencies in the training set (Table 26). The quality of the fit was evaluated using the 500 different frequencies contained

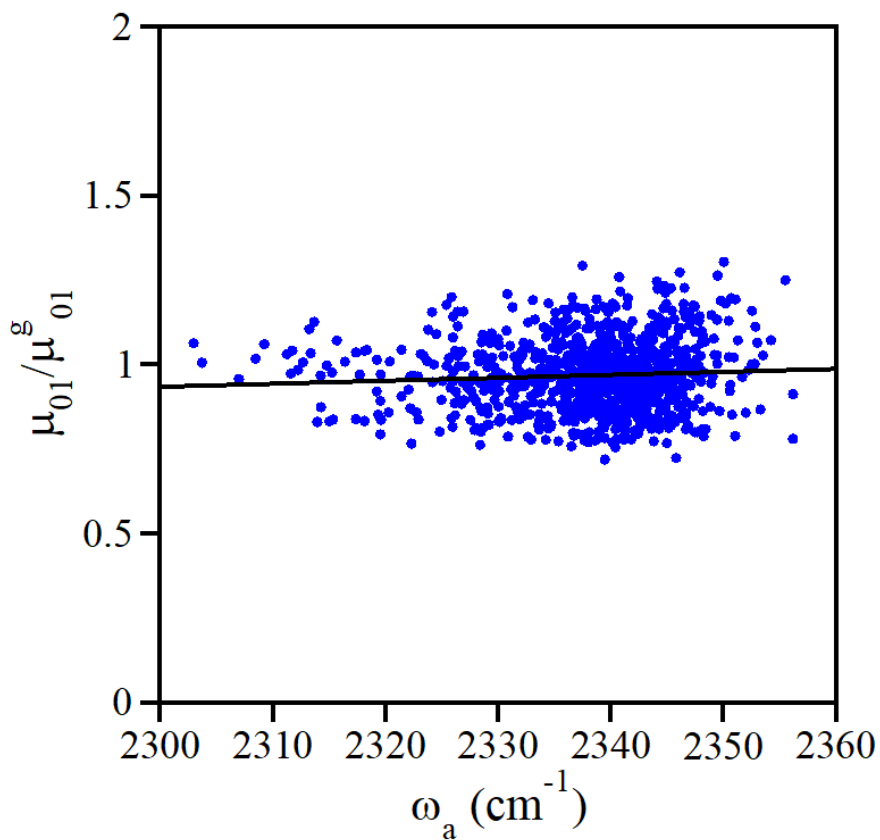


Figure 24: Transition dipole moment integral,  $\mu_{01}$ , of the asymmetric stretch of  $\text{CO}_2$  in 1000  $\text{CO}_2$ - $[\text{C}_4\text{C}_1\text{im}][\text{PF}_6]$  clusters versus the asymmetric stretch vibrational frequency,  $\omega_a$ , where  $\mu_{01}^g$  is the transition dipole moment integral of the asymmetric stretch of  $\text{CO}_2$  in the gas-phase (blue circles). A linear fit of the data (black line) has a slope close to zero indicating that the Condon approximation is reasonable for the asymmetric stretch of  $\text{CO}_2$  in the  $[\text{C}_4\text{C}_1\text{im}][\text{PF}_6]$  IL.

Table 26: Parameters of the spectroscopic map for the CO<sub>2</sub> asymmetric stretch frequency in [C<sub>4</sub>C<sub>1</sub>im][PF<sub>6</sub>]. This map predicts the CO<sub>2</sub> with a regression coefficient  $R = 0.94$  and a root mean squared error of 2.7 cm<sup>-1</sup>. The average shift,  $\langle\Delta\omega\rangle$ , and standard deviation,  $\sigma(\Delta\omega)$ , are reported for each term in the map.

		$\langle\Delta\omega\rangle$ (cm <sup>-1</sup> )	$\sigma(\Delta\omega)$ (cm <sup>-1</sup> )
$\omega_g$	2349.1 cm <sup>-1</sup>	0.0	0.0
$a$	-1160.9 cm <sup>-1</sup>	-6.6	7.0
$b_1$	64.4 cm <sup>-1</sup> a.u. <sup>-1</sup>	-0.1	0.4
$b_2$	93.2 cm <sup>-1</sup> a.u. <sup>-1</sup>	-1.8	0.7
$c_1$	4.70 cm <sup>-1</sup> kcal <sup>-1</sup> mol	-9.5	2.0
$c_2$	-3.55 cm <sup>-1</sup> kcal <sup>-1</sup> mol	7.3	2.1

in the test set (Figure 25). The root-mean-square (RMS) deviation between the test set frequencies and those predicted by Eq. (4.3) was 2.7 cm<sup>-1</sup>, and the value of correlation coefficient for the fit was  $R = 0.94$ . By both metrics, the quality of the spectroscopic map for predicting the CO<sub>2</sub> asymmetric stretch vibrational frequencies in the [C<sub>4</sub>C<sub>1</sub>im][PF<sub>6</sub>] IL is as good or better than previously published maps for other vibrational reporters in conventional solvents. Additionally, when 500 rigid CO<sub>2</sub> snapshots are used as the test set, the same level of accuracy is obtained.

The Condon approximation, that the magnitude of the transition dipole moment is independent of the vibrational frequency of a mode, fails for some solutes that interact in a strong local way with their environment. The most important example is the OH stretch of liquid water. The hydrogen bonds in water polarize the OH bond, increasing the oscillator strength on the red side of the vibrational band, which has a significant effect on the IR absorption line shape.<sup>225,226,245</sup> Similar to the hydrogen bonding of water, the strong local interactions of CO<sub>2</sub> with the ionic liquid anion could, in principle, cause the Condon approximation to fail. However, we find that the Condon approximation for the main band

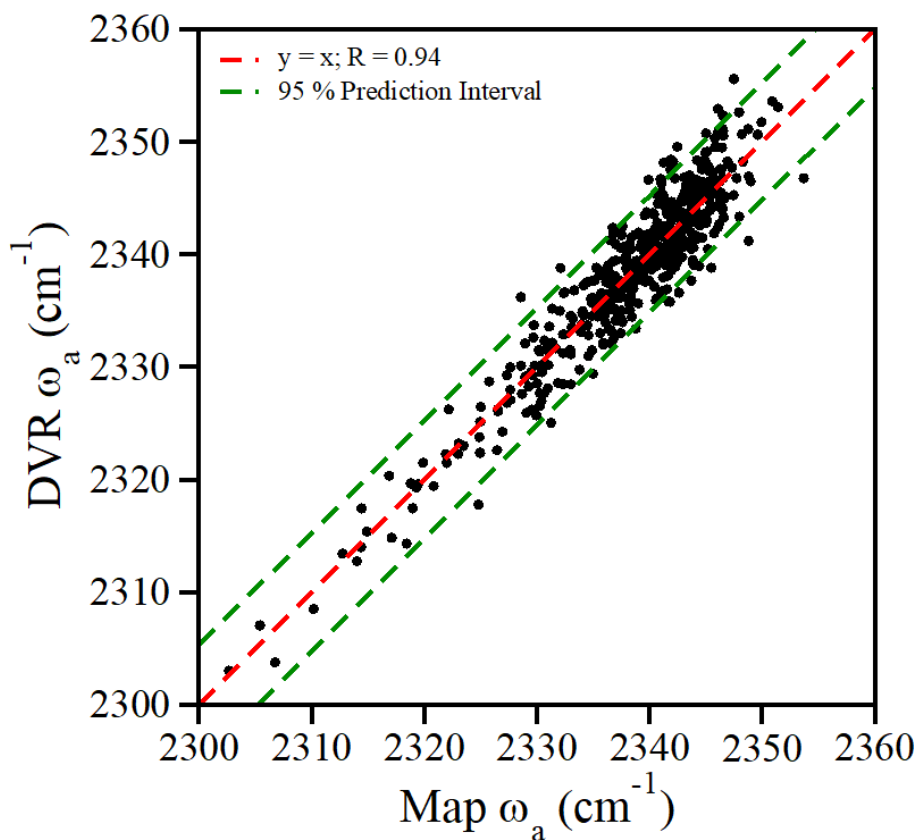


Figure 25: Relationship between  $\text{CO}_2$  asymmetric stretch frequencies in the  $[\text{C}_4\text{C}_1\text{im}][\text{PF}_6]$  IL calculated using the DVR method and those calculated using the spectroscopic map for the 500 test set clusters (black circles). The red line represents a perfect correlation and the 95% prediction interval is indicated with green lines. The spectroscopic map has a regression coefficient of  $R = 0.94$  and a root means squared error of  $2.7 \text{ cm}^{-1}$ .

is adequate (Figure 24). We calculated the transition dipole moment integral,  $\mu_{01}$ , of the asymmetric stretch of CO<sub>2</sub> in 1000 CO<sub>2</sub>-[C<sub>4</sub>C<sub>1</sub>im][PF<sub>6</sub>] clusters. The details of the transition dipole moment integral calculations are provided in the Supporting Information (SI). A plot of  $\mu_{01}$  scaled by  $\mu_{01}^g$ , the transition dipole moment integral of the asymmetric stretch of CO<sub>2</sub> in the gas-phase, versus the asymmetric stretch vibrational frequency,  $\omega_a$ , has a slope close to zero. This confirms that it is reasonable to regard the transition dipole as a constant factor that scales the intensity of linear and non-linear spectra but does not modify their shapes. As a result, we do not treat the environmental dependence of the transition dipole moment in our spectroscopic map; we need only treat the vibrational frequencies.

#### 4.5 PHYSICAL INTERPRETATION OF THE SPECTROSCOPIC MAP

The average contribution to the CO<sub>2</sub> asymmetric stretch vibrational frequency from each of the map components is listed in Table 27. This data demonstrates that the Lennard-Jones potential energy is an important predictor of the vibrational frequency of CO<sub>2</sub> solvated in [C<sub>4</sub>C<sub>1</sub>im][PF<sub>6</sub>], while the electrostatic potential plays a secondary role. This contrasts many prior spectroscopic maps where solvatochromic frequency shifts were based purely on the electrostatics of the environment.<sup>211,224,246</sup> This finding is perhaps surprising at first, because one might expect electrostatics to dominate the interactions of a solute with charged solvent molecules; however, one has to consider that (1) CO<sub>2</sub> is not dipolar or charged, and as such will not interact with uniform electric fields very strongly, and (2) the ionic liquid, particularly the [C<sub>4</sub>C<sub>1</sub>im]<sup>+</sup> butyl tails, have large domains where the dominant interactions are dispersive. These points make it conceivable that van der Waals effects dominate the CO<sub>2</sub>-IL interaction.

To further unravel the origin of the impact of CO<sub>2</sub>-IL interactions on the vibrational signature of CO<sub>2</sub>, we use the fact that the LJ contributions to the spectroscopic map can be further decomposed. In particular, we separate the LJ term into its repulsive ( $\sim r^{-12}$ ) and attractive ( $\sim r^{-6}$ ) contributions (Table 27). We find that the attractive and repulsive LJ terms contribute  $-7.0 \text{ cm}^{-1}$  and  $+4.9 \text{ cm}^{-1}$ , respectively, to the overall LJ vibrational

Table 27: Decomposition of the average LJ contribution to the spectroscopic map for the CO<sub>2</sub> asymmetric stretch frequency in [C<sub>4</sub>C<sub>1</sub>im][PF<sub>6</sub>] into attractive and repulsive components.

LJ Component	Site	$\langle \Delta\omega \rangle$ (cm <sup>-1</sup> )
Attractive	O	-21.7
	C	14.8
	Sum	-6.9
Repulsive	O	12.3
	C	-7.4
	Sum	4.9
Total	O	-9.4
	C	7.3
	Sum	-2.1

shift of  $-2.1 \text{ cm}^{-1}$ . The large contribution from the repulsive LJ term is yet another surprise. To aid in identifying the physical origins of the large repulsive LJ contribution, we performed symmetry adapted perturbation theory (SAPT)<sup>158</sup> calculations that decompose the total interaction energy into physically meaningful components. This analysis should be contrasted with the empirical spectroscopic map, where a good fit implies correlation but not necessarily causation. Our SAPT calculations yield energy contributions, but it should nevertheless be possible to estimate the relative importance of different interactions for vibrational frequencies. The SAPT decomposition supports the previous discussion in that electrostatic interactions (electrostatics, induction) plus the exchange (exchange repulsion, exchange-induction) roughly cancel (total  $-1.3 \text{ kcal mol}^{-1}$ ), whereas dispersive interactions dominate the interaction (total  $-4.7 \text{ kcal mol}^{-1}$  from dispersion plus exchange-dispersion). However, the SAPT data also reveals that exchange-dispersion (the repulsive dispersion part) is over an order of magnitude smaller than the attractive dispersion contribution (10.1% of the total dispersion interaction). This result has to be contrasted with the  $\sim 40\%$  contribution

that the repulsive LJ potential makes to the vibrations. Since the repulsive LJ contribution is the dominant repulsive interaction incorporated in our model, the SAPT results suggest that the repulsive part of the LJ potential fits an agglomerate of exchange (Pauli) repulsion stemming from charge overlap (74.7% of the repulsive interactions), exchange induction (20.7%) plus exchange dispersion (4.6%).

It is likely that LJ components will be an important component of a spectroscopic map for any neutral and nonpolar solute, or any solvent where dispersive interactions, quantum effects (Pauli exchange, for instance), or higher order electrostatic interactions are particularly important. In our case, it seems logical that a higher potential at the carbon would increase the optimal length of the CO bonds, thus decreasing the local mode and normal mode frequencies. Meanwhile, at the oxygen, a larger potential would generally shorten the bond, increasing the frequency. A similar finding was observed by Brinzer *et al.*<sup>27</sup> However, these components only allow the CO<sub>2</sub> vibration to respond to local effects — the electric field components allow it to respond to longer-range interactions. As in prior works for different solvents and solutes, the coefficients for the two electric field components are different from each other, in this case by a substantial margin. It has been previously established that CO<sub>2</sub> interacts with the anions more strongly than with the cations in an ionic liquid.<sup>39,42,101,247–249</sup> This is reflected in the magnitude of the coefficients related to the two components, and in their average frequency contribution (Tables 26 and 27). In particular, CO<sub>2</sub> is a Lewis acid and should generally interact with negatively charged moieties differently from positively charged ones.

## 4.6 VALIDATION

### 4.6.1 Experimental Frequency Distribution

In order to compare our calculated distributions of CO<sub>2</sub> vibrational frequencies with experiment, we must account for the effects that broaden or narrow the IR absorption line shape beyond the underlying distribution of frequencies. The finite population lifetime of the asym-



metric stretch vibration, reorientation of the CO<sub>2</sub> molecule, and a variety of other effects can broaden the absorption spectrum. On the other hand, fast dynamics can narrow the absorption spectrum (i.e. motional narrowing). A faithful comparison to experiment requires a deconvolution of these contributions to estimate the range of instantaneous frequencies experienced by CO<sub>2</sub>.

2D-IR spectra contain sufficient information to recover the distribution of frequencies, which would be difficult to extract from the linear IR absorption spectrum alone.<sup>27</sup> Within the Kubo multi-exponential ansatz, the width of the frequency distribution is determined by the frequency fluctuation correlation function

$$\langle \delta\omega(t)\delta\omega(0) \rangle = \sum_i^N \Delta_i^2 \exp\left(-\frac{t}{\tau_i}\right) \quad (4.6)$$

where  $\Delta_i^2$  are the variances of frequency modulations, and  $\tau_i$  are the timescales for the respective frequency fluctuations. The width of the frequency distribution is the sum of squares of the different broadening processes

$$\langle \delta\omega^2 \rangle = \sum_i^N \Delta_i^2 \quad (4.7)$$

The contribution of homogeneous processes whose frequency fluctuations are too fast to be resolved (specifically when  $\Delta_i\tau_i \ll 1$ ) can be approximated as  $\delta(t)\Delta_H^2\tau_H$ , which results in a frequency correlation function:

$$\langle \delta\omega(t)\delta\omega(0) \rangle = \delta(t)\Delta_H^2\tau_H + \sum_i^{N-1} \Delta_i^2 \exp\left(-\frac{t}{\tau_i}\right) = \frac{\delta(t)}{T_2^*} + \sum_i^{N-1} \Delta_i^2 \exp\left(-\frac{t}{\tau_i}\right) \quad (4.8)$$

where  $T_2^* \equiv (\Delta_H^2\tau_H)^{-1}$  is the pure dephasing time and  $\delta(t)$  is the Dirac delta function. The pure dephasing time depends on the variance of the fast frequency fluctuations,  $\Delta_H^2$ , and the correlation time for fast motions,  $\tau_H$ , and the two parameters cannot be independently determined. Analyzing the change in shape of the 2D-IR spectra as a function of the waiting time can directly determine the magnitude of frequency modulations related to the sum of exponential decays,  $\sum_i^{N-1} \Delta_i^2$ , in Eq. (4.8). For CO<sub>2</sub> in [C<sub>4</sub>C<sub>1</sub>im][PF<sub>6</sub>] this sum is approximately 2 cm<sup>-1</sup>.<sup>27</sup>

Determining the magnitude of frequency modulations that give rise to the first term in Eq. (4.8) is more complicated. The pure dephasing time ( $T_2^*$ ) is only one contributor to the experimentally determined dephasing time ( $T_2$ ), which also depends on the population ( $T_1$ ), and reorientational ( $T_{or}$ ) motions of the molecule,

$$\frac{1}{T_2} = \frac{1}{T_2^*} + \frac{1}{2T_1} + \frac{1}{3T_{or}} \quad (4.9)$$

The experimental dephasing time,  $T_2$ , of the asymmetric stretch of CO<sub>2</sub> in the [C<sub>4</sub>C<sub>1</sub>im][PF<sub>6</sub>] IL is 3.3 ps.<sup>27</sup> Since the experiment was performed in an all-parallel polarization, we cannot unambiguously determine the population and orientation relaxation times. We can estimate them, however, based on the rate of signal decay and the orientational correlation functions determined in a similar ionic liquid.<sup>119,191</sup> Estimates of  $T_1 = 20$  ps and  $T_{or} = 10$  ps, suggest that vast majority contribution to  $T_2$  for CO<sub>2</sub> in [C<sub>4</sub>C<sub>1</sub>im][PF<sub>6</sub>] comes from pure dephasing. Population relaxation and orientational relaxation have a minor effect on the total dephasing time. We estimate a pure dephasing time of  $T_2^* = 4$  ps.

Finally, the variance of the frequency fluctuations,  $\Delta_H^2$ , can be limited to a range by physical constraints on the values of  $\tau_H$ . The lower limit on  $\tau_H$  is governed by the inertial motions of CO<sub>2</sub> and its ionic liquid solvent shells. The timescale of the inertial response in liquid water is in the sub-60 fs range, while that of acetonitrile is 70 fs.<sup>194,250,251</sup> Using 70 fs as a lower limit for  $\tau_H$  places an upper limit on  $\Delta_H$  of 9.7 cm<sup>-1</sup>. Fits to analytical response functions suggests that  $\Delta_H \tau_H \approx 0.2$  is a reasonable estimate of the dynamics that can be resolved using global fitting of the experimental data, which gives an upper limit on  $\tau_H$  of 200 fs, with a corresponding lower limit on  $\Delta_H$  of 6 cm<sup>-1</sup>. Our estimate for the homogeneous width is thus,  $6 < \Delta_H < 10$  cm<sup>-1</sup>. Combining the broadening due to fast and slow motions, the experimentally estimated total frequency width for CO<sub>2</sub> in [C<sub>4</sub>C<sub>1</sub>im][PF<sub>6</sub>] is between 6.3 to 10.2 cm<sup>-1</sup> (Figure 26).

#### 4.6.2 Calculated Frequency Distributions

Figure 27a shows the distribution of CO<sub>2</sub> asymmetric stretch vibrational frequencies computed using the spectroscopic map for 1000 statistically independent snapshots collected

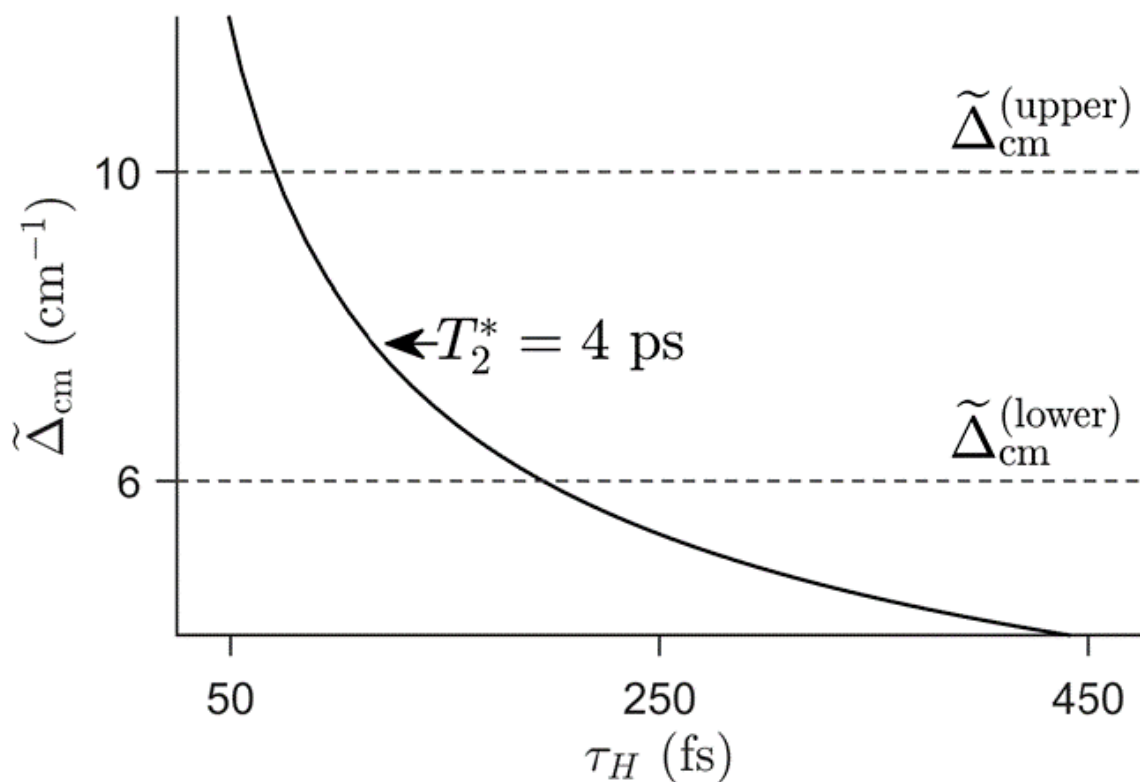


Figure 26: Homogeneous instantaneous linewidth as a function of correlation time for fast motions, with  $T_2^* = 4$  ps, with upper and lower bounds estimated for  $\tilde{\Delta}_H$ . The upper bound, based on an estimated fastest allowed inertial response timescale, and the lower bound, based on a threshold value of  $\Delta_H \tau_H$ , are indicated by dashed horizontal lines. The resulting instantaneous frequency range for homogeneous motions is between 6 to 10  $\text{cm}^{-1}$ .

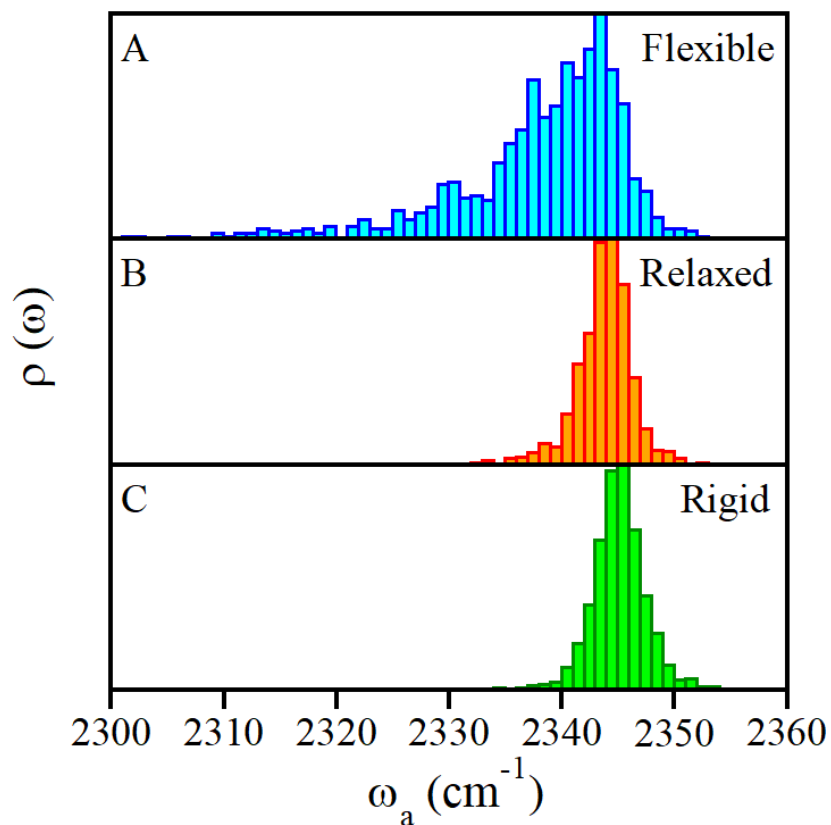


Figure 27: Histograms of the CO<sub>2</sub> asymmetric stretch vibrational frequency,  $\omega_a$ , for 1000 CO<sub>2</sub>-[C<sub>4</sub>C<sub>1</sub>im][PF<sub>6</sub>] clusters. (a) Clusters extracted from an MD simulation of flexible CO<sub>2</sub> in the [C<sub>4</sub>C<sub>1</sub>im][PF<sub>6</sub>] IL. (b) Clusters extracted from an MD simulation of flexible CO<sub>2</sub> in the [C<sub>4</sub>C<sub>1</sub>im][PF<sub>6</sub>] IL, but where the CO<sub>2</sub> geometry is relaxed. (c) Clusters extracted from an MD simulation of rigid CO<sub>2</sub> in the [C<sub>4</sub>C<sub>1</sub>im][PF<sub>6</sub>] IL.

from an MD simulation of flexible CO<sub>2</sub> in [C<sub>4</sub>C<sub>1</sub>im][PF<sub>6</sub>]. These are the same snapshots that were used to parametrize and validate the spectroscopic map in Section 4.4. The distribution is peaked at approximately 2344 cm<sup>-1</sup> and its standard deviation is 7.4 cm<sup>-1</sup>. Both of these values are in reasonable agreement with experiment (2342.5 cm<sup>-1</sup> and 6.3 to 10.2 cm<sup>-1</sup>). Qualitatively, the distribution exhibits a significant asymmetry with a mean frequency of 2339.9 cm<sup>-1</sup> that is about 4 cm<sup>-1</sup> to the red of the peak frequency. The experimental IR absorption line shape, however, does not show signs of such asymmetry in the underlying distribution of frequencies.

The source of the asymmetry in the distribution of frequencies in Figure 27a is the contribution to the spectroscopic map from the CO<sub>2</sub> bend angle, Eq. (4.2). This is illustrated in Figure 27c, where we have calculated the distribution of CO<sub>2</sub> asymmetric stretch vibrational frequencies for 1000 statistically independent snapshots collected from an MD simulation of rigid CO<sub>2</sub> in [C<sub>4</sub>C<sub>1</sub>im][PF<sub>6</sub>]. Since the CO<sub>2</sub> molecule has an angle of 180° in each of the snapshots, the contribution to the calculated vibrational frequency from the CO<sub>2</sub> bend angle is zero. The resulting distribution is correctly symmetric with a mean frequency of 2346.5 cm<sup>-1</sup> and a standard deviation of 2.3 cm<sup>-1</sup>. The calculated distribution is centered 4 cm<sup>-1</sup> to the blue of the experimental distribution, and it is narrower than the lower estimate of the experimental distribution by 4 cm<sup>-1</sup>.

The results in Figure 27a and 27c represent two extremes — one where the CO<sub>2</sub> bend is treated classically (Figure 27a) and another where the CO<sub>2</sub> bend is effectively neglected (Figure 27c). When the CO<sub>2</sub> bend is classical, it is assumed that the CO<sub>2</sub> asymmetric stretch vibrational frequency depends on the instantaneous value of the bend angle, Figure 28 and Eq. (4.2). However, the asymmetric distribution suggests that this approach is incorrect. In fact, a simple thought experiment reinforces the problems associated with regarding the CO<sub>2</sub> bend as a classical variable. Consider a non-rotating CO<sub>2</sub> molecule isolated in the gas phase. If all of the vibrations of the CO<sub>2</sub> molecule are quantum mechanical, the distribution of each of the four vibrations is a delta function. However, if the bend is classical with a kinetic energy commensurate with room temperature, the distribution of asymmetric stretch vibrational frequencies will incorrectly have a finite width. One solution to this conundrum is to adopt a fully quantum mechanical treatment of the CO<sub>2</sub> vibrations. This would require the

construction of a four-dimensional potential energy surface for each of the 1000 benchmark  $\text{CO}_2\text{-}[\text{C}_4\text{C}_1\text{im}][\text{PF}_6]$  clusters, which is computationally intractable.

An alternate strategy is to treat the influence of the  $\text{CO}_2$  bend on the asymmetric stretch vibrational frequency using first-order perturbation theory. Instead of utilizing the instantaneous  $\text{CO}_2$  angle in Eq. (4.2),  $\theta$ , we would instead use the average angle,  $\langle\theta\rangle = \langle\varphi_0|\theta|\varphi_0\rangle$ , where  $\varphi_0(\theta)$  is the ground vibrational wavefunction for the  $\text{CO}_2$  bend. Returning to the  $\text{CO}_2$  in the gas-phase thought experiment, the average angle is constant and equal to  $180^\circ$ . Thus, there would correctly be no contribution to  $\text{CO}_2$  asymmetric stretch vibrational frequency. In contrast, the instantaneous average bend angle will fluctuate away from  $180^\circ$  in an IL because of asymmetric solvation by the solvent. Of course, we do not have access to vibrational wavefunction for the  $\text{CO}_2$  bend for the benchmark  $\text{CO}_2\text{-}[\text{C}_4\text{C}_1\text{im}][\text{PF}_6]$  clusters, nor when we wanted to utilize the spectroscopic map to analyze an MD simulation. An additional approximation is necessary. If we were to regard the  $\text{CO}_2$  bend as harmonic, then the average angle is given by the instantaneous distortion of the  $\text{CO}_2$  geometry by the environment. For the benchmark clusters, the geometry distortion can be determined by optimizing the geometry of the  $\text{CO}_2$  molecule using the classical MD force field and a conjugate gradient minimization while holding fixed both the center-of-mass of the  $\text{CO}_2$ , as well as the configuration of the IL solvent. The map is then used to calculate the vibrational frequency for the relaxed snapshot.

Figure 27b shows the distribution of  $\text{CO}_2$  asymmetric stretch vibrational frequencies computed using the spectroscopic map for 1000 statistically independent snapshots collected from an MD simulation of flexible  $\text{CO}_2$  in  $[\text{C}_4\text{C}_1\text{im}][\text{PF}_6]$  where the  $\text{CO}_2$  bend angle has been relaxed. On average, the relaxed bend angle is  $178.4^\circ$ , and the distribution of frequencies is nearly symmetric with a mean frequency of  $2343.8\text{ cm}^{-1}$  and a standard deviation of  $2.4\text{ cm}^{-1}$ . The mean frequency is in excellent agreement with experiment and differs by only  $1.3\text{ cm}^{-1}$ . Note that this agreement implies that the spectroscopic map is able to accurately capture the solvatochromic shift of the  $\text{CO}_2$  asymmetric stretch vibrational frequency from the gas-phase to the  $[\text{C}_4\text{C}_1\text{im}][\text{PF}_6]$  IL. The width of the distribution is too narrow compared to the estimated width of the experimental distribution of  $6.3$  to  $10.2\text{ cm}^{-1}$ . There are several possible sources for the discrepancy in the width of the distribution, including inaccuracies associated

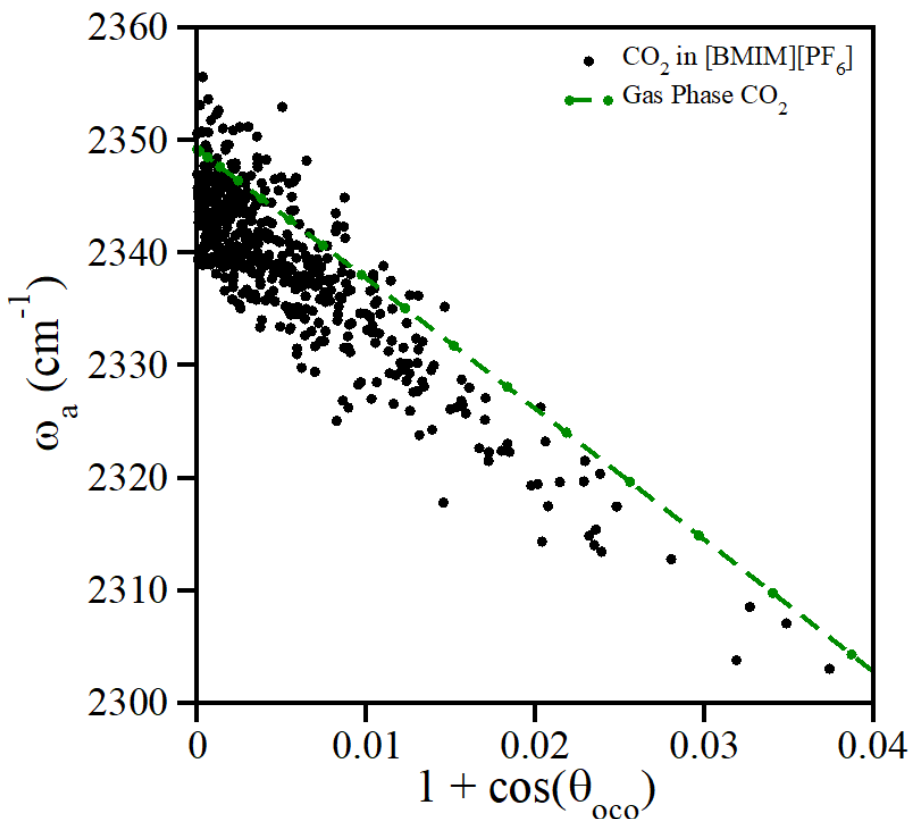


Figure 28: Relationship between the  $\text{CO}_2$  asymmetric stretch vibrational frequency and the OCO angle,  $\theta_{\text{OCO}}$ , for  $\text{CO}_2$  in the gas-phase (green circles) and in the  $[\text{C}_4\text{C}_1\text{im}][\text{PF}_6]$  IL (black circles). The gas-phase data are perfectly correlated with  $1 + \cos \theta_{\text{OCO}}$ . The vibrational frequencies for  $\text{CO}_2$  in the  $[\text{C}_4\text{C}_1\text{im}][\text{PF}_6]$  solvent also show this relationship, but additional solvation effects on the frequency are also present.

with the approximate perturbative approach for the effect of the bend on the asymmetric stretch frequency. However, the overall agreement with experiment is encouraging.

It is instructive to compare the distributions in Figures 27b (relaxed CO<sub>2</sub>) and 27c (rigid CO<sub>2</sub>). Both distributions are symmetric and they have nearly the same widths: 2.4 cm<sup>-1</sup> and 2.3 cm<sup>-1</sup>, respectively. Thus, within the approximate perturbative approach, the bend has very little influence on the width of the distribution. The averages of the distributions differ more significantly: 2343.8 cm<sup>-1</sup> and 2346.5 cm<sup>-1</sup>, respectively. The bend shifts the distribution to the red and into better agreement with experiment. Overall, the role of the bend is relatively minor resulting in a redshift of the distribution by 2.7 cm<sup>-1</sup>. These results suggest several options for how the bend is treated when the map is applied in conjunction with MD simulations to understand the spectroscopy and spectral diffusion dynamics of CO<sub>2</sub> in the [C<sub>4</sub>C<sub>1</sub>im][PF<sub>6</sub>] IL in Paper III<sup>252</sup> in this series. The simplest strategy is to hold the CO<sub>2</sub> rigid and to shift the calculated frequencies by 2.7 cm<sup>-1</sup>. In essence, this would just account for the average effect of the CO<sub>2</sub> angle on the asymmetric stretch frequencies. A more computationally intensive strategy is a simulation with CO<sub>2</sub> flexible but where the geometry of the CO<sub>2</sub> is optimized using the classical force field. The efficacy of these approaches will be evaluated in Paper III.<sup>252</sup>

## 4.7 CONCLUSIONS

In this paper we have developed and validated a spectroscopic map that is the foundation for a molecular interpretation of ultrafast vibrational spectroscopy of CO<sub>2</sub> in ionic liquids. In addition, we have established important insights into the solvatochromic shift of the CO<sub>2</sub> asymmetric stretch vibrational frequency in ILs. We analyzed the physical origin of the vibrational frequency shifts using SAPT energy decomposition schemes. Unlike other vibrational chromophores, such as the OH stretch of HOD in liquid water,<sup>196</sup> electrostatics alone poorly predicts the vibrational frequency. This is consistent with recent studies of the amide I vibration in peptides<sup>229</sup> and nitrile vibrations,<sup>253</sup> where exchange repulsion and dispersion interactions are important for properly describing solvatochromic shifts. While



the most important contributor to the electrostatic part of the spectroscopic map is the field from the anion, both attractive dispersion interactions and repulsive charge overlap forces (Pauli repulsion) play additional important roles. Finally, while the CO<sub>2</sub> bend angle influences the asymmetric stretch frequency, we have shown that the geometry of the CO<sub>2</sub> molecule is only slightly perturbed by the IL, so regarding the CO<sub>2</sub> as rigid is generally sufficient to capture the structural relaxation of the IL relative to the CO<sub>2</sub>.

## 4.8 ACKNOWLEDGEMENTS

The authors thank Prof. Kenneth D. Jordan for helpful discussions. SAC is grateful for financial support from the National Science Foundation (CHE-1565471), the American Chemical Society Petroleum Research Fund (52648-ND6), and the Sustainable Energy Initiative at the University of Notre Dame. SAC and CAD are also thankful for high performance computing resources and support from the Center for Research Computing at the University of Notre Dame. Computational resources were also provided by the Center for Simulation and Modeling at the University of Pittsburgh. SGR acknowledges financial support from the National Science Foundation (CHE-1454105).

## 4.9 SUPPORTING INFORMATION

Details regarding the transition dipole moment integral calculations in Figure 24.

### 4.9.1 Transition Dipole Moment Calculations

The intensity of a vibrational transition,  $\tilde{\nu}_{if}$ , is related to the dipole moment matrix element between the two states,  $\langle \vec{\mu}_{if} \rangle$

$$I(\tilde{\nu}_{if}) = \frac{8\pi^3 N_A}{3hc(4\pi\epsilon_0)} \tilde{\nu}_{if} |\langle \vec{\mu}_{if} \rangle|^2 (N_i - N_f) \quad (4.10)$$

where  $N_A$  is Avogadro’s number,  $N_k$  is the number of particles in the  $k$ th state and  $|\vec{\mu}_{if}|^2$  is the squared norm of the transition dipole moment (TDM) integral between the two states.<sup>254</sup> Because all values in equation 4.10 are constant (at a specific temperature) vibrational intensities for particular transitions are proportional to the squared norm of the TDM vector. Thus, the central property to calculate in order to evaluate the strength of the Condon approximation is  $\langle \vec{\mu}_{if} \rangle$ .

We can calculate the matrix elements of the dipole moment operator in a similar fashion as the bond length matrix elements were calculated in paper 1. Before, we used the value of the bond length at each grid point as a representation of the bond length operator. Similarly, we use the  $x$ ,  $y$ , and  $z$  components of the dipole moment at each grid point (reported by a quantum chemistry program — in this case, Q-Chem<sup>93</sup> — as the appropriate integral over the entire charge density) as a representation of the dipole moment operator,

$$\vec{\mu} = \sum_{k=1}^3 \mu^k \hat{k} \tag{4.11}$$

where  $\hat{k}$  is the  $k^{\text{th}}$  Cartesian basis vector. The dipole moment matrix elements are, for a two dimensional grid,

$$\langle \vec{\mu}_{if} \rangle = \sum_{k=1}^3 \sum_{l=1}^N \sum_{j=1}^N \psi_{jl}^i \mu_{jl}^k \hat{k} \psi_{jl}^f \tag{4.12}$$

where the  $\psi_{jl}^n$  are the vibrational wavefunctions for state  $n$  on grid point  $(j, l)$  returned by the DVR method. We have evaluated the accuracy of this method for CO<sub>2</sub> in two ways. First, we calculate the norm of the TDM integral for the symmetric and asymmetric stretches of CO<sub>2</sub> in the gas phase and compare these to experiment.<sup>255</sup> The results are shown in Table 28, and the accuracy is excellent.

Next, we evaluated the accuracy of this method for CO<sub>2</sub> in solution. A previously used method for evaluating the Condon approximation for vibrational reporters in solution is to (1) optimize the vibrational subsystem of interest with DFT while freezing all other degrees of freedom, (2) calculate the harmonic vibrational frequency and intensity for the vibrational subsystem using the same DFT method, then (3) repeat this for many statistically independent snapshots of the reporter in solution.<sup>226</sup> This process was completed for 25 snapshots of CO<sub>2</sub> in IL solution for the asymmetric stretch. DVR asymmetric stretch frequencies and

TDMs were also calculated for the 25 optimized (post step 1) snapshots. In order to facilitate comparison, the square roots of the intensities were taken. The resulting values and the TDMs were divided by their respective gas phase values. These values are plotted against each other in figure 29. The agreement between the two methods is excellent ( $R = 0.994$ ). This new method has the advantage of being essentially computationally free to perform anytime a DVR calculation has already been done. Due to the possibility of parallelization, DVR calculations can be much more computationally inexpensive than regular vibrational frequency calculations.

Table 28: Transition dipole moments for gas phase stretching modes of CO<sub>2</sub>.

Mode	DVR (D)	Experiment (D) <sup>255</sup>
$\omega_s$	$1.1 \times 10^{-13}$	0.0
$\omega_a$	$3.4 \times 10^{-1}$	$3.3 \times 10^{-1}$

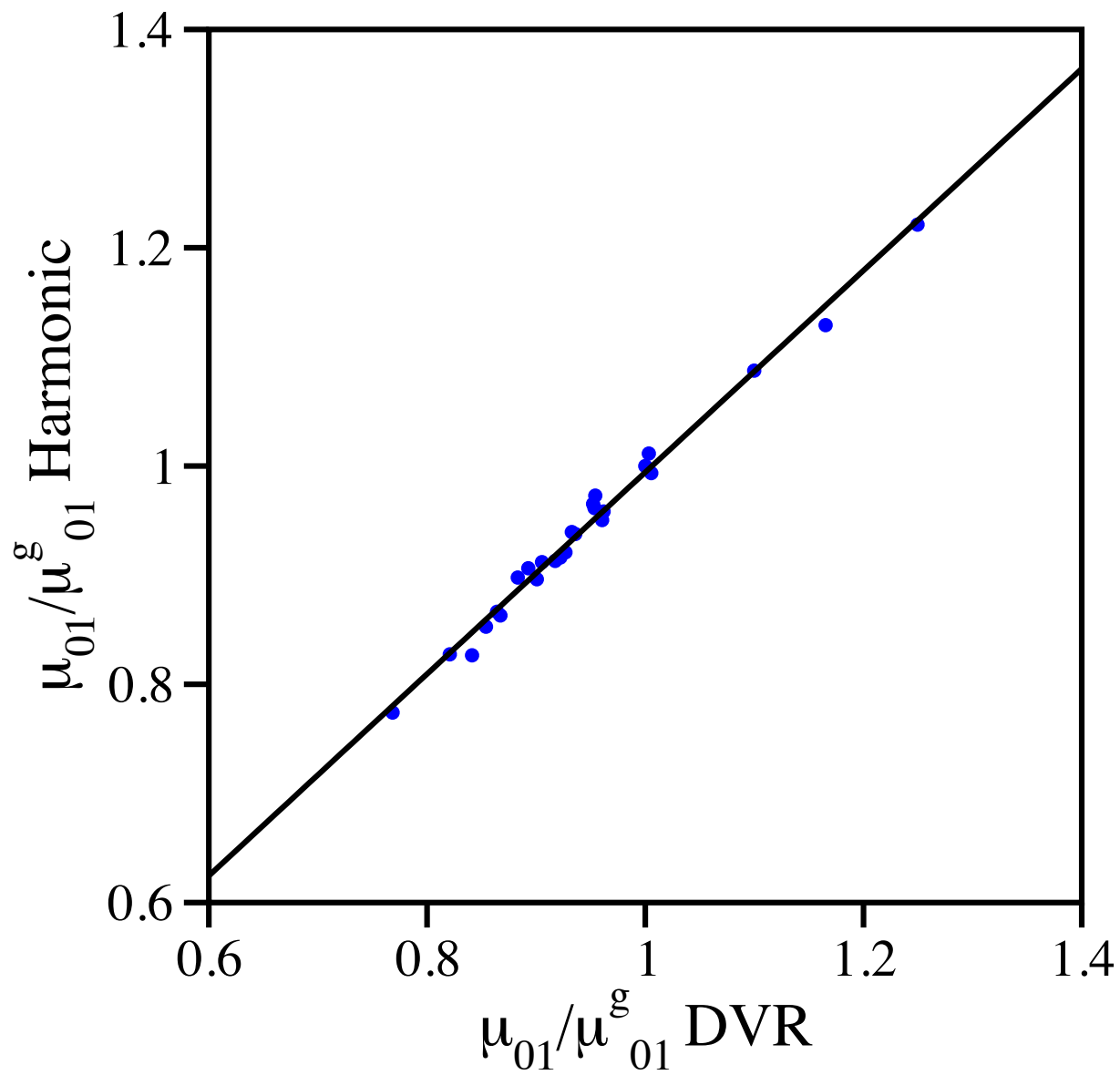


Figure 29: Normalized transition dipole moment for the asymmetric stretch of  $\text{CO}_2$  as calculated by a quantum chemistry program and as calculated by the DVR method (blue dots). The black line is the best fit line,  $y = 0.97x + 0.07$ . The correlation coefficient is 0.994.

## 5.0 A FIRST PRINCIPLES APPROACH FOR PARTITIONING LINEAR RESPONSE PROPERTIES INTO ADDITIVE AND COOPERATIVE CONTRIBUTIONS

The text in this chapter has been adapted from Berquist, E. J.; Lambrecht, D. S. A First Principles Approach for Partitioning Linear Response Properties into Additive and Cooperative Contributions. **2018**, DOI: [10.26434/chemrxiv.5773968.v1](https://doi.org/10.26434/chemrxiv.5773968.v1). The author's contribution to the work included deriving the equations, implementing the algorithm, performing all calculations, and writing the manuscript.

### 5.1 SUMMARY

We present the analytic implementation of linear response equations on top of absolutely localized molecular orbitals (ALMOs) as part of `libresponse`, a library for solving the non-orthogonal molecular response equations for arbitrary operators. In the spirit of the original SCF(MI) and TDDFT(MI) formulations, our method is called linear response for molecular interactions, or LR(MI). Charge transfer was discovered to play an equally significant role in both the ground-state and response iterations.

### 5.2 INTRODUCTION

Experimental spectroscopy is a powerful tool for understanding the structure and function of molecular systems, especially when combined with theory and computation.<sup>2,4,257</sup> Absolutely

localized molecular orbitals (ALMOs)<sup>100</sup> and their associated energy decomposition analysis (ALMO-EDA)<sup>102</sup> provide a intuitive picture of how interactions at the microscopic scale translate to physical insight at the macroscopic scale. ALMO-EDA is capable of separating the total interaction energy of two or more user-defined arbitrary fragments into the following components:

$$\Delta E_{\text{int}} = \Delta E_{\text{gd}} + \Delta E_{\text{frz}} + \Delta E_{\text{pol}} + \Delta E_{\text{CT}}, \quad (5.1)$$

where

- $\Delta E_{\text{gd}}$  is the energy raising due to the distortion from each fragment’s isolated geometry into the cluster geometry;
- $\Delta E_{\text{frz}}$  is the “frozen density” interaction, which accounts for classical electrostatics and Pauli repulsion between the fragments, and is usually positive;
- $\Delta E_{\text{pol}}$  is the polarization energy, which is the self-consistent response of each fragment’s electron density in the field of the other fragments, and is always negative;
- $\Delta E_{\text{CT}}$  is the charge transfer interaction/energy, and is always negative.

This separation is achieved via a constrained self-consistent field calculation, called SCF for molecular interactions (SCF(MI)), producing ALMOs that incorporate fragment polarization but not charge transfer. There are multiple approximations to  $\Delta E_{\text{CT}}$ , which can be further broken apart into a perturbative correction and higher-order effects stemming from a self-consistent treatment, where all orbital constraints are lifted and supersystem SCF iterations are performed (equation 5 in ref. [103]):

$$\Delta E_{\text{CT}} = \Delta E_{\text{CT}}^{\text{RS}} + \Delta E_{\text{CT}}^{\text{HO}}. \quad (5.2)$$

Each of these terms can be broken apart further into a sum of a true electron delocalization term and a BSSE term:

$$\begin{aligned} \Delta E_{\text{CT}}^{\text{RS}} &= \Delta E_{\text{del}}^{\text{RS}} + \Delta E_{\text{BSSE}}^{\text{RS}}, \\ \Delta E_{\text{CT}}^{\text{HO}} &= \Delta E_{\text{del}}^{\text{HO}} + \Delta E_{\text{BSSE}}^{\text{HO}}. \end{aligned} \quad (5.3)$$

Such an energy decomposition is made possible due to the bottom-up construction of ALMOs, where each ALMO is formed from atomic orbitals (AOs) only on a specific fragment,

leading to MOs that are spatially localized onto fragments. In a previous work,<sup>27</sup> we showed that a similar decomposition is possible for peaks in vibrational spectra according to

$$\begin{aligned}\omega_{\text{tot}} &= \omega_{\text{free}} + \Delta\omega_{\text{int}}, \\ \Delta\omega_{\text{int}} &= \Delta\omega_{\text{gd}} + \Delta\omega_{\text{frz}} + \Delta\omega_{\text{pol}} + \Delta\omega_{\text{CT}},\end{aligned}\tag{5.4}$$

where each term has an analogous meaning to the original ALMO-EDA, except changes in the vibrational frequencies  $\{\omega\}$  rather than the interaction energy are used. We seek a more general decomposition of response properties, starting with those based on linear response (LR), where each  $\omega$  in eq. (5.4) is replaced with  $\langle\langle\hat{P};\hat{Q}\rangle\rangle_{\omega_Q}$ , representing the influence of a perturbation  $\hat{Q}$  with frequency  $\omega_Q$  on  $\hat{P}$ .

This is not the first work to seek a fragment- or subsystem-based decomposition of molecular response properties. The LoProp method<sup>258,259</sup> starts from canonical orbitals and performs a series of localizations followed by orthogonalizations, leading to atom- and atom pair-centered contributions. While a top-down approach with only post-SCF localization avoids difficulties with a localized SCF implementation, the use of localized charges and dipole moments requires a finite difference approach beyond first derivatives. Several bottom-up approaches, where localized orbitals or subsystems are formed during the SCF iterations, also exist. More closely related to a many-body expansion (MBE) is the fragment molecular orbital (FMO)<sup>260</sup> calculation of frequency-dependent polarizabilities. The use of MBE-derived expressions requires a response calculation for each term in the fragment expansion, where ALMO-based LR solves the equations for all fragments simultaneously. Response equations have also been formulated within subsystem density functional theory (sDFT).<sup>261,262</sup> sDFT requires the use of an additional kinetic energy functional, while ALMO-based LR can be used with all common exchange-correlation functionals. A comprehensive adaptation of response equations up to dynamic first and second hyperpolarizabilities<sup>263</sup> has been performed with nonorthogonal localized molecular orbitals (NOLMOs).<sup>264</sup> NOLMOs are not connected to an EDA, preventing a property decomposition into EDA-like terms that are the foundation of ALMO-based LR. An attractive fragment-based method designed specifically for the decomposition of molecular interaction energies is symmetry-adapted perturbation theory (SAPT),<sup>265</sup> which is systematically improvable due to its foundations

in perturbation theory. However, because the fragment orbitals are not variationally optimized together (unlike  $\Delta E_{\text{pol}}^{\text{ALMO-EDA}}$ ), the formulation of derivatives even at the uncorrelated monomer approximation (SAPT0) is nontrivial.

The most related methods are those also built on top of ALMOs, specifically ALMO-CIS,<sup>266</sup> ALMO-CIS+CT,<sup>267</sup> TDDFT(MI),<sup>268</sup> and LEA-TDDFT.<sup>269</sup> Compared to ALMO-CIS, our method is more consistent with the index restrictions in TDDFT(MI), and similarly to LEA-TDDFT, we can select individual fragments for allowing charge transfer. For this reason, our method is termed linear response for molecular interactions, or LR(MI).

To this end, we have developed `libresponse`, a molecular response library for non-orthogonal orbitals and arbitrary operators, in the spirit of the `RESPONSE` module of `DALTON`.<sup>270</sup> It is capable of both singlet<sup>271</sup> and triplet<sup>272</sup> response for restricted and unrestricted wavefunctions using Hartree–Fock (HF) or density functional theory (DFT). Properties can be calculated using either the full time-dependent Hartree–Fock (TDHF) or random phase approximation (RPA) equations, or with the configuration interaction with singles (CIS) approximation.<sup>273</sup> Our current implementation is limited to linear response and static properties ( $\omega = 0$ ), with both non-linear response and dynamic properties under current development.

### 5.3 THEORY

For the remainder of the paper, unless otherwise stated, the indices  $i, j, k, l, \dots$  correspond to unoccupied MOs,  $a, b, c, d, \dots$  correspond to virtual/unoccupied MOs,  $\mu, \nu, \lambda, \sigma, \dots$  correspond to AOs, and  $I, J, \dots$  correspond to fragments. There is no distinction between canonical MO and ALMO indices. Comma-separated indices such as  $ia, jb$  correspond to a matrix with a compound index  $ia$  and another compound index  $jb$ .



### 5.3.1 Linear response formalism

The linear response of a molecule in its ground state  $|0\rangle$  to perturbations corresponding to operators  $\hat{P}$  and  $\hat{Q}$  can be written as the sum over all excited states  $\{n\}$ ,

$$\langle\langle\hat{P};\hat{Q}\rangle\rangle_\omega = -\sum_{n>0} \left[ \frac{\langle 0|\hat{P}|n\rangle \langle n|\hat{Q}|0\rangle}{\omega_n - \omega} + \frac{\langle 0|\hat{Q}|n\rangle \langle n|\hat{P}|0\rangle}{\omega_n + \omega} \right], \quad (5.5)$$

where  $\omega$  is the frequency of the perturbation  $\hat{Q}$  and  $\omega_n = E_n - E_0$ . Examples of second-order response properties include the polarizability ( $\hat{P} = \hat{Q} = \hat{\mu}$ ), magnetizability ( $\hat{P} = \hat{Q} = \hat{m}$ ), vibrational frequencies ( $\hat{P} = \hat{Q} = \hat{x}$ ), etc. This paper focuses on molecular response within the static limit ( $\omega = 0$ ),

$$\langle\langle\hat{P};\hat{Q}\rangle\rangle_0 = -\text{Tr} \{ \mathbf{P}^\dagger \mathbf{G}^{-1} \mathbf{Q} \} = -\text{Tr} \{ \mathbf{P}^\dagger \mathbf{X} \}, \quad (5.6)$$

where  $(\mathbf{P})_{ia} = \langle i|\hat{P}|a\rangle$  and  $(\mathbf{Q})_{ia} = \langle i|\hat{Q}|a\rangle$ .  $\mathbf{X} = \mathbf{G}^{-1}\mathbf{Q}$  is the ‘‘response vector’’ corresponding to orbital rotations induced by perturbation  $\hat{Q}$ . The orbital Hessian matrix  $\mathbf{G}$  is defined as

$$\begin{aligned} {}^{RR}\mathbf{G}^\sigma &= \mathbf{A}^\sigma + \mathbf{B}^\sigma \\ {}^{II}\mathbf{G}^\sigma &= \mathbf{A}^\sigma - \mathbf{B}^\sigma \end{aligned} \quad (5.7)$$

for real ( $R$ ) and imaginary ( $I$ ) perturbations, respectively.<sup>274</sup> The spin case  $\sigma$  can either be singlet ( $\sigma = s$ ) or triplet ( $\sigma = t$ ) depending on whether the perturbation operator conserves or flips spin, respectively:

$$A_{ia,jb}^s = \Delta_{ia} + 2(ia|jb) - (ij|ab) \quad (5.8)$$

$$A_{ia,jb}^t = \Delta_{ia} - (ij|ab)$$

$$B_{ia,jb}^s = 2(ia|jb) - (ib|ja) \quad (5.9)$$

$$B_{ia,jb}^t = -(ib|ja)$$

where  $\Delta_{ia} = (\varepsilon_a - \varepsilon_i)\delta_{ia,jb}$  is an orbital energy difference. The CIS or Tamm-Dancoff approximation (TDA) is recovered when  $\mathbf{B} = \mathbf{0}$ . Rather than directly inverting the orbital

Hessian, which would cost  $\mathcal{O}(o^3v^3)$ , one solves iteratively for  $\mathbf{X}$  by expanding the inverse, for example in the real singlet case

$$X_{ia}^{(n+1)} = \frac{Q_{ia} - [4(ia|jb) - (ij|ab) - (ib|ja)] X_{jb}^{(n)}}{\Delta_{ia}} \quad (5.10)$$

Repeated indices imply summation (Einstein sum convention). Here, the superscript in parentheses denotes the iteration number. To avoid the expensive transformation of two-electron repulsion integrals from atomic orbital into molecular orbital basis, one performs the contraction over orbital pair  $jb$  in the atomic orbital basis,

$$(\mathbf{A}^s + \mathbf{B}^s)_{ia,jb} X_{jb} = C_{\mu i} [4(\mu\nu|\lambda\sigma)D_{\lambda\sigma}^X - (\mu\lambda|\nu\sigma)D_{\lambda\sigma}^X - (\mu\sigma|\lambda\nu)D_{\lambda\sigma}^X] C_{\nu a} \quad (5.11)$$

with the perturbed density matrix  $D_{\mu\nu}^X = C_{\lambda j} X_{jb} C_{\sigma b}$ .

To work with *non-orthogonal orbitals*, we replace the  $\Delta_{ia}^{-1}$  denominator in eq. (5.10) with a full matrix inversion of

$$E_{ia,jb} = F_{ab}S_{ij} - F_{ij}S_{ab}, \quad (5.12)$$

where  $\mathbf{F}$  and  $\mathbf{S}$  are the MO-basis Fock and overlap matrix, respectively

### 5.3.2 Adaptation of linear response equations for absolutely localized MOs

The first requirement for adapting the non-orthogonal linear response equations to the ALMO formalism is the projection of the occupied subspace from the virtual subspace. This is because the original fragment-local virtual ALMOs are not orthogonal to occupied ALMOs on other fragments. We form “modified” or “projected” virtual orbitals,

$$|\phi_a\rangle = N_a (1 - \hat{D}_{\text{occ}}) |\psi_a\rangle \quad (5.13)$$

where  $\hat{D}_{\text{occ}} \rightarrow (\mathbf{D}_{\text{occ}})_{\mu\nu} = C_{\mu i} C_{\nu i}$ . Otherwise, this contamination would manifest as spurious low-lying excited states, leading to artificial poles in the polarizability. This requirement for projection is unimportant for ground-state SCF(MI) calculations, where the virtual orbitals do not have an effect on the final result (energy).

The second requirement is to maintain consistency with the fragment-local nature of SCF(MI), where CT is not allowed. Each occupied-virtual MO contribution to the molecular response is restricted to be within the same fragment. This corresponds to forcing all  $ia$  (or  $jb$ ) pairs to be zero when  $i \in F_I, a \in (F_J \neq F_I)$ . Algorithmically, we do this by always working in the full CT-allowed basis, then zeroing out CT-disallowed matrix elements. For the energy denominator, the zero rows and columns must be removed to avoid introducing artificial singularities. The advantage of working in the full CT basis and zeroing matrix elements is twofold: it prevents code duplication due to no need for separate non-orthogonal and ALMO-based response routines, and it enables selection of individual fragments to quantify their CT contribution to the final response property. This is in contrast to ALMO-CIS+CT,<sup>267</sup> which starts from the reduced CT-disallowed basis and adds a distance-dependent CT correction. An outline of the algorithm is available in the SI.

Perhaps the most important fundamental distinction between our formulation of the ALMO response equations and the ALMO-CIS formulation is in the two-electron contribution, where we restrict MO indices but do not restrict AO indices. This means that for a given  $J_{\mu\nu}^X = (\mu\nu|\lambda\sigma)D_{\lambda\sigma}^X$ , summations run over AO indices on all fragments, and both  $(\mu\nu|\lambda\sigma)$  and  $D_{\lambda\sigma}^X$  for  $\lambda \in F_I, \sigma \in (F_J \neq F_I)$  may be non-zero. Additionally, AO indices cannot be restricted due to the projection in eq. (5.13), where each MO can have contributions from all AOs.

In principle, the ALMO formulation presented here can be extended to arbitrary-order response, as long as each compound  $ia$  index is constrained to be within a fragment.

### 5.3.3 Decomposition of linear response properties into local contributions and interaction mechanisms

As with other localized orbital schemes, there is a question of how to define charge transfer. Our implementation allows for the introduction of CT in several stages. This allows for further decomposition of the CT term into qualitative contributions from individual fragments. This is done by starting from the polarized SCF(MI) wavefunction and allowing transitions between the occupied space of a single fragment into the virtual space of all fragments. This

can be denoted as “frz + pol + CT( $F \rightarrow$  all)”, where  $F$  is an individual fragment. The difference between this result and “frz + pol” can be viewed as how important CT is for that specific fragment  $F$ . From the symmetry-adapted perturbation theory<sup>275</sup> perspective, it can also be viewed as a BSSE correction; the larger the difference, the more deficient the fragment-localized basis is. CT can also be allowed during the response calculation even if it was not during the wavefunction calculation; this corresponds to performing non-orthogonal response with no restrictions using the polarized SCF(MI) wavefunction, represented as “frz + pol + CT(all  $\rightarrow$  all) [blocked]”. The sum over all fragments should be qualitatively equal to “frz + pol + CT(all  $\rightarrow$  all) [blocked]”, though it will not be exactly equal due to higher-order effects, similar to the  $\Delta E_{\text{CT}}^{\text{HO}}$  term in ALMO-EDA. Finally, “frz + pol + CT(all  $\rightarrow$  all) [blocked]” corresponds to an unrestricted response calculation on top of the fully-relaxed SCF wavefunction, analogous to a supersystem calculation. The difference between the blocked result and the supersystem result shows the effect of allowing CT in the underlying wavefunction.

This multi-step decomposition of the CT term is similar to our previous work on the effect of CT on the vibrational spectra of CO<sub>2</sub> in ionic liquids.<sup>27,276</sup> Being able to turn CT on and off during the geometry optimization and numerical Hessian calculation allowed us to see how CT contributes to the potential energy surface at multiple stages, but our analytic formulation of the response equations has significant advantages. It enables the identification of fragment-specific contributions to the total response, which is not possible in a numerical response calculation, even with an underlying SCF(MI) wavefunction. This is in addition to all the other advantages of solving analytic rather than numerical equations, namely calculation time, uncoupled and perturbative approximations to fully iterative results, lack of finite-difference error, frequency-dependent perturbations, and response to applied magnetic fields without complex energies; see ref. [1] for a discussion.

Borrowing terminology from ref. [276], CT may enter the response calculation through multiple mechanisms. We consider two: the “wavefunction mechanism”, where CT is (dis)-allowed during the SCF iterations, corresponding to either an SCF(MI) or a canonical SCF calculation, and the “response mechanism”, where CT is (dis)allowed during the response calculation by restricting occupied-virtual MO contributions to be within fragments as dis-

cussed in section 5.3.2. This results in four principal permutations of (dis)allowing CT effects, abbreviated as “ $a/b$ ”, where  $a$  represents the wavefunction mechanism,  $b$  represents the response mechanism, and  $a, b \in \{\text{on, off}\}$ :

- “on/on” is a standard response calculation, and is the “frz + pol + CT(all  $\rightarrow$  all) [super]” result.
- “off/off” is the most consistent with SCF(MI) and TDDFT(MI), and is the “frz + pol” result.
- “off/on” attempts to recover CT effects during the response calculation that are missing from the underlying SCF(MI) wavefunction, and is the “frz + pol + CT(all  $\rightarrow$  all) [blocked]” result. This is not the same as the single perturbative Roothaan step (RS) correction in SCF(MI), but is a fully-iterative response calculation on top of ALMOs.
- “on/off” corresponds to allowing CT during SCF iterations but not the response calculations. This decomposition is not considered, as neither the occupied nor the virtual canonical orbitals can generally be assigned to specific fragments.

For the purposes of this work, we also do not consider the effect of disallowing CT during geometry optimization through the “geometry mechanism”, though in principle it is technically feasible and potentially non-negligible. See ref. [277] for an in-depth analysis of the different potential energy surfaces that comprise ALMO-EDA. Neglecting the geometry mechanism means our approach is consistent with a vertical decomposition rather than an adiabatic one.

## 5.4 METHODS

HF was used for all calculations as a proof of concept to avoid spurious overdelocalization and CT effects due to self-interaction error. As a result, it should provide a lower bound on any possible CT effects during the SCF(MI) calculation. For basis sets, we chose the Karlsruhe def2- family,<sup>278,279</sup> in particular def2-SVP and def2-SVPD. It is well-known that the original ALMO procedure has a strong basis-set dependence, with overestimation of  $\Delta E_{\text{pol}}$

and underestimation of  $\Delta E_{\text{CT}}$  due to the overlapping tails of AOs on different fragments, allowing electron density to “leak” across fragments during SCF(MI).<sup>280</sup> As a result, there is no clear basis set limit to  $\Delta E_{\text{pol}}$ . This informs our basis set choices, which should give qualitatively correct results within the original ALMO definition, with a check on minimal augmentation optimized for use in electric field response calculations.

For a test property, we chose the static dipole polarizability, defined as

$$\alpha_{ab}(0) = - \langle \langle \hat{\mu}_a; \hat{\mu}_b \rangle \rangle_0 = - \left. \frac{\partial^2 E}{\partial \varepsilon_a \partial \varepsilon_b} \right|_{\varepsilon_a, \varepsilon_b=0}, \quad (5.14)$$

where  $a, b \in \{x, y, z\}$  and the energy derivative method based on derivatives with respect to applied electric fields  $\varepsilon$  is equivalent to the linear response calculation.

For a test system, the argon—lithium cation dimer,  $\text{Ar}\cdots\text{Li}^+$ , was chosen due to the reasonable polarizability of the bare (closed-shell) argon atom, which should be quantitatively influenced by the almost unscreened nuclear charge of the lithium cation. The optimized distances are 2.4106 Å using def2-SVP and 2.4297 Å using def2-SVPD. For the distance dependence results, points are spaced at 0.05 Å from 1.25 to 5.00 Å, then spaced at 0.25 Å from 5.00 to 10.00 Å.

All calculations were performed using a development version of Q-CHEM<sup>93</sup> compiled with `libresponse`, our generalized non-orthogonal molecular response library. Thresholds were set to  $10^{-14}$  for integral screening,  $10^{-11}$  for the DIIS error norm in SCF convergence, and  $10^{-8}$  for the DIIS error norm in the analytic response iterations. The Stoll projector equations were used for minimizing the underlying SCF(MI) wavefunction.<sup>100,281</sup> For numerical polarizability calculations, 2nd-order finite difference from energies was used, unless otherwise stated, as differences between 1st- and 2nd-order results were below  $10^{-7} a_0^3$  in all cases. Both 1st- and 2nd-order finite difference polarizability calculations used a finite field step size of  $1.88973 \times 10^{-5} a_0^3$ . All SCF(MI) analytic polarizability calculations use projected virtual orbitals as defined by eq. (5.13).

Table 29: Polarizability results for the argon—lithium cation dimer at the HF/def2-SVP level. The geometry is optimized at the same level. All values have units of  $a_0^3$ .

def2-SVP	analytical			numerical		
	$\alpha_{\perp}$	$\alpha_{\parallel}$	$\alpha_{\text{iso}}$	$\alpha_{\perp}$	$\alpha_{\parallel}$	$\alpha_{\text{iso}}$
monomer A (Ar)	4.3702	4.3702	4.3702	4.3702	4.3702	4.3702
monomer B (Li <sup>+</sup> )	0.1593	0.1593	0.1593	0.1593	0.1593	0.1593
A + B	4.5295	4.5295	4.5295	4.5295	4.5295	4.5295
frz + pol	4.5042	4.5581	4.5222	4.5036	4.5696	4.5256
+ CT(A → all)	4.7040	5.5191	4.9757	—	—	—
+ CT(B → all)	0.1594	0.1586	0.1591	—	—	—
+ CT(all → all) [blocked]	4.8409	5.7540	5.1452	—	—	—
+ CT(all → all) [super]	5.2142	7.8649	6.0978	5.2142	7.8649	6.0978

## 5.5 RESULTS AND DISCUSSION

### 5.5.1 Equilibrium distance

Tables 29 and 30 show equilibrium distance results for the argon—lithium cation dimer at the HF/def2-SVP and HF/def2-SVPD levels, respectively. Due to molecular symmetry, two of the three polarizability tensor principal components are identical; these correspond to the perpendicular polarizability  $\alpha_{\perp}$ , or how easy it is to shift the electron density perpendicular to the interatomic axis. The parallel polarizability  $\alpha_{\parallel}$  corresponds to the principal component that lies along the axis of interaction between the two atoms and corresponds to the bond polarizability. Monomer A and B refer to the isolated atoms; parallel and perpendicular axes have no meaning here due to their spherical symmetry, as all principal components are identical. An approximation to the frozen density interaction in response is the sum of the response of the isolated fragments (atoms), shown as “A + B”. Note that this is not the true frozen monomer response.<sup>277</sup> “frz + pol” is where CT is disallowed between fragments

Table 30: Polarizability results for the argon—lithium cation dimer at the HF/def2-SVPD level. The geometry is optimized at the same level. All values have units of  $a_0^3$ .

def2-SVPD	analytical			numerical		
	$\alpha_{\perp}$	$\alpha_{\parallel}$	$\alpha_{\text{iso}}$	$\alpha_{\perp}$	$\alpha_{\parallel}$	$\alpha_{\text{iso}}$
monomer A (Ar)	10.3735	10.3735	10.3735	10.3735	10.3735	10.3735
monomer B (Li <sup>+</sup> )	0.1603	0.1603	0.1603	0.1603	0.1603	0.1603
A + B	10.5338	10.5338	10.5338	10.5338	10.5338	10.5338
frz + pol	10.0787	10.6399	10.2658	10.0875	10.6599	10.2783
+ CT(A → all)	10.0157	10.4900	10.1738	—	—	—
+ CT(B → all)	0.1606	0.1614	0.1609	—	—	—
+ CT(all → all) [blocked]	10.1395	10.8004	10.3598	—	—	—
+ CT(all → all) [super]	10.0724	11.0645	10.4031	10.0724	11.0645	10.4031

using the algorithm described above, consistent with the original SCF(MI) and conceptually analogous to TDDFT(MI). The remaining entries are those described in section 5.3.3.

**5.5.1.1 Validation of analytic ALMO results** As a correctness check, our analytic formulation is validated against finite-field calculations in the last three columns. An advantage of the analytic formulation is the ability to further decompose the CT term; it is not possible to perform the “CT(A → all)”, “CT(B → all)”, and “CT(all → all) [blocked]” calculations with a finite-field approach. For both basis sets presented, there is no quantitative difference between the analytic and numerical polarizabilities except for “frz + pol”. This should not be interpreted as an error, since our definition of the response equations differs from the no-CT formulation presented by ALMO-CIS, and there is a slight delocalization of the virtual space due to the projection (see section 5.6). Still, the largest percent difference is 0.25% and 0.19% in  $\alpha_{\parallel}$  for def2-SVP and def2-SVPD, respectively, showing our definition should be valid as long as the ALMO approximation holds.



Table 31: Percentage changes in  $\alpha_{\parallel}$  due to the two CT mechanisms. The reference (denominator) polarizability used is shown in *italics*. All calculations used a bond length of 2.4297 Å.

percentage change due to mechanism	basis set	
	def2-SVP	def2-SVPD
response ( <i>off/on</i> - <i>off/off</i> )	20.8	1.5
wavefunction ( <i>on/on</i> - <i>off/on</i> )	26.8	2.4

**5.5.1.2 CT mechanism analysis** From tables 29 and 30, the importance of charge transfer via both the wavefunction and response mechanisms can be calculated as the difference between “frz + pol”, “CT(all  $\rightarrow$  all) [blocked]”, and “CT(all  $\rightarrow$  all) [super]”; these are “off/off”, “off/on”, and “on/on”, respectively. The effect of each mechanism is shown in table 31.

For def2-SVP, it is clear that each successive relaxation of CT restrictions has a non-negligible effect on the final property, where  $\alpha_{\parallel}^{\text{frz} + \text{pol}}$  is 58.0% and 96.2% of the supermolecular value for the two basis sets. This representation makes it clear that in addition to larger polarizabilities ( $\sim 2\times$ ), there are other fundamental differences caused by the use of diffuse functions in def2-SVPD.

Calculated this way, the wavefunction mechanism may include some higher-order effects, as the pure wavefunction mechanism, calculated as “on/off” - “off/off”, cannot be isolated due to the “on/off” term (see section 5.3.3). The size of these higher-order effects in the wavefunction can be estimated by the percent contribution of higher-order effects in ALMO-EDA charge transfer, shown in table 32.

Using the uncorrected results for consistency, the new estimates of percentage changes due to the wavefunction mechanism become 20.3% and 2.0%, showing the presence of CT in the underlying wavefunction and in the response calculation are equally important in predicting molecular properties.

Table 32: Percentage of ALMO-EDA charge transfer due to higher-order effects, calculated as  $100 * \frac{\Delta E_{CT}^{HO}}{\Delta E_{CT}^{RS} + \Delta E_{CT}^{HO}}$ . All calculations used a bond length of 2.4297 Å.

BSSE corrected?	basis set	
	def2-SVP	def2-SVPD
no	24.2	17.1
yes	25.0	19.4

**5.5.1.3 Basis set dependence** The results in table 31 show that charge transfer, or electron delocalization across fragments, is relatively unimportant when using the diffuse basis set. This may be indicative of either basis set incompleteness or overlap between the tails of the diffuse basis functions (see section 5.4). The former argument is supported by the BSSE term being 50% of the total CT term with def2-SVP but only 18% with def2-SVPD, while the latter argument is supported by a doubling of  $\Delta E_{pol}$  from  $-3.3$  to  $-6.6$  kcal mol $^{-1}$  when adding the diffuse functions (see the SI for all ALMO-EDA results).

To investigate these effects directly on the polarizability, rather than through an EDA, table 33 shows the decoupled effect of the nuclear charge and basis functions provided by the lithium cation on the argon. Since the lithium cation is not very polarizable ( $\alpha^{Li^+}$  is  $\approx 4\%$  of  $\alpha^{Ar}$  at the HF/def2-SVP level), only effects on the argon atom are considered. Also presented are the lowest singlet and triplet excitation energies calculated using RPA (TD-HF). Percentages of each value compared to the full Ar $\cdots$ Li $^+$  result within each basis set are shown in the SI.

In order to test the importance of nuclear charge without the influence of basis functions, the “PC(+)” and “PC(-)” entries represent replacing the lithium atom with positive and negative point charges of unit magnitude, respectively. To test the importance of the basis set without the influence of a nuclear charge, “Gh(Li)” represents replacing the full lithium atom with only its basis functions, leaving a “ghost atom” with no (effective) nuclear charge. This should separate the Ar $\cdots$ Li $^+$  interaction into a purely electrostatic effect from the

nucleus, a delocalization-like effect due to increased spatial flexibility for motion of electrons during SCF and response iterations, and some non-additive effects.

As expected, compared to a bare argon atom, the presence of a ghost function both increases the polarizability and decreases the first excitation energy. Intuitively, there are now additional molecular orbitals within reasonable spatial distance that electron density can spread on to as part of the ground-state calculation and can be excited to as part of the excited-state calculation. The difference is much less pronounced for the diffuse basis, where it is the presence of point charges that has the larger effect. This indicates that there is overlap with the tails of only the additional diffuse basis functions at the distance of the lithium center, and is caused by basis set incompleteness present even at the canonical SCF level. The point charges are in comparatively empty space in the non-diffuse basis, and the ghost function is unimportant for the diffuse basis. There,  $\alpha_{\parallel}^{\text{PC}(+)} \gg \alpha_{\parallel}^{\text{Ar}\cdots\text{Li}^+}$  due to electron screening on the lithium atom, which leads to a reduced effective nuclear charge that is less capable of polarizing the argon atom.

As further confirmation that any basis set effects are not due to a breakdown in the original ALMO formulation, we performed ALMO-EDA calculations where the polarization term does not suffer from CT contamination and has a well-defined basis set limit.<sup>282</sup> With this second-generation EDA, the CT energy increases by only  $10^{-4}$  kcal mol<sup>-1</sup> in both basis sets. Since there is no mixing of one fragment’s ALMOs into the other due to overlapping AO tails, these differences are true basis set effects as long as the fragment electron densities are not penetrating each other.

### 5.5.2 Distance dependence

An extension of the validation in section 5.5.1.1 is to compare our analytic implementation of “ALMO frz + pol” with finite field calculations using both energies and dipole moments as a function of interatomic distance. Short-range results for both basis sets are presented in figure 30. Only at shorter than equilibrium distances is there visible deviation between the analytic and numerical “ALMO frz + pol” results, most likely due to the virtual space delocalization. However, this is near the region where the electron clouds of the two atoms

Table 33: Decomposition of polarizabilities into point charge and basis function contributions. All calculations used canonical MOs and a distance of 2.4297 Å from argon to the other center(s).

basis set	structure	$\alpha_{\perp}$ (a.u.)	$\alpha_{\parallel}$ (a.u.)	$\bar{\alpha}$ (a.u.)	${}^t E_{0 \rightarrow \text{lowest}}^{\text{RPA}}$ (eV)	${}^s E_{0 \rightarrow \text{lowest}}^{\text{RPA}}$ (eV)
def2-SVP	Ar⋯PC(−)	4.38	4.35	4.37	23.26	25.09
def2-SVP	Ar	4.37	4.37	4.37	23.69	25.50
def2-SVP	Ar⋯PC(+)	4.36	4.38	4.37	23.42	25.24
def2-SVP	Ar⋯Gh(Li)	4.89	6.23	5.34	12.26	12.58
def2-SVP	Ar⋯Li <sup>+</sup>	5.21	7.88	6.10	11.90	12.18
def2-SVPD	Ar⋯PC(−)	10.94	10.33	10.74	10.92	11.68
def2-SVPD	Ar	10.37	10.37	10.37	12.61	13.22
def2-SVPD	Ar⋯PC(+)	10.04	11.94	10.67	9.90	10.48
def2-SVPD	Ar⋯Gh(Li)	10.39	10.42	10.40	12.05	12.41
def2-SVPD	Ar⋯Li <sup>+</sup>	10.07	11.06	10.40	11.98	12.34

penetrate each other, and  $\Delta E_{\text{frz}}$  is an order of magnitude larger than either  $\Delta E_{\text{pol}}$  or  $\Delta E_{\text{CT}}$  due to Pauli repulsion, so this difference does not detract from the accuracy of the analytic implementation.

Figure 31 shows how the different forms of CT restriction vary as a function of interatomic distance. Similar to the counterpoise (CP) correction for interaction energies, the “BSSE-corrected canonical” polarizability for two monomers A and B is defined as

$$\alpha^{\text{BSSE-corrected}} = \alpha^{AB}(AB) - [(\alpha^A(AB) - \alpha^A(A)) + (\alpha^B(AB) - \alpha^B(B))], \quad (5.15)$$

where, for example,  $\alpha^A(AB)$  is the polarizability of monomer A in the combined (supermolecular) basis of both monomers.

The difference from the BSSE correction is large in def2-SVP but negligible in def2-SVPD, confirming our results from section 5.5.1.3. For both basis sets, the argon polarizability is the major contributor, provided that the entire dimer basis virtual space is available. The difference between the argon and blocked ALMO curves is due to the small but non-zero polarizability of the lithium cation, plus a mutual or higher-order polarization effect that appears at shorter than equilibrium distances for def2-SVP. Most notably, there is a quantitative difference between the blocked ALMO and canonical results even with def2-SVPD, showing that the wavefunction mechanism is large, and only allowing CT during the response calculation cannot recover these effects.

## 5.6 CONCLUSIONS AND FUTURE WORK

We have presented an implementation of linear response molecular properties for ALMOs, along with a decomposition of charge transfer effects, applied to a system with a substantial CT interaction. We discovered that for the static polarizability, charge transfer plays an equally important role in the response calculation as it does in the underlying wavefunction. Additionally, our results confirm that the ALMO and LR(MI) approximations are valid as long as the basis set is not deficient for the system, as is the case with def2-SVP.

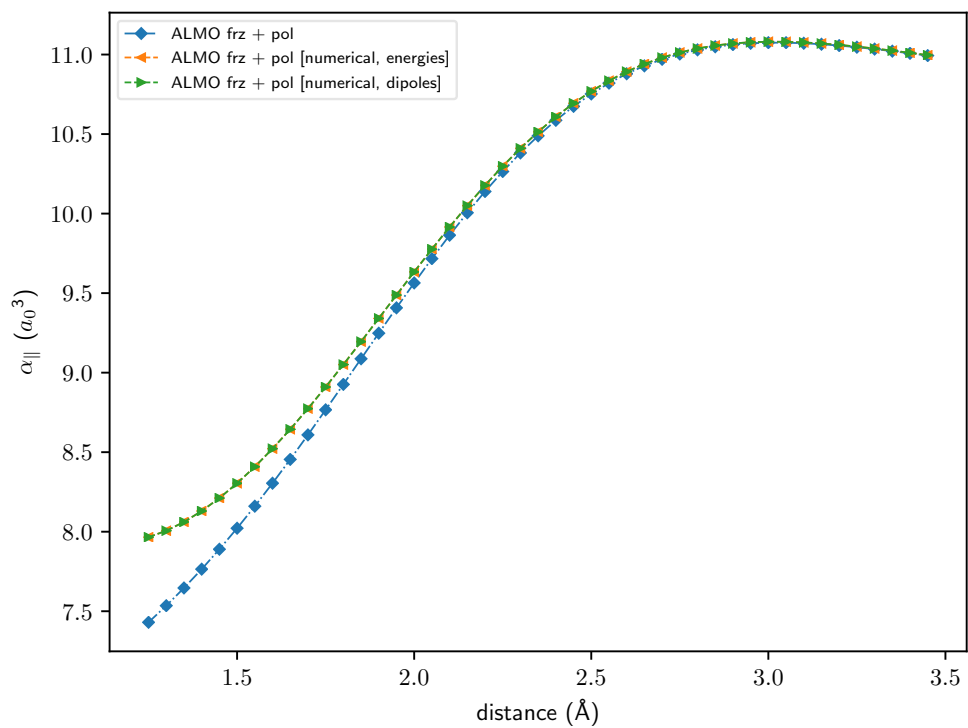
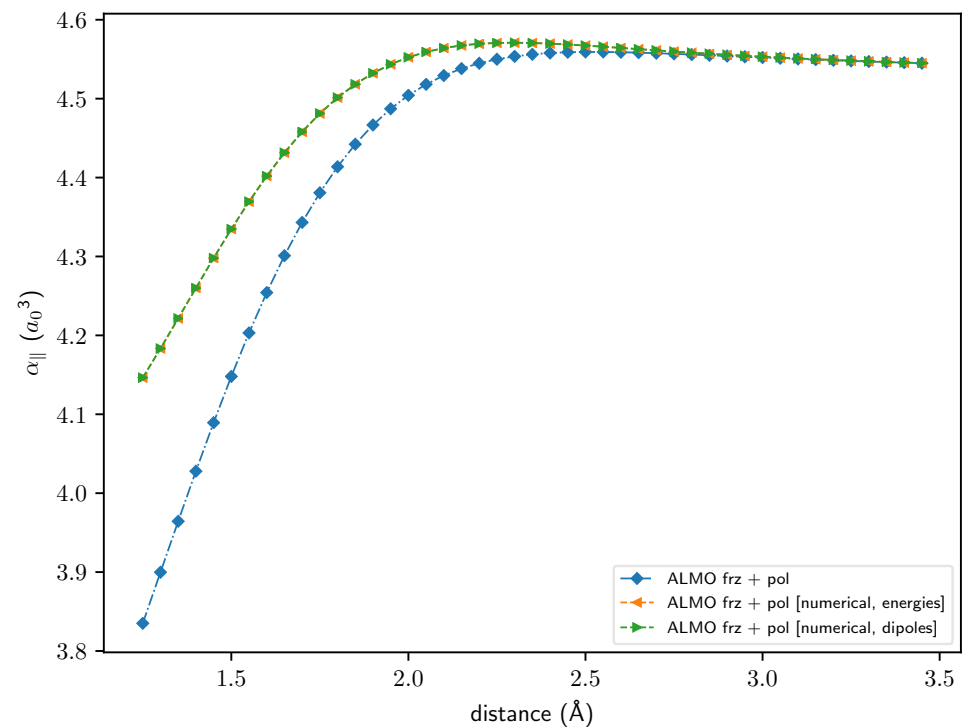


Figure 30: Distance dependence of analytic and numerical ALMO polarizabilities for both basis sets.

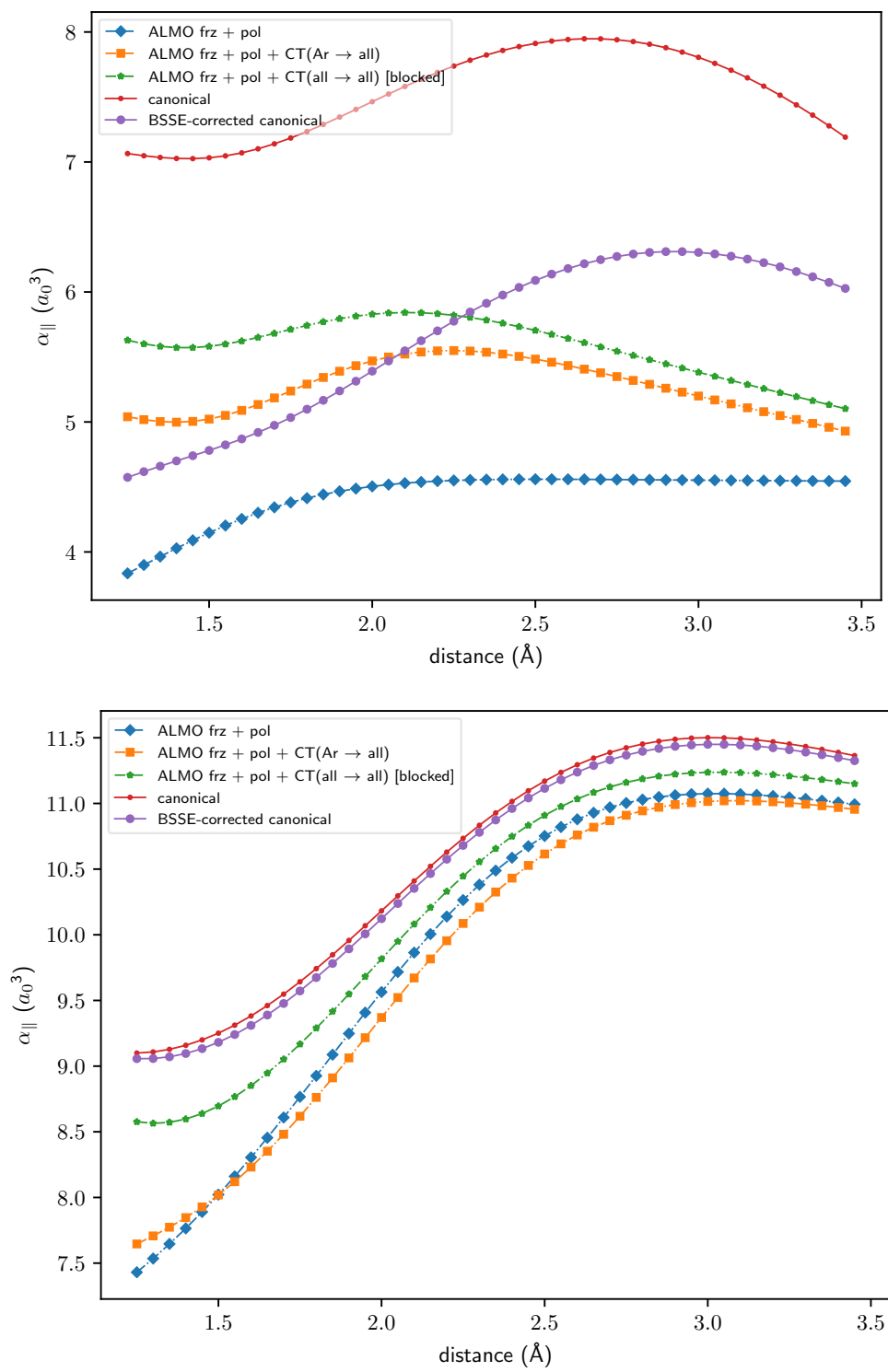


Figure 31: Distance dependence of CT restrictions on polarizabilities for both basis sets.

A question not addressed in this work or other ALMO-based work on excitation energies<sup>266,268</sup> is the effect of non-locality on the projected virtual space. To properly assign fragment-localized contributions to molecular response, both the occupied and virtual ALMOs must be spatially localized to individual fragments. Projecting the occupied space out of the virtual space ensures occupied-virtual orthogonality between fragments, but removes the fragment locality of the virtual space; that is, each virtual MO can no longer be uniquely assigned to a specific fragment. However, we expect that the error introduced by a delocalized virtual space is small compared to the error from using unprojected orbitals, which are less representative of the true potential energy surface for the reasons discussed in section 5.3.2. Future work will use LoProp-type approaches on top of projected ALMOs to investigate the magnitude of these effects. In this way, LR(MI) can be a sensitive test for how modification of the virtual space affects molecular properties.

Our development of a library for calculating molecular response properties of arbitrary operators with non-orthogonal orbitals opens many doors for future development. `libresponse` is available in Q-CHEM 5.0.2 and at <https://github.com/LambrechtLab/libresponse> under the 3-clause BSD license as an Armadillo-based<sup>283</sup> C++ library. It can also be used as a PSI4<sup>284</sup> plugin.

## 5.7 ACKNOWLEDGEMENTS

E.J.B. thanks Q-CHEM for a summer internship opportunity, Evgeny Epifanovsky for help with the initial DIIS implementation, and `cclib`<sup>285,286</sup> for the analysis framework.

## 5.8 SUPPORTING INFORMATION

ALMO-EDA results and analysis, additional ghost basis and point charge analysis, additional distance dependence results, sample input files, pseudocode for algorithm.

As a sanity check for all results, we expect certain physical behavior over the range of



Table 34: ALMO-EDA results. Energy units are kcal mol<sup>-1</sup>. All calculations used Hartree-Fock with a bond length of 2.4297 Å.

ALMO-EDA term	basis set	
	def2-SVP	def2-SVPD
$\Delta E_{\text{frz}}$	1.20	2.18
$\Delta E_{\text{pol}}$	-3.32	-6.61
$\Delta E_{\text{del}}^{\text{RS}}$	-4.07	-0.89
$\Delta E_{\text{BSSE}}^{\text{RS}}$	1.38	0.16
$\Delta E_{\text{CT}}^{\text{RS}}$	-2.69	-0.74
$\Delta E_{\text{int}}^{\text{RS}}$	-4.81	-5.17
$\Delta E_{\text{del}}^{\text{SCF}}$	-5.38	-1.08
$\Delta E_{\text{BSSE}}^{\text{SCF}}$	1.79	0.16
$\Delta E_{\text{CT}}^{\text{SCF}}$	-3.59	-0.91
$\Delta E_{\text{int}}^{\text{SCF}}$	-5.70	-5.35
$\Delta E_{\text{HO}}^{\text{SCF}}$	-0.90	-0.18

Table 35: Analysis of ALMO-EDA terms from table 34.

	basis set	
	def2-SVP	def2-SVPD
$\Delta E_{\text{pol}} + \Delta E_{\text{CT}}^{\text{RS}}$ (kcal/mol)	-6.01	-7.34
$\Delta E_{\text{pol}} + \Delta E_{\text{CT}}^{\text{SCF}}$ (kcal/mol)	-6.91	-7.52
100 * $\Delta E_{\text{BSSE}}^{\text{RS}} / \Delta E_{\text{CT}}^{\text{RS}}$ (%)	-51.4	-21.2
100 * $\Delta E_{\text{BSSE}}^{\text{SCF}} / \Delta E_{\text{CT}}^{\text{SCF}}$ (%)	-49.9	-17.7
100 * $\Delta E_{\text{CT}}^{\text{RS}} / (\Delta E_{\text{pol}} + \Delta E_{\text{CT}}^{\text{RS}})$ (%)	44.8	10.0
100 * $\Delta E_{\text{CT}}^{\text{SCF}} / (\Delta E_{\text{pol}} + \Delta E_{\text{CT}}^{\text{SCF}})$ (%)	51.9	12.2

Table 36: Percentage of supermolecular result for point charge and ghost function polarizabilities. All calculations used Hartree–Fock with canonical MOs and a distance of 2.4297 Å from argon to the other center(s).

basis set	structure	$\alpha_{\perp}$	$\alpha_{\parallel}$	$\bar{\alpha}$	${}^t E_{0 \rightarrow \text{lowest}}^{\text{RPA}}$	${}^s E_{0 \rightarrow \text{lowest}}^{\text{RPA}}$
def2-SVP	Ar...PC(−)	84.1	55.2	71.6	195.5	206.0
def2-SVP	Ar	83.9	55.5	71.6	199.1	209.4
def2-SVP	Ar...PC(+)	83.7	55.6	71.6	196.8	207.2
def2-SVP	Ar...Gh(Li)	93.9	79.1	87.5	103.0	103.3
def2-SVPD	Ar...PC(−)	108.6	93.4	103.3	91.2	94.7
def2-SVPD	Ar	103.0	93.8	99.7	105.3	107.1
def2-SVPD	Ar...PC(+)	99.7	108.0	102.6	82.6	84.9
def2-SVPD	Ar...Gh(Li)	103.2	94.2	100.0	100.6	100.6

interatomic distances. In particular, as the interatomic distance approaches the limit of infinite separation,

- The canonical SCF, numerical SCF(MI), and analytic SCF(MI) results all converge to the same value, which is the sum of polarizabilities the two isolated atoms. CT appears to be an important contributor until approximately 5 Å, where the decay behavior of the canonical SCF changes.
- The point charge, and “frz + pol + CT(Ar → all)” results all converge to the same value, which is the polarizability of the isolated argon atom.

### 5.8.1 Input files

Listing 1: Sample Q-CHEM input file for “ALMO frz + pol” polarizability. Geometry is from HF/def2-SVPD.

```
1 $rem
```

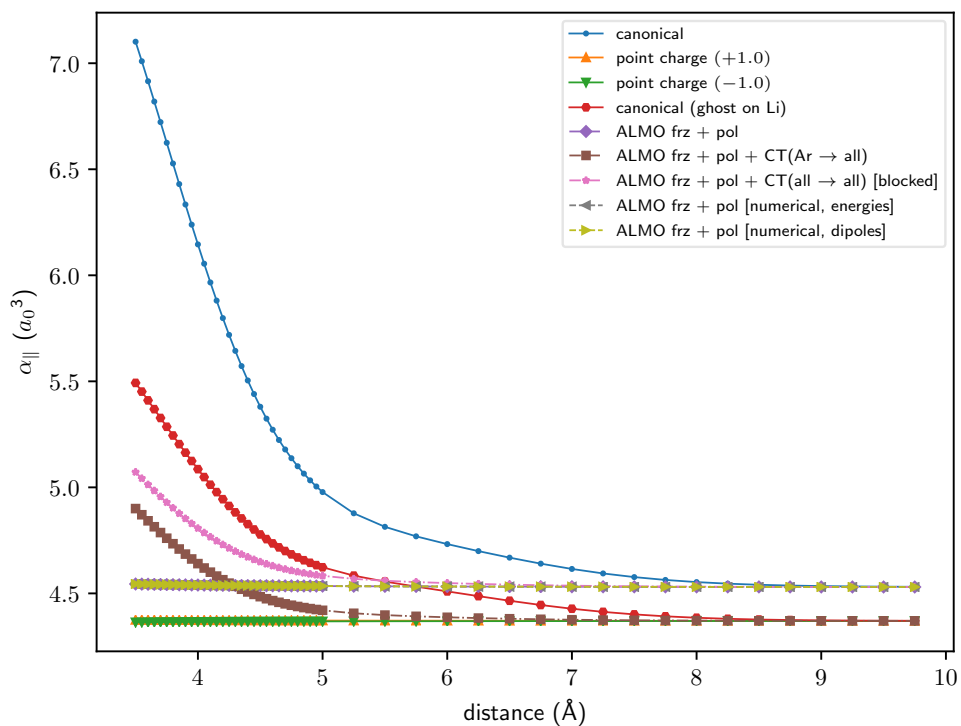
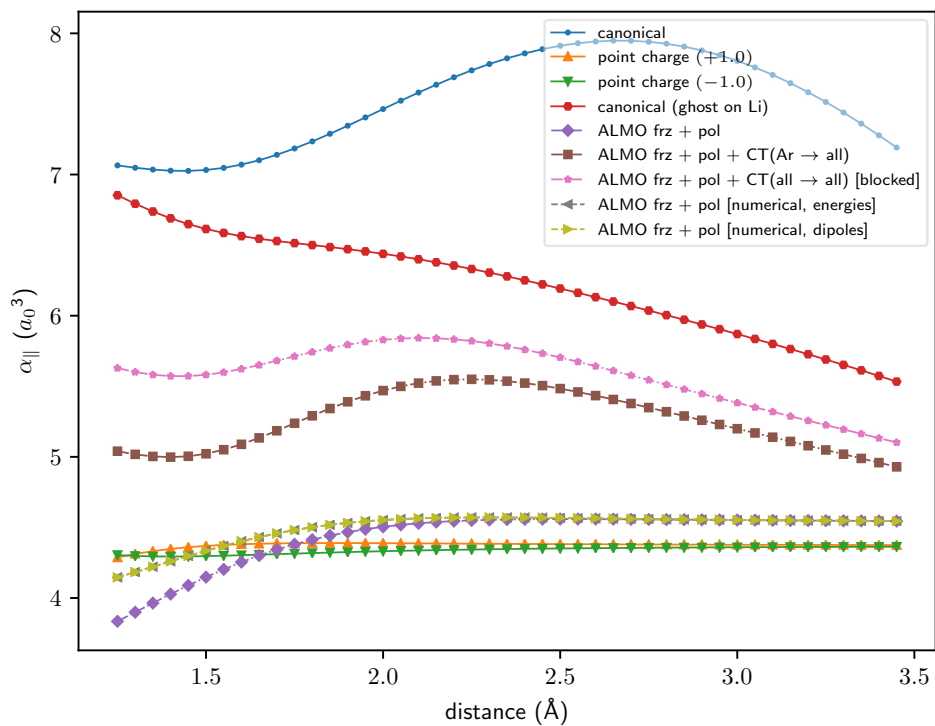


Figure 32: Short- and long-range interatomic separation dependence of the static polarizability parallel to the coordination axis. All calculations are at the HF/def2-SVP level.

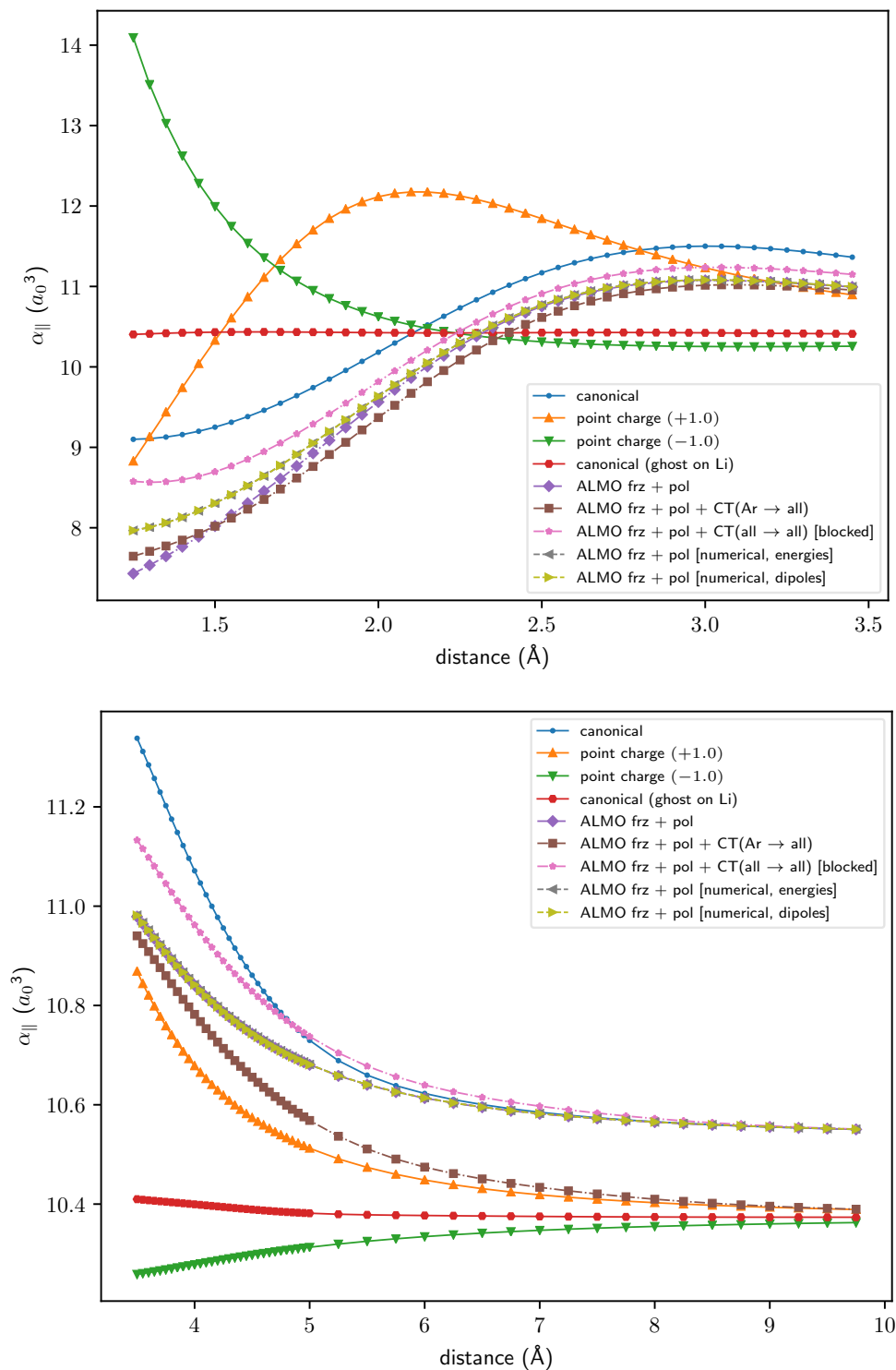


Figure 33: Short- and long-range interatomic separation dependence of the static polarizability parallel to the coordination axis. All calculations are at the HF/def2-SVPD level.

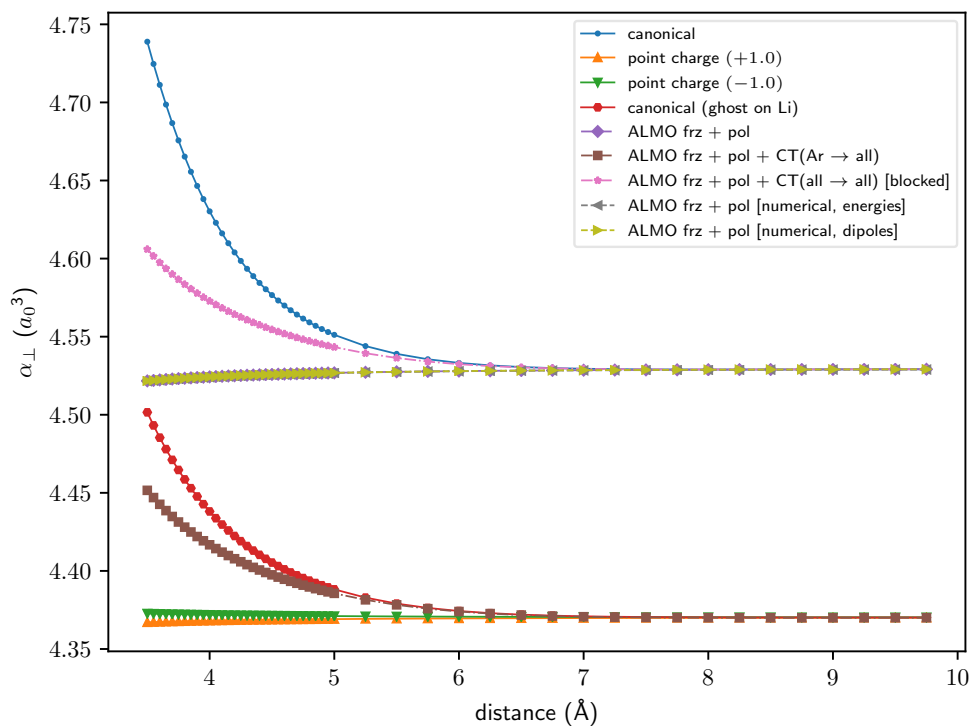
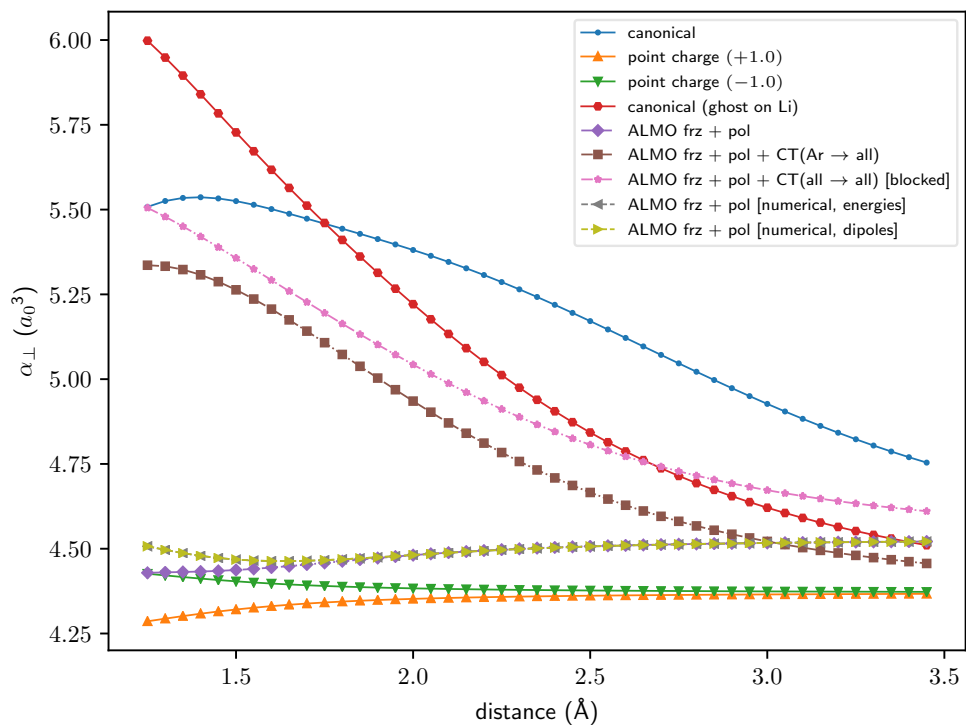


Figure 34: Short- and long-range interatomic separation dependence of the static polarizability perpendicular to the coordination axis. All calculations are at the HF/def2-SVP level.

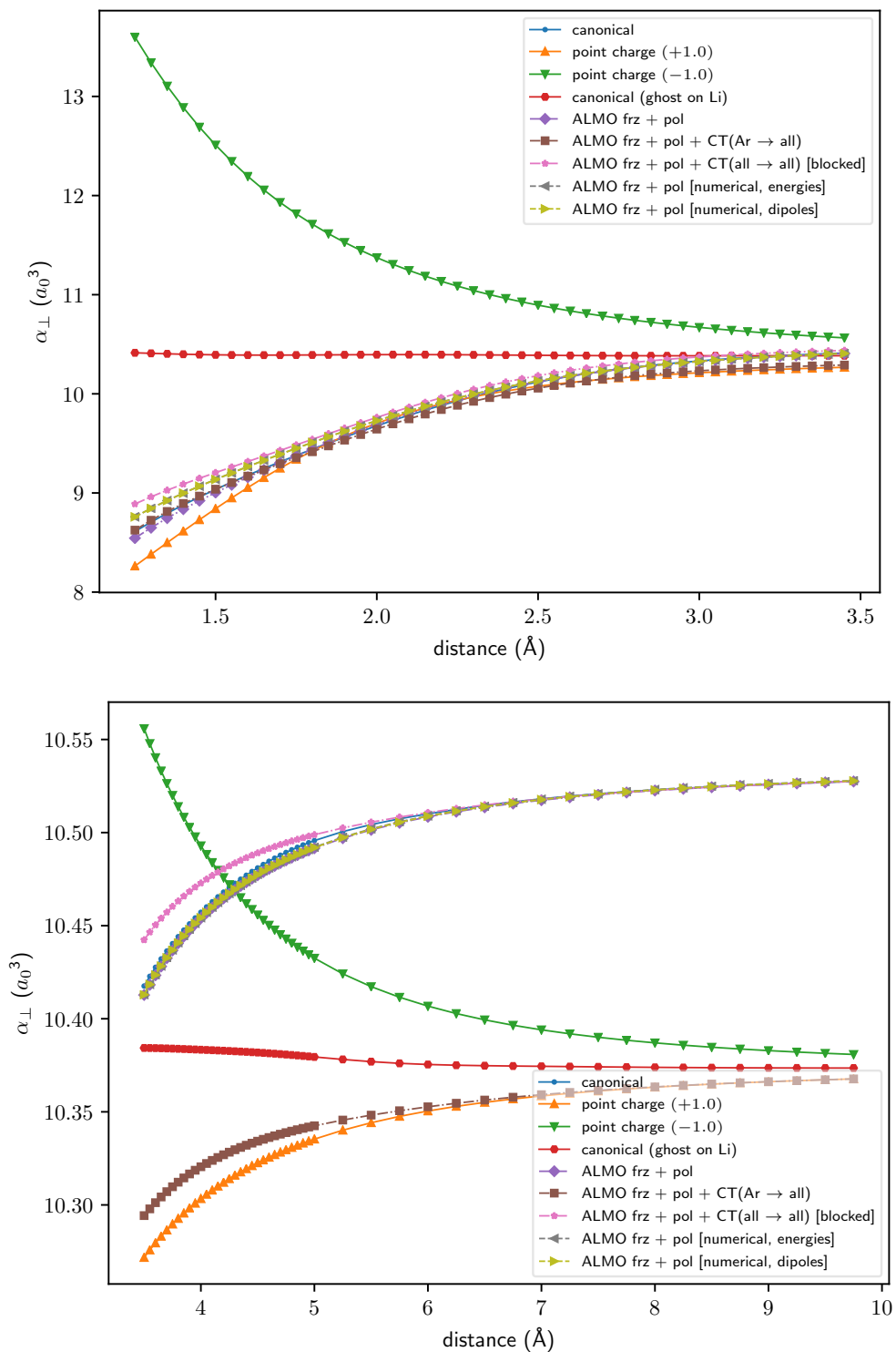


Figure 35: Short- and long-range interatomic separation dependence of the static polarizability perpendicular to the coordination axis. All calculations are at the HF/def2-SVPD level.

```

2  jobtype = polarizability
3  method = hf
4  basis = def2-svpd
5  scf_convergence = 11
6  thresh = 14
7  scf_max_cycles = 1000
8  symmetry = false
9  sym_ignore = true
10 cc_symmetry = false
11 scf_print_frgm = false
12 frgm_method = stoll
13 frgm_lpcorr = 0
14 $end
15
16 $response
17 solver = diis
18 maxiter = 1000
19 conv = 8
20 _almo_do_virt_relocalization = false
21 _almo_project_virts = true
22 _frgm_response_idx = 0
23 _mask_rhsvec_mo = true
24 _mask_rspvec_guess_mo = true
25 _mask_product_mo = true
26 _mask_ediff_mo = true
27 _mask_rspvec_mo = true
28 _mask_form_results_mo = true
29 $end
30
31 $molecule
32 1 1
33 --
34 0 1
35 Ar      0.0000000000000000      0.0000000000000000      -0.857713330500000
36 --
37 1 1
38 Li      0.0000000000000000      0.0000000000000000      1.571999044800000
39 $end

```

Using the template from listing 1:

- To perform the “ALMO frz + pol + CT(all → all) [blocked]” calculations, set `_frgm_response_idx = 0` and all `_mask_*` options to `false`.
- To perform the “ALMO frz + pol + CT(Ar → all)” calculations, set `_frgm_response_idx = 1` and all `_mask_*` options to `true`.

Listing 2: Sample Q-CHEM input file for first-generation ALMO-EDA. Geometry is from

HF/def2-SVPD.

```
1 $rem
2   jobtype = eda
3   method = hf
4   basis = def2-svpd
5   scf_convergence = 11
6   thresh = 14
7   scf_max_cycles = 1000
8   symmetry = false
9   sym_ignore = true
10  cc_symmetry = false
11  scf_print_frgm = false
12  frgm_method = stoll
13  frgm_lpcorr = rs_exact_scf
14  eda_bsse = true
15 $end
16
17 $molecule
18 1 1
19 --
20 0 1
21 Ar      0.0000000000000000      0.0000000000000000      -0.857713330500000
22 --
23 1 1
24 Li      0.0000000000000000      0.0000000000000000      1.571999044800000
25 $end
```

Listing 3: Sample Q-CHEM input file for second-generation ALMO-EDA. Geometry is from HF/def2-SVP.

```
1 $rem
2   method = hf
3   basis = def2-svpd
4   scf_convergence = 11
5   thresh = 14
6   scf_max_cycles = 1000
7   symmetry = false
8   sym_ignore = true
9   cc_symmetry = false
10  scf_print_frgm = false
11  gen_scfman = true
12  frgm_method = stoll
13  frgm_lpcorr = rs_exact_scf
14  eda2 = 1
15 $end
16
17 $molecule
18 1 1
19 --
20 0 1
21 Ar      0.0000000000000000      0.0000000000000000      -0.848158037800000
```



```

22  --
23  1 1
24  Li      0.0000000000000000      0.0000000000000000      1.562443752100000
25  $end

```

### 5.8.2 Algorithm description

Algorithm 1/2 is the general outline of the LR(MI) procedure, without convergence acceleration. The key difference between our implementation and a general CPHF solver is the zeroing of vector and matrix elements. Again, any contraction between  $ia$  or  $ia, jb$  indices may be restricted, but transformations from  $\mu\nu$  to  $ia$  are not.

Missing from this example is the formation of  $\mathbf{A}, \mathbf{B}$ , and the prefactor in the equation for real/imaginary operators. A difference from the paper equations in the implementation is that rather than take a single  $\hat{P}$  and single  $\hat{Q}$ , a list of operators is passed. For example, if  $\text{operators} = [\hat{\mu}, \hat{m}]$ , then  $\langle\langle\hat{\mu}; \hat{\mu}\rangle\rangle$ ,  $\langle\langle\hat{\mu}; \hat{m}\rangle\rangle$ , and  $\langle\langle\hat{m}; \hat{m}\rangle\rangle$  will automatically be formed. Another implementation detail is that each operator carries its own property gradient vectors (the occ-virt MO basis integrals for the right-hand side) and perturbed response vectors, and each operator carries information about whether it is real/imaginary and spin conserving/altering.

---

**Algorithm 1** Static linear response approach within fragment-localized formalism.

---

```
1: procedure SOLVE_LINEAR_RESPONSE(resp, operators, occupations, C, F, S,  $\vartheta$ , maxiter,  
   allow_ct?)  
2:   for  $s \leftarrow 1, N_{spin}$  do  
3:     Transformation of F and S from AO to full MO basis  
4:      $E_{ia,jb} = F_{ab}S_{ij} - F_{ij}S_{ab}$   $\triangleright$  Form non-orthogonal orbital energy matrix E  
5:     if not allow_ct? then  
6:       Zero cross-fragment ia indices and shrink dimensions of E  
7:     end if  
8:     Form inverse for denominator  $\mathbf{E}^{-1}$   
9:   end for  
10:  for  $i \leftarrow 1, N_{operators}$  do  
11:    for  $c \leftarrow 1, N_{components}$  do  
12:       $(\mathbf{Z})_{\mu\nu} \leftarrow operators[i, c, \mu\nu]$   $\triangleright$  Select operator component as perturbation for  
right-hand side  
13:      Transform operator component from AO to occ-virt MO basis and append  
 $\mathbf{Z}_{ia}$  to rhsvecs  
14:      if not allow_ct? then  
15:        Zero cross-fragment ia indices and shrink dimensions of Z  
16:      end if  
17:       $\mathbf{X}^{(0)} \leftarrow \mathbf{0}$   $\triangleright$  Form initial guess for response vector (uncoupled result)  
18:      converged  $\leftarrow$  false  
19:      for  $n \leftarrow 1, maxiter$  do  
20:         $D_{\mu\nu}^X \leftarrow C_{\mu i} X_{ia}^{(n-1)} C_{\nu a}$   $\triangleright$  Form perturbed density  
21:         $J_{\mu\nu}^X[\mathbf{D}^X], K_{\mu\nu}^X[\mathbf{D}^X] \leftarrow fock\_build(\mathbf{D}^X)$   $\triangleright$  Form Coulomb and exchange  
contributions  
22:         $(\mathbf{R}^{(n)})_{\mu\nu} \leftarrow 4\mathbf{J}^X - \mathbf{K}^X - (\mathbf{K}^X)^T$   $\triangleright$  Form half-transformed orbital  
Hessian-vector product (here,  $\mathbf{G} = \mathbf{A}^s + \mathbf{B}^s$ )  
23:        Do second transformation of Hessian-vector product  $(\mathbf{R})_{ia}^{(n)}$   
24:        if not allow_ct? then  
25:          Zero cross-fragment ia indices and shrink dimensions of R, X  
26:        end if  
27:         $X_{ia}^{(n)} \leftarrow (\mathbf{E}^{-1})_{ia,jb} [Z_{jb} - R_{jb}^{(n)}]$   $\triangleright$  Update response vector  
28:        if not allow_ct? then  
29:          Restore dimensions of X  
30:        end if  
31:        if  $\|\mathbf{X}^{(n)} - \mathbf{X}^{(n-1)}\| < \vartheta$  then  
32:          converged  $\leftarrow$  true  
33:          Append  $\mathbf{X}^{(n)}$  to rspvecs  
34:          break  
35:        end if
```

---

---

**Algorithm 2** Continuation of algorithm 1

---

```
36:         end for
37:         if not converged then
38:             crash
39:         end if
40:     end for
41: end for
42: resp ← 0 ▷ Form all possible permutations of property gradient and response vectors
43: for a ← 1, len(rhsvecs) do
44:     for b ← 1, len(rspvecs) do
45:         P ← rhsvecs[a], Q ← rspvecs[b]
46:          $\langle\langle \hat{P}; \hat{Q} \rangle\rangle_0 \leftarrow -P_{ia}Q_{ia}$ 
47:         resp[a, b] ←  $\langle\langle \hat{P}; \hat{Q} \rangle\rangle_0$ 
48:     end for
49: end for
50: end procedure
```

---

## 6.0 Psi4NumPy: AN INTERACTIVE QUANTUM CHEMISTRY PROGRAMMING ENVIRONMENT FOR REFERENCE IMPLEMENTATIONS AND RAPID DEVELOPMENT

The text in this chapter has been adapted from Smith, D.; Burns, L. A.; Sirianni, D. A.; Nascimento, D. R.; Kumar, A.; James, A. M.; Schriber, J. B.; Zhang, T.; Zhang, B.; Abbott, A. S.; Berquist, E. J.; Lechner, M. H.; dos Anjos Cunha, L.; Heide, A. G.; Waldrop, J. M.; King, R. A.; Simmonett, A. C.; Turney, J. M.; Schaefer, H. F.; Evangelista, F. A.; DePrince III, A. E.; Crawford, T. D.; Patkowski, K.; Sherrill, C. D. Psi4NumPy: An Interactive Quantum Chemistry Programming Environment for Reference Implementations and Rapid Development. **2018**, DOI: [10.26434/chemrxiv.5746059.v1](https://doi.org/10.26434/chemrxiv.5746059.v1). The author's contributions to this work were the reference implementation and Jupyter Notebook tutorial for the SCF first hyperpolarizability, presented in sections [6.9.1](#) and [6.9.2](#).

### 6.1 SUMMARY

PSI4NUMPY demonstrates the use of efficient computational kernels from the open-source PSI4 program through the popular NUMPY library for linear algebra in Python to facilitate the rapid development of clear, understandable Python computer code for new quantum chemical methods, while maintaining a relatively low execution time. Using these tools, reference implementations have been created for a number of methods, including self-consistent field (SCF), SCF response, many-body perturbation theory, coupled-cluster theory, configuration interaction, and symmetry-adapted perturbation theory. Further, several reference codes have been integrated into Jupyter notebooks, allowing background and explanatory

information to be associated with the implementation. `PSI4NUMPY` tools and associated reference implementations can lower the barrier for future development of quantum chemistry methods. These implementations also demonstrate the power of the hybrid C++/Python programming approach employed by the `PSI4` program.

## 6.2 INTRODUCTION

Whereas in the past a new quantum chemical (QC) method was commonly presented solely through its equations, perhaps along with a few token values, the more recent expectation is that equations will be accompanied by results from an effective computer program clearly demonstrating the utility of the method. This expectation becomes increasingly burdensome as new computer architectures emerge, since some theories will be naturally more computationally efficient or more difficult to implement than others. The computation expense of most quantum chemical methods creates substantial pressure for methods to be implemented with highly optimized algorithms.

This situation presents a challenge for ongoing development in quantum chemistry, because new theoretical methods are typically complex and their correct implementation is non-trivial. Additionally, computationally efficient codes require a low-level programming language like C++ or Fortran, followed by substantial code profiling, testing, and optimization. Often a method’s first implementation is a rather messy computer program. The researcher may be learning the details of the method as they progress, resulting in “experimental” parts of the code that may never get removed, or data structures that may not be optimal for the final version of the method. Additionally, development is often carried out by graduate students not yet proficient in programming, resulting in unconventional coding styles. Subsequently, a researcher seeking to extend or enhance a method previously developed in-house is often faced with the daunting prospect of deciphering a quite complex existing code.

Still more challenging is implementing or extending an existing method sourced solely from the literature. Often, a paper describing a new quantum chemical method that properly

focuses on scientific detail falls short on algorithmic or numerical detail sufficient for independent reimplementations. Indeed, the methods are so complex that the original equations frequently include typos, which are generally tracked through institutional lore rather than published errata. Additionally, modern approaches often employ combinations of approximations with multiple numerical cutoffs, exacerbating the reproducibility problem. This paradigm is illustrated within a recent comment,<sup>288</sup> whereby several corrections to equations originally published in 2011 for a two-level semi-empirical method<sup>289</sup> were proposed after being re-engineered to reproduce values computed using a binary program distributed with the original publication. Even facilitated through private communication with the method’s author, this cycle of rediscovery and reimplementations is both highly non-trivial and unsustainable. Fortunately, an open-source program<sup>290</sup> has been made available by the commenting author that implements the method and proposed changes, so that further extensions of the method can proceed with this program as a reference.

Such “reference implementations” (easy-to-read, unoptimized computer programs solely targeting the correct result) can be a helpful initial step toward developing or understanding a complex method, yet they are not widely available in quantum chemistry. To our knowledge, reference implementations and benchmarking have only been performed in a large-scale way for density functional theory (DFT) exchange-correlation kernels<sup>291</sup> and periodic boundary condition DFT with pseudopotentials.<sup>292</sup> One factor limiting more widespread use of reference implementations for quantum chemistry is that methods are often so computationally demanding that a basic, unoptimized implementation is too slow for computations on even the smallest molecules. What is needed is an alliance of QC code that is easy to peruse and manipulate with underlying non-QC routines that are fast enough for testing on non-trivial molecules.

Here we present `PSI4NUMPY`, a framework for the creation of clear, readable reference implementations of quantum chemical methods and for the rapid development of new methods. `PSI4NUMPY` takes advantage of `PSI4`’s<sup>284</sup> application programming interface (API) that makes efficient computational kernels written in C++ available from Python, a language that is easy to learn and has become very popular in scientific computing. As a high-level language, Python allows complex tasks to be specified with relatively few lines of code.

PSI4NUMPY capitalizes on the straightforward conversion of PSI4 tensors to NUMPY<sup>293</sup> and Numerical Python’s (NUMPY’s) own low-level back end to ensure that all data arrays can use the optimized Basic Linear Algebra Subroutines (BLAS) library<sup>294</sup> for common linear algebra operations. The wide user base of NUMPY ensures constant updates and bug fixes. PSI4NUMPY has been packaged for minimal setup, requiring only 3 minutes, with no preinstalled compilers necessary on 64-bit Linux, Mac, and Windows. Here we introduce the main elements of the PSI4NUMPY framework and illustrate them with a substantial collection of reference implementations for standard quantum chemical methods and numerical techniques. The PSI4NUMPY is built entirely on Free and Open Source Software (FOSS)<sup>295</sup> as shown in Fig. 36 to ensure a barrierless entry to quantum chemistry programming.

Several of the reference implementations have been augmented by tutorial-style introductions to the relevant theory. The PSI4NUMPY tutorial collection includes self-consistent field (SCF), DFT,<sup>296</sup> many-body perturbation theory (MBPT),<sup>297</sup> symmetry-adapted perturbation theory (SAPT),<sup>156,298</sup> coupled-cluster (CC),<sup>299</sup> and configuration interaction (CI)<sup>300,301</sup> theories, with additional sections detailing the theory and implementation of linear response, geometry optimizations, and Verlet integrators. It is our hope that PSI4NUMPY and the accompanying reference code will lower the barrier to implementing quantum chemical methods.

Shortly before submission, we discovered the Quantum Chemistry Program Exchange (QCPE),<sup>302,303</sup> whose goals of (self-contained) software accessibility, algorithm explication, and free software “publishing” the PSI4NUMPY project shares. The general tools embraced by PSI4NUMPY (GitHub for communication, NUMPY for linear algebra, Python for interfacing, and Jupyter for illumination) further allow rapid prototyping and educational objectives. In this manner PSI4NUMPY can be thought as a modern successor to QCPE built to serve the flexible needs of the community.

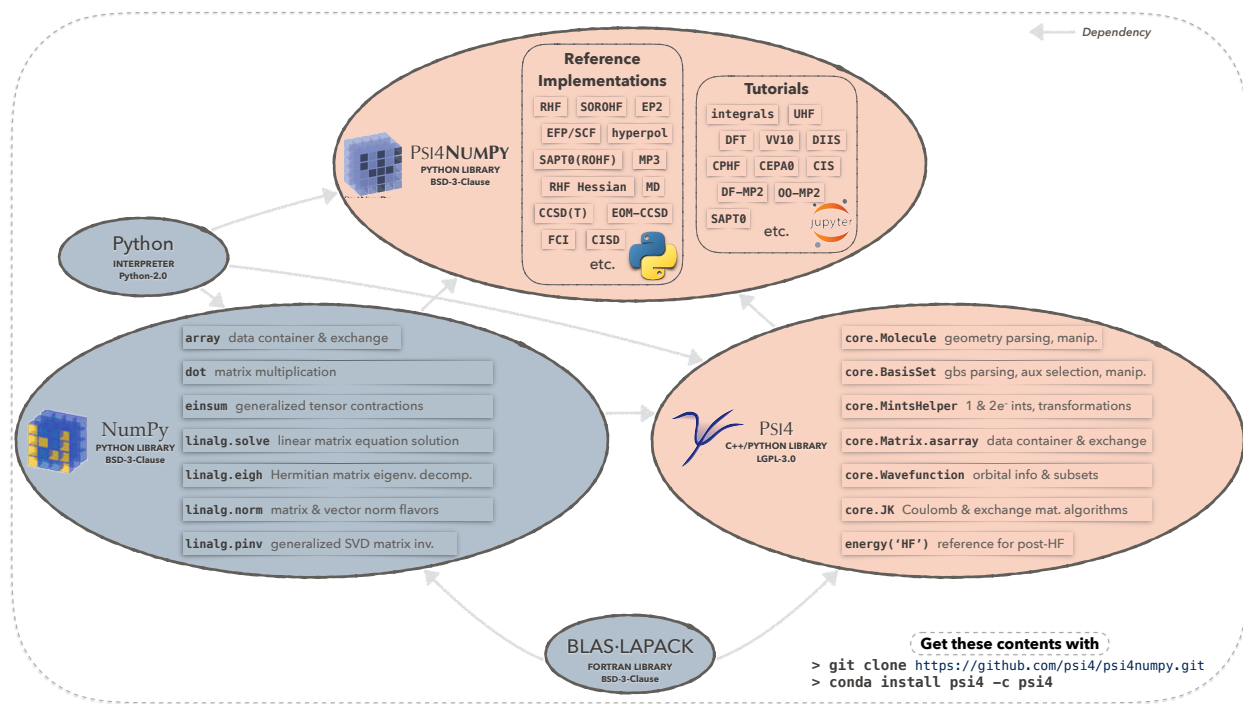


Figure 36: Psi4NUMPY draws linear algebra tools from NUMPY and fundamental quantum chemistry structures from Psi4 to bring together a practical and convenient environment for code development, verification, and exploration. The most important data structures and functions are shown for NUMPY and Psi4 as well as representative tutorial and reference implementations presently in Psi4NUMPY.



## 6.3 BASIC TOOLS

The basic premise of `PSI4NUMPY` is to leverage `PSI4` to generate quantum chemistry-specific quantities and the `NUMPY` library<sup>293</sup> for all other tensor manipulations. The latest version of `PSI4` has added the capability to import `PSI4` as a Python module as well as continuing to be called in an executable fashion. In this way, both the `PSI4` and `NUMPY` libraries can be loaded into a single Python script and used in cooperation.

A key capacity in this enterprise is seamless translation between `NUMPY` and `PSI4` data classes. For example, converting from a `NUMPY` array to a `PSI4` matrix and back again can be easily accomplished:

Listing 4: Basic example of `PSI4`↔`NUMPY` interoperability

```
1 import numpy, psi4
2 np_array = numpy.zeros((5, 5))
3 psi4_matrix = psi4.core.Matrix.from_array(np_array)
4 new_np_array = numpy.array(psi4_matrix)
```

At the core of this procedure is `NUMPY`'s `array_interface`<sup>304</sup> protocol, a basic specification for dense matrices consisting of

1. (a) the starting memory location for an in-memory array
2. (b) the overall “shape” of the array [( $n$ ,) for a vector, ( $n$ ,  $m$ ) for a matrix, etc.]
3. (c) the type of data involved (`double64`, `int32`, etc.)

This specification is compact and widely used amongst the scientific Python community in a variety of scenarios. Using the `array_interface`, it becomes straightforward to allow `NUMPY` access to a `PSI4` data class, allowing both `PSI4` and `NUMPY` to access and manipulate the same data. For example, the below will overwrite the `PSI4` Matrix class in place with a random `NUMPY` array:

Listing 5: Mutating a `PSI4` Matrix directly from a `NUMPY` routine

```
1 psi4_matrix.np[:] = numpy.random.rand(5, 5)
```

In this way the typical separation between general tensor frameworks and custom quantum chemistry data structures is removed.

A description of the full set of capabilities of the `array_interface` is available in the PSI4 documentation: <http://psicode.org/psi4manual/master/numpy.html>.

### 6.3.1 Wavefunction Objects

In PSI4 all built-in methodologies have the option to return a Wavefunction object that holds basic information about the previous computation or, in some cases, holds functions for readily computing advanced quantities. Obtaining the Wavefunction object in this manner is straightforward:

Listing 6: Initializing a PSI4 computation from Python

```
1 mol = psi4.geometry("""
2 0
3 H 1 0.96
4 H 1 0.96, 104.5
5 """)
6 hf_e, hf_wfn = psi4.energy("HF/cc-pVDZ", molecule=mol, return_wfn=True)
```

Once a Wavefunction object is obtained, a variety of attributes can be queried using standard Python syntax:

Listing 7: Obtaining PSI4 wavefunction data in Python

```
1 # Number of doubly occupied orbitals
2 docc = hf_wfn.ndocc()
3 # Alpha orbital coefficient matrix
4 Ca = hf_wfn.Ca()
5 # Occupied subset of the alpha orbitals
6 Ca_occ = hf_wfn.Ca_subset("AO", "OCC")
```

In addition to generating useful information after a computation, a Wavefunction object can also be passed as reference state to a further computation. For PSI4NUMPY, this means that reference implementations of post-Hartree–Fock methods (MPn, CCSD, etc.) need not re-code their own Hartree–Fock program; this simultaneously reduces code duplication and increases readability, both of which are cornerstones of the PSI4NUMPY project.

### 6.3.2 Integrals

PSI4 offers a wide selection of efficient C++ tools accessible directly in Python. These tools are largely object-based and capable of storing quantities in memory or on disk. One such object is the `libmints` library,<sup>284</sup> which is currently the primary interface for computing one- and two-electron integrals in PSI4. This library is accessible through the `MintsHelper` class that directs the efficient computation and storage of molecular integrals Python-side:

Listing 8: Computing atomic orbital-basis integrals from PSI4 via Python

```
1 # Create instance of MintsHelper using primary basis set
2 mints = psi4.core.MintsHelper(primary_basis)
3 # Compute one-electron AO overlap matrix
4 S = mints.ao_overlap()
5 # Compute core Hamiltonian matrix
6 T = mints.ao_kinetic()
7 V = mints.ao_potential()
8 H = T + V
9 # Compute two-electron integrals in AO basis in memory
10 I_ao = mints.ao_eri()
```

Each of the above `MintsHelper` class methods returns a PSI4 matrix which can be converted to a NUMPY array using `numpy.asarray(matrix)` or modified in place with the `matrix.np` accessor.

In addition to computing molecular integrals, the `libmints` library also performs optimized electron repulsion integral (ERI) transformations. For example, the  $\mathcal{O}(N^5)$  transformation of the two-electron integrals between the atomic orbital and molecular orbital basis, given by

$$(ia|jb) = [[C_{\mu i} [C_{\nu a}(\mu\nu|\lambda\sigma)]] C_{\lambda j}] C_{\sigma b}, \quad (6.1)$$

can be performed easily with:

Listing 9: Calling PSI4's AO-to-MO transformation from Python

```
1 # Occupied and virtual subsets of SCF orbital coefficient matrices
2 Ca_occ = hf_wfn.Ca_subset("AO", "OCC")
3 Ca_virt = hf_wfn.Ca_subset("AO", "VIR")
4 # AO basis to MO basis in-memory ERI transform
5 I_mo = mints.mo_transform(Ca_occ, Ca_virt, I_ao, Ca_occ, Ca_virt)
```

In this manner, arbitrary ERI transformations may be performed, allowing both speed and flexibility for constructing reference implementations.

### 6.3.3 Coulomb and Exchange (JK) Matrix Objects

A key component in SCF-level theories is the contraction of the 4-index electron repulsion integrals with the 2-index density matrix to form  $J$  and  $K$  matrices:

$$J_{\lambda\sigma}[D] \equiv (\lambda\sigma|\mu\nu)D_{\mu\nu}, \quad (6.2)$$

$$K_{\lambda\sigma}[D] \equiv (\lambda\mu|\sigma\nu)D_{\mu\nu} \quad (6.3)$$

PSI4 provides objects for computing generalized Coulomb (J) and Exchange (K) matrices, with specialized algorithms for integral-direct, PK supermatrix,<sup>305</sup> or density fitting (DF) scenarios. For the DF-JK object, it is often advantageous to use a factorized form of the density matrix,

$$D_{\mu\nu} \equiv C_{\mu p}^{\text{left}} C_{\nu p}^{\text{right}}, \quad (6.4)$$

where  $p$  is a general MO index. For example, in canonical Restricted Hartree Fock (RHF), the density matrix takes the form of

$$D_{\mu\nu}^{\text{RHF}} = C_{\mu i} \chi_{ia} C_{\nu a}, \quad (6.5)$$

where  $i$  runs only over occupied orbitals. The computation of the RHF JK matrices can be translated directly to Python code with the following lines:

Listing 10: Building two-electron Fock matrix components using PSI4 via Python

```

1 # Create a JK object in the current primary basis set
2 jk = psi4.core.JK.build(primary_basis)
3 # Add the occupied parts of the SCF orbital matrix
4 jk.add_C_left(C_occupied)
5 jk.add_C_right(C_occupied)
6 # Perform the computation and obtain the J and K matrices
7 jk.compute()
8 J = jk.J()
9 K = jk.K()

```

In this fashion, virtually any SCF-level theory can be formulated at the `PSI4NUMPY` layer by handling only 2-D arrays with `NUMPY` (typically by threaded vendor BLAS) and leaving the 3- and 4-D arrays to `PSI4` libraries (using optimized C++ routines). Thus, SCF-level theories can be implemented with the same efficiency as their pure C++ counterparts.

To illustrate this point, the `PSI4 SCF` program is compared against a `PSI4NUMPY` implementation on an Intel i7-5930K processor with the adenine-thymine complex in the aug-cc-pVTZ basis (1127 basis functions) using a density-fitted JK build on six cores. The `PSI4 SCF` program took 250 seconds while the `PSI4NUMPY` implementation took 245 seconds. This should not be surprising as each spent 94% total wall time computing the J and K quantities (both implementations used 18 SCF iterations) and all other operations of nonnegligible cost use the same BLAS implementations.

## 6.4 RAPID DEVELOPMENT

A key component of the `PSI4NUMPY` framework is to provide an easy-to-use development environment for rapid prototyping. Vital to this is `NUMPY`'s `einsum` function that performs arbitrary tensor contractions using Einstein summation syntax. For example, the atomic orbital to molecular orbital 4-index transformation of Eq. (6.1) and code snippet 9 could be accomplished by:

Listing 11: Performing an AO-to-MO transformation using `NUMPY`'s `einsum`

```

1 I_mo = numpy.einsum("pi,qa,pqrs,rj,sb->iajb",
2                   Ca_occ, Ca_virt, I_ao,
3                   Ca_occ, Ca_virt)
```

Recently, one of us (D.G.A.S.) modified `NUMPY`'s `einsum` function so that it will automatically factorize the incoming tensor expression to reduce the cost of the operation from naive  $\mathcal{O}(N^8)$  to the conventional  $\mathcal{O}(N^5)$  version. This feature is available in `NUMPY` 1.12 and onwards, with additional optimizations and BLAS usage occurring in `NUMPY` 1.14. In addition, a drop-in replacement for the `einsum` function, which makes optimal use of vendor BLAS, can be found through the Optimized Einsum project.<sup>306</sup>

Using the `einsum` function, it is straightforward to transcribe existing equations directly into working code without a compilation stage. While the resulting program is not as efficient for post-SCF level theories as a full implementation in a low-level language, the code is easy to read and modify without the need for compilation, allowing considerable flexibility when prototyping. In addition, the resulting program will provide correct answers for the given expressions, sparing the developer any worry whether low-level code is correct.

As an example of rapid prototyping, we use a temporary CCSD quantity in the Direct Product Decomposition formalism.<sup>307</sup> For virtual indices  $a, b, c, d$  and occupied indices  $i, j, k$ , Equation 8 of Ref. [307] is written as:

$$W_{jaci} = \langle ja||ci \rangle + t_i^d \langle ja||cd \rangle - t_k^a \langle jk||ci \rangle - \left( \frac{1}{2} t_{ik}^{da} + t_i^d t_k^a \right) \langle jk||cd \rangle, \quad (6.6)$$

which can be directly translated into a function:

Listing 12: Example of forming a coupled cluster intermediate using `NUMPY` and `einsum`

```

1 def build_Wjaci(T1, T2, MO):
2     Wjaci = MO[o, v, v, o].copy()
3     Wjaci += numpy.einsum("jid,jacd->jaci", T1, MO[o, v, v, v])
4     Wjaci -= numpy.einsum("ka,jkci->jaci", T1, MO[o, o, v, o])
5     tmp = 0.5 * T2 + numpy.einsum("jid,ka->ikda", T1, T1)
6     Wjaci -= numpy.einsum("ikda,jkcd->jaci", tmp, MO[o, o, v, v])
7     return Wjaci

```

Here, `MO` holds the 4-index antisymmetrized integrals, `T1` and `T2` the current amplitudes, and the `o`, `v` quantities are Python-based slices so that `MO[o, v, v, v]` returns the occupied-virtual-virtual-virtual blocks of the antisymmetrized integrals.

To our knowledge, the first implementations of Symmetry-Adapted Perturbation Theory with Complete Active Space SCF references [SAPT(CASSCF)], fourth-order Electron Propagator Theory, and transcorrelated theories have all been achieved using these rapid prototyping techniques.

## 6.5 ACCESS AND CONTRIBUTIONS

To ensure ease of community access to the PSI4NUMPY project, all software dependencies are made available as binary Conda packages<sup>308</sup> either by us (e.g., PSI4), or by ContinuumIO or Intel (e.g., NUMPY, Matplotlib, Jupyter). Through this route, binary distributions are installable in a single line to all common computing platforms, so users are not required to compile, link against the correct libraries, or debug runtime issues. We hope that the ready accessibility of these tools facilitates their use in methods development and in the creation of additional publicly available reference implementations.

To lower the barrier to contribution, guidance is included in the repository regarding attribution, citations, and testing. Though the authors adhere to Python software development best practices in their other projects, they resist advanced Python syntax, organization, file linking, or other jargon-ized code in PSI4NUMPY in favor of straightforward scripts and Jupyter notebooks for ease of community involvement. Educators are encouraged to base lessons and labs off this work and are also referred to the PSI4EDUCATION project.<sup>309</sup>

## 6.6 REFERENCE IMPLEMENTATIONS

To illustrate the PSI4NUMPY tools, and to provide a resource to the quantum chemistry methods development community, we have created a number of reference implementations and made them publicly available on GitHub at <https://github.com/psi4/psi4numpy>. We intend to add to this collection over time. Given the wide spectrum of quantum chemical methods, we also encourage submissions from other developers.

The PSI4NUMPY reference implementations, while not necessarily as efficient as optimized versions in a low-level language, furnish at least the basic requirements for a programmer to reproduce the methodology. These references provide a medium to explain minute details that might be included in a corresponding paper and to record algorithmic tricks used to improve numerical stability or computational efficiency. In addition, these clear implementations will make explicit any important steps that might not be mentioned in a

paper because they are assumed to be background knowledge in a given subfield of quantum chemistry.

Programmers can use these reference implementations to obtain intermediate quantities to validate a new implementation at every step, ensuring accuracy and assisting in the process of debugging a new program. These reference implementations can also be used as starting points for either building upon existing methodologies or exploring new methodologies in combination with the rapid prototyping aspects of this project.

Current reference implementations include

1. Self-Consistent Field
  - a. Restricted simple and DIIS<sup>98</sup>-accelerated Hartree–Fock
  - b. Restricted, Unrestricted, and Restricted Open-Shell Hartree–Fock
  - c. Restricted, Unrestricted, and Restricted Open-Shell Hartree–Fock time-independent orbital Hessians
  - d. Restricted time-dependent Hartree–Fock and coupled-perturbed Hartree–Fock for dipole polarizabilities
  - e. Restricted nuclear gradients and Hessians
2. Many-Body Perturbation Theory
  - a. Canonical and density-fitted MP2
  - b. Spin-integrated and spin-orbital MP3
  - c. Arbitrary-order MP
3. Coupled-Cluster
  - a. Simple and DIIS-accelerated CCSD
  - b. CCSD(T)
  - c. CCSD dipole polarizabilities
  - d. Time-dependent equation-of-motion CCSD
4. Configuration Interaction
  - a. Excited-state CIS
  - b. Canonical and Davidson–Liu CISD
  - c. Full configuration interaction
5. Symmetry-Adapted Perturbation Theory



The Fock matrix,  $\mathbf{F}$ , has elements  $F_{\mu\nu}$  given (in the atomic orbital basis) as

$$F_{\mu\nu} = H_{\mu\nu} + 2(\mu\nu|\lambda\sigma)D_{\lambda\sigma} - (\mu\lambda|\nu\sigma)D_{\lambda\sigma},$$

where  $D_{\lambda\sigma}$  is an element of the one-particle density matrix  $\mathbf{D}$ , constructed from the orbital coefficient matrix  $\mathbf{C}$ :

$$D_{\lambda\sigma} = C_{\sigma i} C_{\lambda i}$$

Using the above equations, the Fock matrix can be constructed as:

```
In [1]: D = numpy.einsum('pi,qi->pq', Cocc, Cocc)

J = numpy.einsum('pqrs,rs->pq', I, D)
K = numpy.einsum('prqs,rs->pq', I, D)

F = H + 2 * J - K
```

Figure 37: Extract from a Jupyter notebook demonstrating the construction of a SCF Fock matrix where  $\mathbf{I}$  is the 4-index electron repulsion integral array and  $\mathbf{C}_{occ}$  is the occupied orbital matrix.

- a. Restricted and Restricted Open-Shell SAPT0
- b. Atomic orbital implementation of SAPT0
6. Electron Propagator Theory
  - a. Spin-integrated and spin-orbital EP2
  - b. Spin-orbital EP3

### 6.6.1 Jupyter Notebook integration

As a service to the community, some of the reference implementations have been augmented by additional, tutorial-style background information on various subfields of quantum chemistry. We found it convenient to add this additional information using the Jupyter notebook web application,<sup>310</sup> a popular Integrated Development Environment (IDE) for interactive computing in several programming languages that is starting to be adopted by chemists.<sup>311</sup> An example for restricted Hartree–Fock can be found in Fig. 37.

These documents may be unique within quantum chemistry in that they focus not only on theoretical considerations but also on the details of a method's implementation, such as *why* certain programming choices were made. For example, the comparison between a general matrix inversion and solving a set of linear equations demonstrates instability issues that often plague the former technique. Such illustrations should make the Jupyter implementations useful both to new users in quantum chemistry and to experienced users interested in exploring new subfields.

Current tutorial-style Jupyter reference implementations include

1. Introductions to the PSI4NUMPY methodology
2. Introduction to Hartree–Fock, DIIS, and density fitting
3. Density Functional Theory: grids, LDA kernels, VV10 dispersion, and asymptotic corrections
4. Møller–Plesset: canonical and density-fitted reference implementations of MP2
5. Molecular Properties: Integrals, CPHF, CIS
6. Symmetry-Adapted Perturbation Theory: Canonical and atomic orbital algorithms
7. Orbital-Optimized Methods: OMP2
8. Coupled-Cluster Approximations: CEPA0, CCD
9. Geometry Optimization Techniques: Internal Coordinates, Hessian guesses, and advanced Newton-Raphson methods

Molecular-dynamics tutorials include

1. Periodic Lennard-Jones simulation with Verlet integrators
2. Periodic Ewald Electrostatic summation

## 6.7 CONCLUSIONS

We believe that the benefits of the PSI4NUMPY framework to the computational chemistry community are threefold. Beginning researchers can use the PSI4NUMPY reference implementations for *education*. Reference implementations convey not just the underlying

mathematical formulas of a given theory, but how to implement these formulas in a manner that avoids common pitfalls such as ill-conditioned numerical equations. PSI4NUMPY is likely the most interactive educational resource available in this field: thanks to the Jupyter Notebook format, the learners can explore the implementation step by step and easily try out various modifications and additional approximations.

More advanced researchers who need to reimplement and/or modify a given computational chemistry approach can use the PSI4NUMPY reference implementations for *validation*, taking advantage of the code that, thanks to the extensive use of the NUMPY `einsum` functionality, provides a nearly one-to-one correspondence between the terms in a formula and the lines of Python code. As a result, it is trivial to switch off, for debugging purposes, any subset of terms as well as generate an arbitrary intermediate without even recompiling any code. This feature should be contrasted with the situation when one tries to validate their code against a C++/Fortran implementation from an established electronic-structure package. Once the relevant fragment of code that does the actual computation is found (which is not always trivial), various terms are typically combined in nontrivial ways to improve computational performance. As a result, getting out a specific intermediate for checking the implementation in progress often requires substantive changes to the reference code, not to mention its recompilation. In addition, we include the programmed formulas together with their implementation in the Jupyter Notebook to alleviate difficulties associated with incompatible notation or even errors in the originally published expressions.

Finally, for researchers who want to develop new functionality, PSI4NUMPY is a highly valuable platform for *initial implementation* that is efficient enough for meaningful testing, quick to generate, easy to debug, and has few opportunities for programming errors. All underlying quantum-chemistry building blocks such as integrals, orbitals, density matrices, and CI vectors are efficiently computed by PSI4 and readily imported in the NUMPY format. In particular, a PSI4NUMPY implementation of any one-electron theory such as HF or DFT is already close to optimal as the most expensive operations are all written in terms of generalized Coulomb and exchange matrices which are supplied by PSI4. Some of us, together with their collaborators, have already taken advantage of the PSI4NUMPY capabilities to rapidly generate pilot implementations of brand new electronic-structure approaches.

## 6.8 ACKNOWLEDGEMENTS

This work was supported in part by the U.S. National Science Foundation through grants ACI-1449723 and CHE-1566192 for D.G.A.S, L.A.B., D.A.S, and C.D.S; CHE-1661604 for B.Z., A.S.A, J.M.T., and H.F.S; CHE-1554354 for D.R.N and A.E.D. B. Z. contributions to this work were also supported by a Software Fellowship from the Molecular Sciences Software Institute, which is funded by the U.S. National Science Foundation (ACI-1547580). M.H.L. acknowledges financial support by the Studienstiftung des Deutschen Volkes. K.P. is supported by the U.S. National Science Foundation CAREER award CHE-1351978.

## 6.9 SUPPORTING INFORMATION

Documents reproducing all currently available reference implementations and interactive tutorials are available free of charge via the Internet at <https://zenodo.org/record/1134320>. For all future materials, please see <https://github.com/psi4/psi4numpy>. The code found in sections 6.9.1 and 6.9.2 was merged in commit hash 066f378.

### 6.9.1 Hyperpolarizability Reference Implementation

```
1  """
2  Helper classes and functions for molecular properties requiring
3  solution of CPHF equations.
4  """
5
6  __authors__   = "Daniel G. A. Smith"
7  __credits__  = ["Daniel G. A. Smith", "Eric J. Berquist"]
8
9  __copyright__ = "(c) 2014-2017, The Psi4NumPy Developers"
10 __license__   = "BSD-3-Clause"
11 __date__     = "2017-8-30"
12
13 import time
14 import numpy as np
15 np.set_printoptions(precision=5, linewidth=200, suppress=True)
16 import psi4
17
18 import os.path
19 import sys
```

```

20  dirname = os.path.dirname(os.path.abspath(__file__))
21  sys.path.append(os.path.join(dirname, '../..Self-Consistent-Field'))
22  from helper_HF import DIIS_helper
23
24
25  class helper_CPHF(object):
26
27      def __init__(self, mol, numpy_memory=2):
28
29          self.mol = mol
30          self.numpy_memory = numpy_memory
31
32          # Compute the reference wavefunction and CPHF using Psi
33          scf_e, self.scf_wfn = psi4.energy('SCF', return_wfn=True)
34
35          self.C = self.scf_wfn.Ca()
36          self.Co = self.scf_wfn.Ca_subset("AO", "OCC")
37          self.Cv = self.scf_wfn.Ca_subset("AO", "VIR")
38          self.epsilon = np.asarray(self.scf_wfn.epsilon_a())
39
40          self.nbf = self.scf_wfn.nmo()
41          self.nocc = self.scf_wfn.nalpha()
42          self.nvir = self.nbf - self.nocc
43
44          # Integral generation from Psi4's MintsHelper
45          self.mints = psi4.core.MintsHelper(self.scf_wfn.basiset())
46
47          # Get nbf and ndocc for closed shell molecules
48          print('\nNumber of occupied orbitals: %d' % self.nocc)
49          print('Number of basis functions: %d' % self.nbf)
50
51          # Grab perturbation tensors in MO basis
52          nCo = np.asarray(self.Co)
53          nCv = np.asarray(self.Cv)
54          self.tmp_dipoles = self.mints.so_dipole()
55          self.dipoles_xyz = []
56          for num in range(3):
57              Fso = np.asarray(self.tmp_dipoles[num])
58              Fia = (nCo.T).dot(Fso).dot(nCv)
59              Fia *= -2
60              self.dipoles_xyz.append(Fia)
61
62          self.x = None
63          self.rhsvecs = None
64
65      def run(self, method='direct', omega=None):
66          self.method = method
67          if self.method == 'direct':
68              if not omega:
69                  self.solve_static_direct()
70              else:
71                  self.solve_dynamic_direct(omega=omega)
72          elif self.method == 'iterative':
73              if not omega:

```

```

74         self.solve_static_iterative()
75     else:
76         self.solve_dynamic_iterative(omega=omega)
77     else:
78         raise Exception("Method %s is not recognized" % self.method)
79     self.form_polarizability()
80
81 def solve_static_direct(self):
82     # Run a quick check to make sure everything will fit into memory
83     I_Size = (self.nbf ** 4) * 8.e-9
84     oNNN_Size = (self.nocc * self.nbf ** 3) * 8.e-9
85     ovov_Size = (self.nocc * self.nocc * self.nvir * self.nvir) * 8.e-9
86     print("\nTensor sizes:")
87     print("ERI tensor           %4.2f GB." % I_Size)
88     print("oNNN MO tensor         %4.2f GB." % oNNN_Size)
89     print("ovov Hessian tensor    %4.2f GB." % ovov_Size)
90
91     # Estimate memory usage
92     memory_footprint = I_Size * 1.5
93     if I_Size > self.numpy_memory:
94         psi4.core.clean()
95         raise Exception("Estimated memory utilization (%4.2f GB) exceeds numpy_memory
96     ↪ \
97                             limit of %4.2f GB." % (memory_footprint, self.numpy_memory))
98
99     # Compute electronic Hessian
100    print('\nForming Hessian...')
101    t = time.time()
102    docc = np.diag(np.ones(self.nocc))
103    dvir = np.diag(np.ones(self.nvir))
104    eps_diag = self.epsilon[self.nocc:].reshape(-1, 1) - self.epsilon[:self.nocc]
105
106    # Form [0,N,N,N] MO tensor, oN^4 cost
107    MO = np.asarray(self.mints.mo_eri(self.Co, self.C, self.C, self.C))
108
109    H = np.einsum('ai,ij,ab->iajb', eps_diag, docc, dvir)
110    H += 4 * MO[:, self.nocc:, :self.nocc, self.nocc:]
111    H -= MO[:, self.nocc:, :self.nocc, self.nocc:].swapaxes(0, 2)
112
113    H -= MO[:, :self.nocc, self.nocc:, self.nocc:].swapaxes(1, 2)
114
115    print('...formed Hessian in %.3f seconds.' % (time.time() - t))
116
117    # Invert Hessian (o^3v^3)
118    print('\nInverting Hessian...')
119    t = time.time()
120    Hinv = np.linalg.inv(H.reshape(self.nocc * self.nvir, -1)).reshape(self.nocc,
121    ↪ self.nvir, self.nocc, self.nvir)
122    print('...inverted Hessian in %.3f seconds.' % (time.time() - t))
123
124    # Form perturbation response vector for each dipole component
125    self.x = []
126    for numx in range(3):

```

```

126         xcomp = np.einsum('iajb,ia->jb', Hinv, self.dipoles_xyz[numx])
127         self.x.append(xcomp.reshape(-1))
128
129     self.rhsvecs = []
130     for numx in range(3):
131         rhsvec = self.dipoles_xyz[numx].reshape(-1)
132         self.rhsvecs.append(rhsvec)
133
134     def solve_dynamic_direct(self, omega=0.0):
135         # Adapted completely from TDHF.py
136
137         eps_v = self.epsilon[self.nocc:]
138         eps_o = self.epsilon[:self.nocc]
139
140         t = time.time()
141         I = self.mints.ao_eri()
142         v_ijab = np.asarray(self.mints.mo_transform(I, self.Co, self.Co, self.Cv,
143             ↪ self.Cv))
144         v_iajb = np.asarray(self.mints.mo_transform(I, self.Co, self.Cv, self.Co,
145             ↪ self.Cv))
146         print('Integral transform took %.3f seconds\n' % (time.time() - t))
147
148         # Since we are time dependent we need to build the full Hessian:
149         # | A B |      | D S | | x |   | b |
150         # | B A | - w | S -D | | -x | = | -b |
151
152         # Build A and B blocks
153         t = time.time()
154         A11 = np.einsum('ab,ij->iajb', np.diag(eps_v), np.diag(np.ones(self.nocc)))
155         A11 -= np.einsum('ij,ab->iajb', np.diag(eps_o), np.diag(np.ones(self.nvir)))
156         A11 += 2 * v_iajb
157         A11 -= v_ijab.swapaxes(1, 2)
158         A11 *= 2
159
160         B11 = -2 * v_iajb
161         B11 += v_ijab.swapaxes(0, 2)
162         B11 *= 2
163
164         # Reshape and jam it together
165         nov = self.nocc * self.nvir
166         A11.shape = (nov, nov)
167         B11.shape = (nov, nov)
168
169         Hess1 = np.hstack((A11, B11))
170         Hess2 = np.hstack((B11, A11))
171         Hess = np.vstack((Hess1, Hess2))
172
173         S11 = np.zeros_like(A11)
174         D11 = np.zeros_like(B11)
175         S11[np.diag_indices_from(S11)] = 2
176
177         S1 = np.hstack((S11, D11))
178         S2 = np.hstack((D11, -S11))
179         S = np.vstack((S1, S2))

```

```

178     S *= omega
179     print('Hessian formation took %.3f seconds\n' % (time.time() - t))
180
181     t = time.time()
182     Hinv = np.linalg.inv(Hess - S)
183     print('Hessian inversion took %.3f seconds\n' % (time.time() - t))
184
185     self.x = []
186     self.rhsvecs = []
187     for numx in range(3):
188         rhsvec = self.dipoles_xyz[numx].reshape(-1)
189         rhsvec = np.concatenate((rhsvec, -rhsvec))
190         xcomp = Hinv.dot(rhsvec)
191         self.rhsvecs.append(rhsvec)
192         self.x.append(xcomp)
193
194     def solve_static_iterative(self, maxiter=20, conv=1.e-9, use_diis=True):
195
196         # Init JK object
197         jk = psi4.core.JK.build(self.scf_wfn.basisset())
198         jk.initialize()
199
200         # Add blank matrices to the jk object and numpy hooks to C_right
201         npC_right = []
202         for xyz in range(3):
203             jk.C_left_add(self.Co)
204             mC = psi4.core.Matrix(self.nbf, self.nocc)
205             npC_right.append(np.asarray(mC))
206             jk.C_right_add(mC)
207
208         # Build initial guess, previous vectors, diis object, and C_left updates
209         self.x = []
210         x_old = []
211         diis = []
212         ia_denom = - self.epsilon[:self.nocc].reshape(-1, 1) + self.epsilon[self.nocc:]
213         for xyz in range(3):
214             self.x.append(self.dipoles_xyz[xyz] / ia_denom)
215             x_old.append(np.zeros(ia_denom.shape))
216             diis.append(DIIS_helper())
217
218         # Convert Co and Cv to numpy arrays
219         Co = np.asarray(self.Co)
220         Cv = np.asarray(self.Cv)
221
222         print('\nStarting CPHF iterations:')
223         t = time.time()
224         for CPHF_ITER in range(1, maxiter + 1):
225
226             # Update jk's C_right
227             for xyz in range(3):
228                 npC_right[xyz][:] = Cv.dot(self.x[xyz].T)
229
230             # Compute JK objects
231             jk.compute()

```



```

232
233     # Update amplitudes
234     for xyz in range(3):
235         # Build J and K objects
236         J = np.asarray(jk.J()[xyz])
237         K = np.asarray(jk.K()[xyz])
238
239         # Build new guess
240         X = self.dipoles_xyz[xyz].copy()
241         X -= (Co.T).dot(4 * J - K.T - K).dot(Cv)
242         X /= ia_denom
243
244         # DIIS for good measure
245         if use_diis:
246             diis[xyz].add(X, X - x_old[xyz])
247             X = diis[xyz].extrapolate()
248         self.x[xyz] = X.copy()
249
250     # Check for convergence
251     rms = []
252     for xyz in range(3):
253         rms.append(np.max((self.x[xyz] - x_old[xyz]) ** 2))
254         x_old[xyz] = self.x[xyz]
255
256     avg_RMS = sum(rms) / 3
257     max_RMS = max(rms)
258
259     if max_RMS < conv:
260         print('CPHF converged in %d iterations and %.2f seconds.' % (CPHF_ITER,
261             ↪ time.time() - t))
262         self.rhsvecs = []
263         for numx in range(3):
264             rhsvec = self.dipoles_xyz[numx].reshape(-1)
265             self.rhsvecs.append(rhsvec)
266             self.x[numx] = self.x[numx].reshape(-1)
267         break
268
269     print('CPHF Iteration %3d: Average RMS = %3.8f Maximum RMS = %3.8f' %
270         (CPHF_ITER, avg_RMS, max_RMS))
271
272 def solve_dynamic_iterative(self, omega=0.0, maxiter=20, conv=1.e-9, use_diis=True):
273
274     # Init JK object
275     jk = psi4.core.JK.build(self.scf_wfn.basisset())
276     jk.initialize()
277
278     # Add blank matrices to the JK object and NumPy hooks to
279     # C_right; there are 6 sets of matrices to account for X and Y
280     # vectors separately.
281     npC_right = []
282     for xyz in range(6):
283         jk.C_left_add(self.Co)
284         mC = psi4.core.Matrix(self.nbf, self.nocc)
285         npC_right.append(np.asarray(mC))

```

```

285         jk.C_right_add(mC)
286
287     # Build initial guess, previous vectors, diis object, and C_left updates
288     x_l, x_r = [], []
289     x_l_old, x_r_old = [], []
290     diis_l, diis_r = [], []
291     ia_denom_l = self.epsilon[self.nocc:] - self.epsilon[:self.nocc].reshape(-1, 1) -
        ↪ omega
292     ia_denom_r = self.epsilon[self.nocc:] - self.epsilon[:self.nocc].reshape(-1, 1) +
        ↪ omega
293     for xyz in range(3):
294         x_l.append(self.dipoles_xyz[xyz] / ia_denom_l)
295         x_r.append(self.dipoles_xyz[xyz] / ia_denom_r)
296         x_l_old.append(np.zeros(ia_denom_l.shape))
297         x_r_old.append(np.zeros(ia_denom_r.shape))
298         diis_l.append(DIIS_helper())
299         diis_r.append(DIIS_helper())
300
301     # Convert Co and Cv to numpy arrays
302     Co = np.asarray(self.Co)
303     Cv = np.asarray(self.Cv)
304
305     print('\nStarting CPHF iterations:')
306     t = time.time()
307     for CPHF_ITER in range(1, maxiter + 1):
308
309         # Update jk's C_right; ordering is Xx, Xy, Xz, Yx, Yy, Yz
310         for xyz in range(3):
311             npC_right[xyz][:] = Cv.dot(x_l[xyz].T)
312             npC_right[xyz + 3][:] = Cv.dot(x_r[xyz].T)
313
314         # Perform generalized J/K build
315         jk.compute()
316
317         # Update amplitudes
318         for xyz in range(3):
319             # Build J and K objects
320             J_l = np.asarray(jk.J()[xyz])
321             K_l = np.asarray(jk.K()[xyz])
322             J_r = np.asarray(jk.J()[xyz + 3])
323             K_r = np.asarray(jk.K()[xyz + 3])
324
325             # Build new guess
326             X_l = self.dipoles_xyz[xyz].copy()
327             X_r = self.dipoles_xyz[xyz].copy()
328             X_l -= (Co.T).dot(2 * J_l - K_l).dot(Cv)
329             X_r -= (Co.T).dot(2 * J_r - K_r).dot(Cv)
330             X_l /= ia_denom_l
331             X_r /= ia_denom_r
332
333         # DIIS for good measure
334         if use_diis:
335             diis_l[xyz].add(X_l, X_l - x_l_old[xyz])
336             X_l = diis_l[xyz].extrapolate()

```

```

337         diis_r[xyz].add(X_r, X_r - x_r_old[xyz])
338         X_r = diis_r[xyz].extrapolate()
339         x_l[xyz] = X_l.copy()
340         x_r[xyz] = X_r.copy()
341
342     # Check for convergence
343     rms = []
344     for xyz in range(3):
345         rms_l = np.max((x_l[xyz] - x_l_old[xyz]) ** 2)
346         rms_r = np.max((x_r[xyz] - x_r_old[xyz]) ** 2)
347         rms.append(max(rms_l, rms_r))
348         x_l_old[xyz] = x_l[xyz]
349         x_r_old[xyz] = x_r[xyz]
350
351     avg_RMS = sum(rms) / 3
352     max_RMS = max(rms)
353
354     if max_RMS < conv:
355         print('CPHF converged in %d iterations and %.2f seconds.' % (CPHF_ITER,
356             ↪ time.time() - t))
357         self.rhsvecs = []
358         for numx in range(3):
359             rhsvec = self.dipoles_xyz[numx].reshape(-1)
360             self.rhsvecs.append(np.concatenate((rhsvec, -rhsvec)))
361             self.x.append(np.concatenate((x_l[numx].reshape(-1),
362                 x_r[numx].reshape(-1))))
363         break
364
365     print('CPHF Iteration %3d: Average RMS = %3.8f Maximum RMS = %3.8f' %
366         (CPHF_ITER, avg_RMS, max_RMS))
367
368     def form_polarizability(self):
369         self.polar = np.empty((3, 3))
370         for numx in range(3):
371             for numf in range(3):
372                 self.polar[numx, numf] = self.x[numx].dot(self.rhsvecs[numf])
373
374 if __name__ == '__main__':
375     print('\n')
376     print('@test_CPHF running CPHF.py')
377
378     from CPHF import *
379
380     from helper_CPHF import helper_CPHF
381
382     helper = helper_CPHF(mol)
383
384     print('\n')
385     print('@test_CPHF running solve_static_direct')
386
387     helper.solve_static_direct()
388     helper.form_polarizability()
389     assert np.allclose(polar, helper.polar, rtol=0, atol=1.e-5)

```

```

390     print('\n')
391     print('@test_CPHF running solve_static_iterative')
392
393     helper.solve_static_iterative()
394     helper.form_polarizability()
395     assert np.allclose(polar, helper.polar, rtol=0, atol=1.e-5)
396
397     f = 0.0
398
399     print('\n')
400     print('@test_CPHF running solve_dynamic_direct ({}).format(f))
401
402     helper.solve_dynamic_direct(omega=f)
403     helper.form_polarizability()
404     assert np.allclose(polar, helper.polar, rtol=0, atol=1.e-5)
405
406     print('\n')
407     print('@test_CPHF running solve_dynamic_iterative ({}).format(f))
408
409     helper.solve_dynamic_iterative(omega=f)
410     helper.form_polarizability()
411     assert np.allclose(polar, helper.polar, rtol=0, atol=1.e-5)
412
413     f = 0.0773178
414     ref = np.array([
415         [8.19440121, 0.00000000, 0.00000000],
416         [0.00000000, 12.75967150, 0.00000000],
417         [0.00000000, 0.00000000, 10.25213939]
418     ])
419
420     print('\n')
421     print('@test_CPHF running solve_dynamic_direct ({}).format(f))
422
423     helper.solve_dynamic_direct(omega=f)
424     helper.form_polarizability()
425     assert np.allclose(ref, helper.polar, rtol=0, atol=1.e-5)
426
427     print('\n')
428     print('@test_CPHF running solve_dynamic_iterative ({}).format(f))
429
430     helper.solve_dynamic_iterative(omega=f)
431     helper.form_polarizability()
432     assert np.allclose(ref, helper.polar, rtol=0, atol=1.e-5)
433
434     """
435     1     """
436     2     A reference implementation to compute the first dipole
437     3     hyperpolarizability  $\beta$  from a restricted HF reference using the
438     4      $2n+1$  rule from perturbation theory.
439     5
440     6     References:
441     7     Equations taken from [Karna:1991:487], http://dx.doi.org/10.1002/jcc.540120409
442     8     """
443     9
444     10    __authors__ = "Eric J. Berquist"

```

```

11 __credits__ = ["Eric J. Berquist"]
12
13 __copyright__ = "(c) 2014-2017, The Psi4NumPy Developers"
14 __license__ = "BSD-3-Clause"
15 __date__ = "2017-08-26"
16
17 from itertools import permutations, product
18
19 import numpy as np
20 np.set_printoptions(precision=5, linewidth=200, suppress=True)
21 import psi4
22 from helper_CPHF import helper_CPHF
23
24 # Memory for Psi4 in GB
25 psi4.set_memory('2 GB')
26 psi4.core.set_output_file("output.dat", False)
27
28 mol = psi4.geometry("""
29 O
30 H 1 1.1
31 H 1 1.1 2 104
32 symmetry c1
33 """)
34
35 # Set options for CPHF
36 psi4.set_options({"basis": "aug-cc-pvdz",
37                  "scf_type": "direct",
38                  "df_scf_guess": False,
39                  "e_convergence": 1e-9,
40                  "d_convergence": 1e-9})
41
42 # Compute the (first) hyperpolarizability corresponding to static
43 # fields, beta(0;0,0), eqns. (IV-2a) and (VII-4).
44
45 helper = helper_CPHF(mol)
46 # For the 2n+1 rule, the quadratic response starting quantities must
47 # come from linear response.
48 helper.run()
49
50 na = np.newaxis
51 moenergies = helper.epsilon
52 C = np.asarray(helper.C)
53 Co = helper.Co
54 Cv = helper.Cv
55 nbf, norb = C.shape
56 nocc = Co.shape[1]
57 nvir = norb - nocc
58 nov = nocc * nvir
59 x = np.asarray(helper.x)
60 ncomp = x.shape[0]
61 integrals_ao = np.asarray([np.asarray(dipole_ao_component)
62                             for dipole_ao_component in helper.tmp_dipoles])
63
64 # form full MO-basis dipole integrals

```

```

65 integrals_mo = np.empty(shape=(ncomp, norb, norb))
66 for i in range(ncomp):
67     integrals_mo[i] = (C.T).dot(integrals_ao[i]).dot(C)
68
69 # repack response vectors to [norb, norb]; 1/2 is due to X + Y
70 U = np.zeros_like(integrals_mo)
71 for i in range(ncomp):
72     U[i, :nocc, nocc:] = 0.5 * x[i].reshape(nocc, nvir)
73     U[i, nocc:, :nocc] = -0.5 * x[i].reshape(nocc, nvir).T
74
75 # form G matrices from perturbation and generalized Fock matrices; do
76 # one more Fock build for each response vector
77 jk = psi4.core.JK.build(helper.scf_wfn.basisset())
78 jk.initialize()
79 G = np.empty_like(U)
80 R = psi4.core.Matrix(nbf, nocc)
81 npR = np.asarray(R)
82 for i in range(ncomp):
83     V = integrals_mo[i]
84
85     # eqn. (III-1b) Note: this simplified handling of the response
86     # vector transformation for the Fock build is insufficient for
87     # frequency-dependent response.
88     jk.C_clear()
89     # Psi4's JK builders don't take a density, but a left set of
90     # coefficients with shape [nbf, nocc] and a right set of
91     # coefficients with shape [nbf, nocc]. Because the response vector
92     # describes occ -> vir transitions, we perform ([nocc, nvir] *
93     # [nbf, nvir]^T)^T.
94     L = Co
95     npR[:] = x[i].reshape(nocc, nvir).dot(np.asarray(Cv).T).T
96     jk.C_left_add(L)
97     jk.C_right_add(R)
98     jk.compute()
99     # 1/2 is due to X + Y
100    J = 0.5 * np.asarray(jk.J()[0])
101    K = 0.5 * np.asarray(jk.K()[0])
102
103    # eqn. (21b)
104    F = (C.T).dot(4 * J - K.T - K).dot(C)
105    G[i] = V + F
106
107 # form epsilon matrices, eqn. (34)
108 E = G.copy()
109 omega = 0
110 for i in range(ncomp):
111     eoU = (moenergies[... , na] + omega) * U[i]
112     Ue = U[i] * moenergies[na]
113     E[i] += (eoU - Ue)
114
115 # Assume some symmetry and calculate only part of the tensor.
116 # eqn. (VII-4)
117 hyperpolarizability = np.zeros(shape=(6, 3))
118 off1 = [0, 1, 2, 0, 0, 1]

```

```

119 off2 = [0, 1, 2, 1, 2, 2]
120 for r in range(6):
121     b = off1[r]
122     c = off2[r]
123     for a in range(3):
124         t11 = 2 * np.trace(U[a].dot(G[b]).dot(U[c])[:nocc, :nocc])
125         t12 = 2 * np.trace(U[a].dot(G[c]).dot(U[b])[:nocc, :nocc])
126         t13 = 2 * np.trace(U[c].dot(G[a]).dot(U[b])[:nocc, :nocc])
127         tr1 = np.trace(U[c].dot(U[b]).dot(E[a])[:nocc, :nocc])
128         tr2 = np.trace(U[b].dot(U[c]).dot(E[a])[:nocc, :nocc])
129         tr3 = np.trace(U[c].dot(U[a]).dot(E[b])[:nocc, :nocc])
130         tr4 = np.trace(U[a].dot(U[c]).dot(E[b])[:nocc, :nocc])
131         tr5 = np.trace(U[b].dot(U[a]).dot(E[c])[:nocc, :nocc])
132         tr6 = np.trace(U[a].dot(U[b]).dot(E[c])[:nocc, :nocc])
133         t1 = t11 + t12 + t13
134         tr = tr1 + tr2 + tr3 + tr4 + tr5 + tr6
135         hyperpolarizability[r, a] = -2 * (t1 - tr)
136
137 ref_static = np.array([
138     [ 0.00000001,  0.00000000,  0.22843772],
139     [ 0.00000000,  0.00000000, -25.35476040],
140     [ 0.00000000,  0.00000000, -10.84023375],
141     [ 0.00000000,  0.00000000,  0.00000000],
142     [ 0.22843772,  0.00000000,  0.00000000],
143     [ 0.00000000, -25.35476040,  0.00000000]
144 ])
145 assert np.allclose(ref_static, hyperpolarizability, rtol=0.0, atol=1.0e-3)
146 print('\nFirst dipole hyperpolarizability (static):')
147 print(hyperpolarizability)
148
149 # Compute the (first) hyperpolarizability corresponding to
150 # second-harmonic generation, beta(-2w;w,w), eqns. (IV-2c) and
151 # (VII-1). Because two different frequencies are involved, the linear
152 # response equations must be solved twice.
153
154 print('Setting up for second-harmonic generation (SHG) calculation...')
155 # In SHG, the first frequency is doubled to obtain the second
156 # frequency. All variables containing '1' correspond to the first
157 # (set) frequency, and all variables containing '2' correspond to the
158 # second (doubled) frequency.
159 f1 = 0.0773178
160 f2 = 2 * f1
161
162 print('\nForming response vectors for {} a.u.'.format(f1))
163 helper1 = helper_CPHF(mol)
164 helper1.solve_dynamic_direct(omega=f1)
165 helper1.form_polarizability()
166 print(helper1.polar)
167 print('\nForming response vectors for {} a.u.'.format(f2))
168 helper2 = helper_CPHF(mol)
169 helper2.solve_dynamic_direct(omega=f2)
170 helper2.form_polarizability()
171 print(helper2.polar)
172

```

```

173  rspvecs1 = helper1.x
174  rspvecs2 = helper2.x
175
176  # repack response vectors to [norb, norb]
177  U1 = np.zeros_like(integrals_mo)
178  U2 = np.zeros_like(integrals_mo)
179  for i in range(ncomp):
180      U1[i, :nocc, nocc:] = rspvecs1[i][nov:].reshape(nocc, nvir)
181      U1[i, nocc:, :nocc] = rspvecs1[i][:nov].reshape(nocc, nvir).T
182      U2[i, :nocc, nocc:] = rspvecs2[i][nov:].reshape(nocc, nvir)
183      U2[i, nocc:, :nocc] = rspvecs2[i][:nov].reshape(nocc, nvir).T
184
185  G1 = np.empty_like(U1)
186  G2 = np.empty_like(U2)
187  R1_l = psi4.core.Matrix(nbf, nocc)
188  R1_r = psi4.core.Matrix(nbf, nocc)
189  R2_l = psi4.core.Matrix(nbf, nocc)
190  R2_r = psi4.core.Matrix(nbf, nocc)
191  npR1_l = np.asarray(R1_l)
192  npR1_r = np.asarray(R1_r)
193  npR2_l = np.asarray(R2_l)
194  npR2_r = np.asarray(R2_r)
195  jk.C_clear()
196  jk.C_left_add(Co)
197  jk.C_right_add(R1_l)
198  jk.C_left_add(Co)
199  jk.C_right_add(R1_r)
200  jk.C_left_add(Co)
201  jk.C_right_add(R2_l)
202  jk.C_left_add(Co)
203  jk.C_right_add(R2_r)
204  nCo = np.asarray(Co)
205  # Do 4 Fock builds at a time: X/Y vectors for both frequencies; loop
206  # over operator components
207  for i in range(3):
208      V = integrals_mo[i]
209
210      x1 = U1[i, :nocc, :]
211      y1 = U1[i, :, :nocc]
212      x2 = U2[i, :nocc, :]
213      y2 = U2[i, :, :nocc]
214      npR1_l[:] = C.dot(x1.T)
215      npR1_r[:] = C.dot(y1)
216      npR2_l[:] = C.dot(x2.T)
217      npR2_r[:] = C.dot(y2)
218
219      jk.compute()
220
221      J1_l = -np.asarray(jk.J()[0])
222      K1_l = -np.asarray(jk.K()[0])
223      J1_r = np.asarray(jk.J()[1])
224      K1_r = np.asarray(jk.K()[1])
225      J2_l = -np.asarray(jk.J()[2])
226      K2_l = -np.asarray(jk.K()[2])

```



```

227     J2_r = np.asarray(jk.J()[3])
228     K2_r = np.asarray(jk.K()[3])
229     J1 = J1_l + J1_r
230     J2 = J2_l + J2_r
231     K1 = K1_l + K1_r.T
232     K2 = K2_l + K2_r.T
233
234     F1 = (C.T).dot(2 * J1 - K1).dot(C)
235     F2 = (C.T).dot(2 * J2 - K2).dot(C)
236     G1[i, ...] = V + F1
237     G2[i, ...] = V + F2
238
239     # form epsilon matrices, eqn. (34), one for each frequency
240     E1 = G1.copy()
241     E2 = G2.copy()
242     for i in range(ncomp):
243         eoU1 = (moenergies[..., na] + f1) * U1[i]
244         Ue1 = U1[i] * moenergies[na]
245         E1[i] += (eoU1 - Ue1)
246         eoU2 = (moenergies[..., na] + f2) * U2[i]
247         Ue2 = U2[i] * moenergies[na]
248         E2[i] += (eoU2 - Ue2)
249
250     # Assume some symmetry and calculate only part of the tensor.
251
252     hyperpolarizability = np.zeros(shape=(6, 3))
253     for r in range(6):
254         b = off1[r]
255         c = off2[r]
256         for a in range(3):
257             t11 = np.trace(U2[a].T.dot(G1[b]).dot(U1[c])[:nocc, :nocc])
258             t12 = np.trace(U1[c].dot(G1[b]).dot(U2[a].T)[:nocc, :nocc])
259             t13 = np.trace(U2[a].T.dot(G1[c]).dot(U1[b])[:nocc, :nocc])
260             t14 = np.trace(U1[b].dot(G1[c]).dot(U2[a].T)[:nocc, :nocc])
261             t15 = np.trace(U1[c].dot(-G2[a].T).dot(U1[b])[:nocc, :nocc])
262             t16 = np.trace(U1[b].dot(-G2[a].T).dot(U1[c])[:nocc, :nocc])
263             tr1 = np.trace(U1[c].dot(U1[b]).dot(-E2[a].T)[:nocc, :nocc])
264             tr2 = np.trace(U1[b].dot(U1[c]).dot(-E2[a].T)[:nocc, :nocc])
265             tr3 = np.trace(U1[c].dot(U2[a].T).dot(E1[b])[:nocc, :nocc])
266             tr4 = np.trace(U2[a].T.dot(U1[c]).dot(E1[b])[:nocc, :nocc])
267             tr5 = np.trace(U1[b].dot(U2[a].T).dot(E1[c])[:nocc, :nocc])
268             tr6 = np.trace(U2[a].T.dot(U1[b]).dot(E1[c])[:nocc, :nocc])
269             t1 = t11 + t12 + t13 + t14 + t15 + t16
270             tr = tr1 + tr2 + tr3 + tr4 + tr5 + tr6
271             hyperpolarizability[r, a] = 2 * (t1 - tr)
272
273     # pylint: disable=C0326
274     ref = np.array([
275         [ 0.00000000,  0.00000000,  1.92505358],
276         [ 0.00000000,  0.00000000, -31.33652886],
277         [ 0.00000000,  0.00000000, -13.92830863],
278         [ 0.00000000,  0.00000000,  0.00000000],
279         [-1.80626084,  0.00000000,  0.00000000],
280         [ 0.00000000, -31.13504192,  0.00000000]

```

```

281 ])
282 ref_avgs = np.array([0.00000000, 0.00000000, 45.69300223])
283 ref_avg = 45.69300223
284
285 thresh = 1.0e-2
286 # assert np.all(np.abs(ref - hyperpolarizability) < thresh)
287
288 print('hyperpolarizability: SHG, (-{}; {}, {}), symmetry-unique components'.format(f2,
↵ f1, f1))
289 print(hyperpolarizability)
290 print('ref')
291 print(ref)
292
293 # Transpose all frequency-doubled quantities (+2w) to get -2w.
294
295 for i in range(ncomp):
296     U2[i] = U2[i].T
297     G2[i] = -G2[i].T
298     E2[i] = -E2[i].T
299
300 # Assume some symmetry and calculate only part of the tensor. This
301 # time, work with the in-place manipulated quantities (this tests
302 # their correctness).
303
304 mU = (U2, U1)
305 mG = (G2, G1)
306 me = (E2, E1)
307
308 hyperpolarizability = np.zeros(shape=(6, 3))
309 off1 = [0, 1, 2, 0, 0, 1]
310 off2 = [0, 1, 2, 1, 2, 2]
311 for r in range(6):
312     b = off1[r]
313     c = off2[r]
314     for a in range(3):
315         t11 = np.trace(mU[0][a].dot(mG[1][b]).dot(mU[1][c])[:nocc, :nocc])
316         t12 = np.trace(mU[1][c].dot(mG[1][b]).dot(mU[0][a])[:nocc, :nocc])
317         t13 = np.trace(mU[0][a].dot(mG[1][c]).dot(mU[1][b])[:nocc, :nocc])
318         t14 = np.trace(mU[1][b].dot(mG[1][c]).dot(mU[0][a])[:nocc, :nocc])
319         t15 = np.trace(mU[1][c].dot(mG[0][a]).dot(mU[1][b])[:nocc, :nocc])
320         t16 = np.trace(mU[1][b].dot(mG[0][a]).dot(mU[1][c])[:nocc, :nocc])
321         tr1 = np.trace(mU[1][c].dot(mU[1][b]).dot(me[0][a])[:nocc, :nocc])
322         tr2 = np.trace(mU[1][b].dot(mU[1][c]).dot(me[0][a])[:nocc, :nocc])
323         tr3 = np.trace(mU[1][c].dot(mU[0][a]).dot(me[1][b])[:nocc, :nocc])
324         tr4 = np.trace(mU[0][a].dot(mU[1][c]).dot(me[1][b])[:nocc, :nocc])
325         tr5 = np.trace(mU[1][b].dot(mU[0][a]).dot(me[1][c])[:nocc, :nocc])
326         tr6 = np.trace(mU[0][a].dot(mU[1][b]).dot(me[1][c])[:nocc, :nocc])
327         t1 = [t11, t12, t13, t14, t15, t16]
328         tr = [tr1, tr2, tr3, tr4, tr5, tr6]
329         hyperpolarizability[r, a] = 2 * (sum(t1) - sum(tr))
330
331 assert np.all(np.abs(ref - hyperpolarizability) < thresh)
332
333 # Assume no symmetry and calculate the full tensor.

```

```

334
335 hyperpolarizability_full = np.zeros(shape=(3, 3, 3))
336
337 # components x, y, z
338 for ip, p in enumerate(list(product(range(3), range(3), range(3)))):
339     a, b, c = p
340     t1, tr = [], []
341     # 1st tuple -> index a, b, c (*not* x, y, z!)
342     # 2nd tuple -> index frequency (0 -> -2w, 1 -> +w)
343     for iq, q in enumerate(list(permutations(zip(p, (0, 1, 1)), 3))):
344         d, e, f = q
345         tlp = (mU[d[1]][d[0]]).dot(mG[e[1]][e[0]]).dot(mU[f[1]][f[0]])
346         tle = np.trace(tlp[:nocc, :nocc])
347         t1.append(tle)
348         trp = (mU[d[1]][d[0]]).dot(mU[e[1]][e[0]]).dot(me[f[1]][f[0]])
349         tre = np.trace(trp[:nocc, :nocc])
350         tr.append(tre)
351     hyperpolarizability_full[a, b, c] = 2 * (sum(t1) - sum(tr))
352 print('hyperpolarizability: SHG, (-{}; {}, {}), full tensor'.format(f2, f1, f1))
353 print(hyperpolarizability_full)
354
355 # Check that the elements of the reduced and full tensors are
356 # equivalent.
357
358 for r in range(6):
359     b = off1[r]
360     c = off2[r]
361     for a in range(3):
362         diff = hyperpolarizability[r, a] - hyperpolarizability_full[a, b, c]
363         assert abs(diff) < 1.0e-14

```

## 6.9.2 Hyperpolarizability Tutorial

```

1  """Tutorial: SCF first hyperpolarizability"""
2
3  __author__      = "Eric J. Berquist"
4  __credit__     = ["Eric J. Berquist"]
5
6  __copyright__  = "(c) 2014-2017, The Psi4NumPy Developers"
7  __license__    = "BSD-3-Clause"
8  __date__       = "2017-12-19"

```

**6.9.2.1 Introduction** In [Tutorial 6a](#), the calculation of linear response properties from analytic derivative theory is presented, the foundation of which are the coupled-perturbed Hartree–Fock (CPHF) or coupled-perturbed self-consistent field (CPSCF) equations. Starting from analytic derivative theory provides a convenient physical picture: how does the total energy of a system change under the influence of one or more internal or external per-

turbations? Continuing the case of an external electric field, the total energy of a system can be represented with a series expansion:

$$E(\mathbf{E}) = \sum_{n=0}^{\infty} \frac{1}{n!} E^{(n)}(\mathbf{a}) \cdot (\mathbf{E} - \mathbf{a})^n, \quad (6.7)$$

where the electric field is  $\mathbf{E} = \vec{E} = (E_x, E_y, E_z)$  and  $\mathbf{a}$  is the expansion point. In practice, we always expand around  $\mathbf{a} = \mathbf{0}$ , so it is a Maclaurin series:

$$E(\mathbf{E}) = \sum_{n=0}^{\infty} \frac{1}{n!} \mathbf{E}^{(n)}(\mathbf{0}) \cdot \mathbf{E}^n. \quad (6.8)$$

Expanding the above to the first 4 explicit terms gives

$$E(\mathbf{E}) \approx E^{(0)}(\mathbf{0}) + \mathbf{E}^{(1)}(\mathbf{0}) \cdot \mathbf{E} + \frac{1}{2} \mathbf{E}^{(2)}(\mathbf{0}) \cdot \mathbf{E}^2 + \frac{1}{6} \mathbf{E}^{(3)}(\mathbf{0}) \cdot \mathbf{E}^3, \quad (6.9)$$

where we identify

$$E^{(0)} \rightarrow \text{the unperturbed ground-state energy} \quad (6.10)$$

$$E_a^{(1)} \rightarrow -\mu_a, \text{ the dipole moment} \quad (6.11)$$

$$E_{ab}^{(2)} \rightarrow -\alpha_{ab}, \text{ the polarizability} \quad (6.12)$$

$$E_{abc}^{(3)} \rightarrow -\beta_{abc}, \text{ the first hyperpolarizability} \quad (6.13)$$

The first hyperpolarizability is the leading-order term that describes the *nonlinear* response of a system to an external electric field. Each term in the series expansion increases the rank of the coefficient by one: the ground-state energy is a scalar, the dipole is a length 3 vector, the polarizability is a 3-by-3 matrix, and the first hyperpolarizability is a 3-by-3-by-3 tensor.

Translated into the language of analytic derivative theory, is it represented as

$$\beta_{abc} = \left. \frac{\partial^3 E}{\partial E_a \partial E_b \partial E_c} \right|_{\mathbf{E}=\mathbf{0}}, \quad (6.14)$$

though it is not yet clear how to take derivatives of the energy beyond what is presented in [tutorial 6a](#). Additionally, nothing has been stated about time dependence; everything to this point has been the static case, where the strength of fields do not vary with time. We will first incorporate time dependence, and equations for derivative theory will result.

**6.9.2.2 Notation** Before going further, some notational conventions should be mentioned. When used as field indices,  $a, b, c, \dots \in \{x, y, z\}$ , the three Cartesian directions.

For matrix indices,  $\mu, \nu, \lambda, \sigma, \dots$  label atomic orbitals (AOs)/basis functions,  $i, j, k, l, \dots$  label occupied molecular orbitals (MOs),  $a, b, c, d, \dots$  label unoccupied/virtual MOs, and  $p, q, r, s, \dots$  label all MOs. Einstein summation is used, so repeated indices are contracted over.

**6.9.2.3 Derivation** Again, write the total Hamiltonian as the sum of unperturbed and perturbed components

$$\begin{aligned}\hat{H}(\mathbf{E}, t) &= \hat{H}^{(0)} + \hat{V}(\mathbf{E}, t) \\ \hat{V}(\mathbf{E}, t) &= -\boldsymbol{\mu} \cdot \mathbf{E}(e^{\pm i\omega t} + 1)\end{aligned}\tag{Karna 2}$$

where part of the external field now oscillates with some characteristic frequency  $\omega$ . This can be incorporated into the time-dependent Schrödinger equation, which for a stationary state obeys

$$\left[ \hat{H}^{(0)} + \hat{V}(\mathbf{E}, t) - i \frac{\partial}{\partial t} \right] \psi(t) = 0,\tag{Karna 3}$$

$$FC - i \frac{\partial}{\partial t} SC = SC\epsilon,\tag{Karna 5}$$

$$\frac{\partial}{\partial t} C^\dagger SC = 0,\tag{Karna 6}$$

where the full definition of the Fock matrix is

$$F_{\mu\nu} = h_{\mu\nu} + D_{\lambda\sigma}[2J_{\mu\nu\lambda\sigma} - K_{\mu\nu\lambda\sigma}]\tag{Karna 9}$$

and the density matrix is defined as

$$D_{\mu\nu} = C_{\mu p} n_{pq} C_{\nu q}^\dagger,\tag{Karna 10}$$

where the diagonal occupation number matrix  $n_{ii} = 2$  and  $n_{aa} = 0$  for RHF.

In general, the MO coefficients are perturbation- and time-dependent, but the basis functions themselves are not. This means that when the series expansion for the perturbation

above is performed on other quantities, only  $F$ ,  $C$ ,  $\epsilon$ , and  $D$  are affected. For example, the Lagrangian multiplier matrix  $\epsilon$  can be expanded as

$$\epsilon(\mathbf{E}) = \epsilon^0 + E_a \epsilon^a + \frac{1}{2!} E_a E_b \epsilon^{ab} + \frac{1}{3!} E_a E_b E_c \epsilon^{abc} + \dots \quad (\text{Karna 17c})$$

where  $a, b, c, \dots \in \{x, y, z\}$ , and

$$\epsilon^a = e^{\pm i\omega t} \epsilon^a(\pm\omega) + \epsilon^a(0), \quad (\text{Karna 19a})$$

$$\epsilon^{ab} = e^{\pm 2i\omega t} \epsilon^{ab}(\pm\omega, \pm\omega) + e^{\pm i\omega t} \{ \epsilon^{ab}(0, \pm\omega) + \epsilon^{ab}(\pm\omega, 0) \} + \epsilon^{ab}(\pm\omega, \mp\omega) + \epsilon^{ab}(0, 0), \quad (\text{Karna 19b})$$

$$\begin{aligned} \epsilon^{abc} = & e^{\pm 3i\omega t} \epsilon^{abc}(\pm\omega, \pm\omega, \pm\omega) + e^{\pm 2i\omega t} \{ \epsilon^{abc}(0, \pm\omega, \pm\omega) + \epsilon^{abc}(\pm\omega, 0, \pm\omega) + \epsilon^{abc}(\pm\omega, \pm\omega, 0) \} \\ & + e^{\pm i\omega t} \{ \epsilon^{abc}(\pm\omega, \pm\omega, \mp\omega) + \epsilon^{abc}(\pm\omega, \mp\omega, \pm\omega) + \epsilon^{abc}(\mp\omega, \pm\omega, \pm\omega) \} \\ & + e^{\pm i\omega t} \{ \epsilon^{abc}(0, 0, \pm\omega) + \epsilon^{abc}(0, \pm\omega, 0) + \epsilon^{abc}(\pm\omega, 0, 0) \} \\ & + \{ \epsilon^{abc}(0, \pm\omega, \mp\omega) + \epsilon^{abc}(\pm\omega, 0, \mp\omega) + \epsilon^{abc}(\pm\omega, \mp\omega, 0) \} + \epsilon^{abc}(0, 0, 0), \end{aligned} \quad (\text{Karna 19c})$$

showing that each order of the expansion consists of all possible phase combinations. For the first hyperpolarizability, only quantities with at most two field indices are required. Each permutationally unique subterm of the expansion corresponds to a different physical

observable:

$$(0) \rightarrow \text{static polarizability} \rightarrow \alpha(0; 0) = -Tr[H^a D^b(0)]$$

(Karna IV-1a)

$$(\pm\omega) \rightarrow \text{dynamic polarizability} \rightarrow \alpha(\mp\omega; \pm\omega) = -Tr[H^a D^b(\pm\omega)]$$

(Karna IV-1b)

$$(0, 0) \rightarrow \text{static (first) hyperpolarizability} \rightarrow \beta(0; 0, 0) = -Tr[H^a D^{bc}(0, 0)]$$

(Karna IV-2a)

$$(0, \pm\omega) \rightarrow \text{electrooptic Pockels effect (EOPE)} \rightarrow \beta(\mp\omega; 0, \pm\omega) = -Tr[H^a D^{bc}(0, \pm\omega)]$$

(Karna IV-2b)

$$(\pm\omega, \pm\omega) \rightarrow \text{second harmonic generation (SHG)} \rightarrow \beta(\mp 2\omega; \pm\omega, \pm\omega) = -Tr[H^a D^{bc}(\pm\omega, \pm\omega)]$$

(Karna IV-2c)

$$(\pm\omega, \mp\omega) \rightarrow \text{optical rectification} \rightarrow \beta(0; \pm\omega, \mp\omega) = -Tr[H^a D^{bc}(\pm\omega, \mp\omega)]$$

(Karna IV-2d)

where each property is calculated as the trace over the AO-basis dipole matrices  $H^a$  with the appropriate perturbed density. The task now comes down to calculating the necessary perturbed density for the phenomenon of interest. The second-order densities required for the four different first hyperpolarizabilities are

$$D^{ab}(\pm\omega, \pm\omega) = C^{ab}(\pm\omega, \pm\omega)nC^{0\dagger} + C^a(\pm\omega)nC^{b\dagger}(\mp\omega) \\ + C^b(\pm\omega)nC^{a\dagger}(\mp\omega) + C^0nC^{ab\dagger}(\mp\omega, \mp\omega),$$

(Karna III-2a)

$$D^{ab}(0, \pm\omega) = C^{ab}(0, \pm\omega)nC^{0\dagger} + C^a(0)nC^{b\dagger}(\mp\omega) \\ + C^b(\pm\omega)nC^{a\dagger}(0) + C^0nC^{ab\dagger}(0, \mp\omega),$$

(Karna III-2b)

$$D^{ab}(\pm\omega, \mp\omega) = C^{ab}(\pm\omega, \mp\omega)nC^{0\dagger} + C^a(\pm\omega)nC^{b\dagger}(\pm\omega) \\ + C^b(\mp\omega)nC^{a\dagger}(\mp\omega) + C^0nC^{ab\dagger}(\mp\omega, \pm\omega),$$

(Karna III-2c)

$$D^{ab}(0, 0) = C^{ab}(0, 0)nC^{0\dagger} + C^a(0)nC^{b\dagger}(0) \\ + C^b(0)nC^{a\dagger}(0) + C^0nC^{ab\dagger}(0, 0).$$

(Karna III-2d)

Table 37: Orders of MO coefficient derivatives and rotation matrices needed for energy derivatives following Wigner’s  $2n + 1$  rule. Reproduced from Ref. [7].

	CI: MO/CI space	MCSCF: MO/CI space	RHF: MO space
Energy, $E$	$C_{\mu}^i, C_I$	$C_{\mu}^i, C_I$	$C_{\mu}^i$
First Derivative, $\frac{\partial E}{\partial a}$	$U^a, C_I$	$C_{\mu}^i, C_I$	$C_{\mu}^i$
Second Derivative, $\frac{\partial^2 E}{\partial a \partial b}$	$U^{ab}, \frac{\partial C_I}{\partial a}$	$U^a, \frac{\partial C_I}{\partial a}$	$U^a$
Third Derivative, $\frac{\partial^3 E}{\partial a \partial b \partial c}$	$U^{abc}, \frac{\partial C_I}{\partial a}$	$U^a, \frac{\partial C_I}{\partial a}$	$U^a$
Fourth Derivative, $\frac{\partial^4 E}{\partial a \partial b \partial c \partial d}$	$U^{abcd}, \frac{\partial^2 C_I}{\partial a \partial b}$	$U^{ab}, \frac{\partial^2 C_I}{\partial a \partial b}$	$U^{ab}$
Fifth Derivative, $\frac{\partial^5 E}{\partial a \partial b \partial c \partial d \partial e}$	$U^{abcde}, \frac{\partial^2 C_I}{\partial a \partial b}$	$U^{ab}, \frac{\partial^2 C_I}{\partial a \partial b}$	$U^{ab}$

Already a few important insights about the equations are revealed: Each perturbation index always carries its respective frequency, and the positive and negative frequencies are related by the Hermitian adjoint (except for  $C(-\omega) = -C^0 U^\dagger(+\omega)$ , Karna eq. 40). We also see the appearance of terms like  $C^{ab}$ , which will require  $U^{ab}$  originating from the second-order CPHF. Computationally, this is undesirable due to the increased number of iterative calculations that must be performed, so we borrow a trick that most prominently appears in perturbation theory.

#### 6.9.2.4 Wigner’s $2n + 1$ rule

From Schaefer,<sup>7</sup> page 25:

When the wavefunction is determined up to the  $n$ th order, the expectation value (electronic energy) of the the system is resolved, according to the results of perturbation theory, up to the  $(2n + 1)$ st order. This principle is called Wigner’s  $2n + 1$  theorem.<sup>21,22</sup>

Since the first hyperpolarizability is calculated as a third derivative of the energy, perturbed coefficients with only one field index should be required. From table 37, we can also see why SCF gradients ( $\frac{\partial E}{\partial R_A}$ , where  $R_A$  is the  $A$ -th Cartesian component of nucleus  $R$ ) avoid the need to solve for  $U$  matrices.



**6.9.2.5 Final Expressions** To this point, most work has been in the AO basis, but it is conceptually easier to work in the MO basis, in particular due to the use of the  $\epsilon$  equations

$$\epsilon^a(\pm\omega) = G^a(\pm\omega) + \epsilon^0 U^a(\pm\omega) - U^a(\pm\omega)\epsilon^0 \pm \omega U^a(\pm\omega), \quad (\text{Karna 34})$$

where the  $G$  matrices are the MO-basis Fock matrices

$$G^{ab\dots} = C^{0\dagger} F^{ab\dots} C^0, \quad (6.15)$$

and the  $U$  matrices are the MO-basis perturbation parameters

$$C^{ab\dots} = C^0 U^{ab\dots}, \quad (6.16)$$

which will be discussed in the implementation. The final expression for the static hyperpolarizability is

$$\begin{aligned} \beta_{abc}(0; 0, 0) = & Tr[n\{U^a(0)G^b(0)U^c(0) + U^c(0)G^b(0)U^a(0) + U^b(0)G^c(0)U^a(0) \\ & + U^a(0)G^c(0)U^b(0) + U^c(0)G^a(0)U^b(0) + U^b(0)G^a(0)U^c(0)\}] \\ & - Tr[n\{U^a(0)U^c(0)\epsilon^b(0) + U^c(0)U^a(0)\epsilon^b(0) + U^b(0)U^a(0)\epsilon^c(0) \\ & + U^a(0)U^b(0)\epsilon^c(0) + U^c(0)U^b(0)\epsilon^a(0) + U^b(0)U^c(0)\epsilon^a(0)\}]. \end{aligned} \quad (\text{Karna VII-4})$$

By noticing that each term corresponds to a unique permutation of the field indices, it can be rewritten as

$$\beta_{abc} = Tr \left[ n \sum \mathcal{P}(d, e, f) U^d G^e U^f \right] - Tr \left[ n \sum \mathcal{P}(d, e, f) U^d U^e \epsilon^f \right], \quad (6.17)$$

where the permutation indices are initially assigned as  $d = a, e = b, f = c$ . The frequency notation has also been dropped, since each  $abc$  (and therefore each  $def$ ) will always carry the appropriate field index, making this the most general form of the first hyperpolarizability. If the indices  $abc$  are also permuted, then all 27 components of the first hyperpolarizability tensor will be computed.

**6.9.2.6 Computational Procedure** The basic quantities we need are the matrices  $C$ ,  $\mu$ ,  $F$ ,  $\epsilon$ , and  $U$ . The MO coefficients  $C$  are already obtained from the ground-state calculation, along with  $\epsilon^0$  (the MO energies) and  $F^0$  (the AO-basis Fock matrix). The dipole matrices  $\mu$  are needed for the linear response (polarizability) calculation, which results in response vectors that compose the off-diagonal blocks of the  $U^a$  matrix.  $G^a$  is obtained from  $F^a$ , which comes from performing a single Fock build with the perturbed density  $D^a$ . Finally,  $\epsilon^a$  can be constructed.

Although the expressions so far are general for any frequency and first-order non-linear optical response, the tutorial implementation will cover the static case. For second-harmonic generation, see the reference implementation.

```

1 import numpy as np
2 np.set_printoptions(3, linewidth=100, suppress=True)      # when we inspect the
   ↪ vectors/matrices,
3
   # use a prettier format for
   ↪ printing
4 import psi4

```

The energy and density convergence criteria are tightened from defaults, as response properties are sensitive to the quality of the ground-state wavefunction.

```

1 mol = psi4.geometry('''
2     O
3     H 1 0.9435
4     H 1 0.9435 2 105.9443
5     symmetry c1
6 ''')
7 psi4.set_options({
8     "basis": "aug-cc-pVDZ",
9     "scf_type": "direct",
10    "df_scf_guess": False,
11    "e_convergence": 1e-9,
12    "d_convergence": 1e-9,
13 })

1 # This is to enable testing outside of the notebook environment.
2 import sys
3 try:
4     get_ipython()
5     sys.path.append('../..//Response-Theory/Self-Consistent-Field')
6 except NameError:
7     import os.path
8     dirname = os.path.dirname(os.path.abspath(__file__))
9     sys.path.append(os.path.join(dirname, '../..//Response-Theory/Self-Consistent-Field'))
10
11 from helper_CPHF import helper_CPHF

```

The helper encapsulates the solution of the ground-state wavefunction followed by the frequency-(in)dependent linear response equations,

$$\left[ \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{A}^* \end{pmatrix} - \omega_f \begin{pmatrix} \boldsymbol{\Sigma} & \boldsymbol{\Delta} \\ -\boldsymbol{\Delta}^* & -\boldsymbol{\Sigma}^* \end{pmatrix} \right] \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \begin{pmatrix} \mathbf{V} \\ -\mathbf{V}^* \end{pmatrix}, \quad (6.18)$$

either directly (via matrix inversion in the MO basis) or iteratively (via repeated matrix-vector products using Fock builds). For a HF/DFT reference with canonical orbitals, the above equations reduce to

$$\left[ \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{pmatrix} - \omega_f \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix} \right] \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \begin{pmatrix} \mathbf{V} \\ -\mathbf{V} \end{pmatrix}. \quad (\text{Neese 107})$$

In the static limit ( $\omega_f = 0$ ), the whole superoverlap matrix vanishes, and the CPHF equations can be reduced to those used in [tutorial 6a](#).

```

1 solver = helper_CPHF(mol)
2 solver.run()

1 Number of occupied orbitals: 5
2 Number of basis functions: 41
3
4 Tensor sizes:
5 ERI tensor          0.02 GB.
6 oNNN MO tensor     0.00 GB.
7 ovov Hessian tensor 0.00 GB.
8
9 Forming Hessian...
10 ...formed Hessian in 0.473 seconds.
11
12 Inverting Hessian...
13 ...inverted Hessian in 0.007 seconds.
```

Because the calculation of  $\beta$  requires  $U^a$ , we also obtain linear response properties from a quadratic response calculation. This holds for any order of response, where lower-order response functions are automatically obtained from higher-order response calculations.

```

1 print(np.around(solver.polar, 4))
1 [[ 7.2587 -0.      0.    ]
2  [-0.      8.7969  0.    ]
3  [ 0.      0.      7.854  ]]
```

```

1  # epsilon0
2  moenergies = solver.epsilon
3  C = np.asarray(solver.C)
4  Co = solver.Co
5  Cv = solver.Cv
6  nbf, norb = C.shape
7  nocc = Co.shape[1]
8  nvir = norb - nocc
9  nov = nocc * nvir
10 # the response vectors X_x, X_y, X_z; Y_x, Y_y, Y_z not needed separately for static
    ↪ response
11 x = np.asarray(solver.x)
12 ncomp = x.shape[0]
13 # reuse the AO-basis dipole integrals
14 integrals_ao = np.asarray([np.asarray(dipole_ao_component)
15                             for dipole_ao_component in solver.tmp_dipoles])
16 print("dimension of response vectors from linear response: {}".format(x.shape))
17 # for dynamic response, this will be (2 * nov)
18 assert x.shape[1] == nov

1  dimension of response vectors from linear response: (3, 180)

```

The foundation of the CPHF equations is that the right-hand side  $\mathbf{V}$  is a perturbation on the wavefunction causing single excitations from the occupied orbitals to virtual orbitals, the coefficients of which are in the response vectors  $\mathbf{X}$ ; the vectors  $\mathbf{Y}$  describe single deexcitations. Because the full (square)  $\mathbf{U}$  matrices are required, all MO-based quantities must be of shape  $[N_{orb}, N_{orb}]$  rather than  $[N_{occ}, N_{vir}]$ .

```

1  # form full MO-basis dipole integrals
2  integrals_mo = np.empty(shape=(ncomp, norb, norb))
3  for i in range(ncomp):
4      integrals_mo[i, ...] = (C.T).dot(integrals_ao[i, ...]).dot(C)

```

Similarly,  $\mathbf{X}$  and  $\mathbf{Y}$  form the off-diagonal blocks of the  $\mathbf{U}$  matrices. They are usually stored as in DALTON, where each vector is of length  $2N_{ov}$ , with  $\mathbf{X}$  on top of  $\mathbf{Y}$ .

```

1  # repack response vectors to [norb, norb]; 1/2 is due to X + Y
2  U = np.zeros_like(integrals_mo)
3  for i in range(ncomp):
4      U[i, :nocc, nocc:] = 0.5 * x[i, ...].reshape(nocc, nvir)
5      U[i, nocc:, :nocc] = -0.5 * x[i, ...].reshape(nocc, nvir).T

```

A minor implementation detail: because this was not a frequency-dependent calculation, only  $\mathbf{X} + \mathbf{Y}$  needs to be calculated; as they are identical, this leads to the prefactor of 1/2.

```

1  # form G matrices from perturbation and generalized Fock matrices; do
2  # one more Fock build for each response vector
3  jk = psi4.core.JK.build(solver.scf_wfn.basisset())
4  jk.initialize()

```

```

5  G = np.empty_like(U)
6  R = psi4.core.Matrix(nbf, nocc)
7  npR = np.asarray(R)
8  for i in range(ncomp):
9      V = integrals_mo[i, ...]
10
11     # eqn. (III-1b)
12     # Note: this simplified handling of the response vector
13     # transformation for the Fock build is insufficient for
14     # frequency-dependent response. 1/2 is due to X + Y
15     jk.C_clear()
16     L = Co
17     npR[...] = x[i, ...].reshape(nocc, nvir).dot(np.asarray(Cv).T).T
18     jk.C_left_add(L)
19     jk.C_right_add(R)
20     jk.compute()
21     J = 0.5 * np.asarray(jk.J())[0]
22     K = 0.5 * np.asarray(jk.K())[0]
23
24     # eqn. (21b)
25     F = (C.T).dot(4 * J - K.T - K).dot(C)
26     G[i, ...] = V + F
27
28     # form epsilon matrices, eqn. (34)
29     E = G.copy()
30     omega = 0
31     for i in range(ncomp):
32         eoU = (moenergies[..., np.newaxis] + omega) * U[i, ...]
33         Ue = U[i, ...] * moenergies[np.newaxis, ...]
34         E[i, ...] += (eoU - Ue)
35
36     # Assume some symmetry and calculate only part of the tensor.
37     # eqn. (VII-4)
38     hyperpolarizability = np.zeros(shape=(6, 3))
39     off1 = [0, 1, 2, 0, 0, 1]
40     off2 = [0, 1, 2, 1, 2, 2]
41     for r in range(6):
42         b = off1[r]
43         c = off2[r]
44         for a in range(3):
45             t11 = 2 * np.trace(U[a, ...].dot(G[b, ...]).dot(U[c, ...])[:nocc, :nocc])
46             t12 = 2 * np.trace(U[a, ...].dot(G[c, ...]).dot(U[b, ...])[:nocc, :nocc])
47             t13 = 2 * np.trace(U[c, ...].dot(G[a, ...]).dot(U[b, ...])[:nocc, :nocc])
48             tr1 = np.trace(U[c, ...].dot(U[b, ...]).dot(E[a, ...])[:nocc, :nocc])
49             tr2 = np.trace(U[b, ...].dot(U[c, ...]).dot(E[a, ...])[:nocc, :nocc])
50             tr3 = np.trace(U[c, ...].dot(U[a, ...]).dot(E[b, ...])[:nocc, :nocc])
51             tr4 = np.trace(U[a, ...].dot(U[c, ...]).dot(E[b, ...])[:nocc, :nocc])
52             tr5 = np.trace(U[b, ...].dot(U[a, ...]).dot(E[c, ...])[:nocc, :nocc])
53             tr6 = np.trace(U[a, ...].dot(U[b, ...]).dot(E[c, ...])[:nocc, :nocc])
54             t1 = t11 + t12 + t13
55             tr = tr1 + tr2 + tr3 + tr4 + tr5 + tr6
56             hyperpolarizability[r, a] = -2 * (t1 - tr)
57
58     ref_static = np.array([

```

```

2     [ 0.00000001,  0.00000000, -0.10826460],
3     [ 0.00000000,  0.00000000, -11.22412215],
4     [ 0.00000000,  0.00000000, -4.36450397],
5     [ 0.00000000,  0.00000000, -0.00000001],
6     [-0.10826460, -0.00000001,  0.00000000],
7     [-0.00000001, -11.22412215,  0.00000000]
8 ])
9 assert np.allclose(ref_static, hyperpolarizability, rtol=0.0, atol=1.0e-3)
10 print('First dipole hyperpolarizability (static):')
11 print(hyperpolarizability)

1 First dipole hyperpolarizability (static):
2 [[ -0.      -0.      -0.10826]
3  [ -0.      -0.     -11.22412]
4  [ -0.      -0.     -4.3645 ]
5  [ -0.      -0.       0.      ]
6  [ -0.10826  0.      -0.      ]
7  [  0.      -11.22412 -0.      ]]

```

### 6.9.2.7 References Primary equations and implementation:

- Karna, S. P.; Dupuis, M. Frequency dependent nonlinear optical properties of molecules: Formulation and implementation in the HONDO program. *J. Comput. Chem.* **1991**, *12*, 487–504, DOI: [10.1002/jcc.540120409](https://doi.org/10.1002/jcc.540120409)

$2n + 1$  rule:

- Yamaguchi, Y.; Goddard, J. D.; Osamura, Y.; Schaefer III, H. F., *A New Dimension to Quantum Chemistry: Analytic Derivative Methods in Ab Initio Molecular Electronic Structure Theory*; International Series of Monographs on Chemistry; Oxford University Press: 1994
- Epstein, S. T. General Remainder Theorem. *J. Chem. Phys.* **1968**, *48*, 4725–4726, DOI: [10.1063/1.1668053](https://doi.org/10.1063/1.1668053)
- Epstein, S. T. Constraints and the  $V^{2n+1}$  theorem. *Chem. Phys. Lett.* **1980**, *70*, 311–312, DOI: [10.1016/0009-2614\(80\)85340-1](https://doi.org/10.1016/0009-2614(80)85340-1)

Additional reading:

- Neese, F. Prediction of molecular properties and molecular spectroscopy with density functional theory: From fundamental theory to exchange-coupling. *Coord. Chem. Rev.* **2009**, *253*, 526–563, DOI: [10.1016/j.ccr.2008.05.014](https://doi.org/10.1016/j.ccr.2008.05.014)

## 7.0 FULL PUBLICATION LIST

1. Shao, Y.; Gan, Z.; Epifanovsky, E.; Gilbert, A. T. B.; Wormit, M.; Kussmann, J.; Lange, A. W.; Behn, A.; Deng, J.; Feng, X.; Ghosh, D.; Goldey, M.; Horn, P. R.; Jacobson, L. D.; Kaliman, I.; Khaliullin, R. Z.; Kuš, T.; Landau, A.; Liu, J.; Proynov, E. I.; Rhee, Y. M.; Richard, R. M.; Rohrdanz, M. A.; Steele, R. P.; Sundstrom, E. J.; Woodcock, H. L.; Zimmerman, P. M.; Zuev, D.; Albrecht, B.; Alguire, E.; Austin, B.; Beran, G. J. O.; Bernard, Y. A.; Berquist, E.; Brandhorst, K.; Bravaya, K. B.; Brown, S. T.; Casanova, D.; Chang, C.-M.; Chen, Y.; Chien, S. H.; Closser, K. D.; Crittenden, D. L.; Diedenhofen, M.; DiStasio Jr., R. A.; Do, H.; Dutoi, A. D.; Edgar, R. G.; Fatehi, S.; Fusti-Molnar, L.; Ghysels, A.; Golubeva-Zadorozhnaya, A.; Gomes, J.; Hanson-Heine, M. W. D.; Harbach, P. H. P.; Hauser, A. W.; Hohenstein, E. G.; Holden, Z. C.; Jagau, T.-C.; Ji, H.; Kaduk, B.; Khistyayev, K.; Kim, J.; Kim, J.; King, R. A.; Klunzinger, P.; Kosenkov, D.; Kowalczyk, T.; Krauter, C. M.; Lao, K. U.; Laurent, A. D.; Lawler, K. V.; Levchenko, S. V.; Lin, C. Y.; Liu, F.; Livshits, E.; Lochan, R. C.; Luenser, A.; Manohar, P.; Manzer, S. F.; Mao, S.-P.; Mardirossian, N.; Marenich, A. V.; Maurer, S. A.; Mayhall, N. J.; Neuscamman, E.; Oana, C. M.; Olivares-Amaya, R.; O'Neill, D. P.; Parkhill, J. A.; Perrine, T. M.; Peverati, R.; Prociuk, A.; Rehn, D. R.; Rosta, E.; Russ, N. J.; Sharada, S. M.; Sharma, S.; Small, D. W.; Sodt, A.; Stein, T.; Stück, D.; Su, Y.-C.; Thom, A. J. W.; Tsuchimochi, T.; Vanovschi, V.; Vogt, L.; Vydrov, O.; Wang, T.; Watson, M. A.; Wenzel, J.; White, A.; Williams, C. F.; Yang, J.; Yeganeh, S.; Yost, S. R.; You, Z.-Q.; Zhang, I. Y.; Zhang, X.; Zhao, Y.; Brooks, B. R.; Chan, G. K. L.; Chipman, D. M.; Cramer, C. J.; Goddard, W. A.; Gordon, M. S.; Hehre, W. J.; Klamt, A.; Schaefer, H. F.; Schmidt, M. W.; Sherrill, C. D.; Truhlar, D. G.; Warshel, A.; Xu, X.; Aspuru-Guzik, A.; Baer, R.; Bell, A. T.; Besley, N. A.; Chai, J.-D.; Dreuw, A.; Dunietz, B. D.; Furlani,

- T. R.; Gwaltney, S. R.; Hsu, C.-P.; Jung, Y.; Kong, J.; Lambrecht, D. S.; Liang, W.; Ochsenfeld, C.; Rassolov, V. A.; Slipchenko, L. V.; Subotnik, J. E.; Van Voorhis, T.; Herbert, J. M.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M. Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol. Phys.* **2015**, *113*, 184–215, DOI: [10.1080/00268976.2014.952696](https://doi.org/10.1080/00268976.2014.952696)
2. Brinzer, T.; Berquist, E. J.; Ren, Z.; Dutta, S.; Johnson, C. A.; Krisher, C. S.; Lambrecht, D. S.; Garrett-Roe, S. Ultrafast Vibrational Spectroscopy (2D-IR) of CO<sub>2</sub> in Ionic Liquids: Carbon Capture from Carbon Dioxide's Point of View. *J. Chem. Phys.* **2015**, *142*, 212425, DOI: [10.1063/1.4917467](https://doi.org/10.1063/1.4917467)
  3. Berquist, E. J.; Langner, K. M.; O'Boyle, N. M.; Tenderholt, A. L. Release of cclib version 1.3.1., 2015, DOI: [10.5281/zenodo.15108](https://doi.org/10.5281/zenodo.15108)
  4. Daly, C. A.; Berquist, E. J.; Brinzer, T.; Garrett-Roe, S.; Lambrecht, D. S.; Corcelli, S. A. Modeling Carbon Dioxide Vibrational Frequencies in Ionic Liquids: II. Spectroscopic Map. *J. Phys. Chem. B* **2016**, *120*, 12633–12642, DOI: [10.1021/acs.jpccb.6b09509](https://doi.org/10.1021/acs.jpccb.6b09509)
  5. Berquist, E. J.; Daly, C. A.; Brinzer, T.; Bullard, K. K.; Campbell, Z. M.; Corcelli, S. A.; Garrett-Roe, S.; Lambrecht, D. S. Modeling Carbon Dioxide Vibrational Frequencies in Ionic Liquids: I. Ab Initio Calculations. *J. Phys. Chem. B* **2017**, *121*, 208–220, DOI: [10.1021/acs.jpccb.6b09489](https://doi.org/10.1021/acs.jpccb.6b09489)
  6. Brinzer, T.; Berquist, E. J.; Dutta, S.; Johnson, C. A.; Krisher, C. S.; Lambrecht, D. S.; Garrett-Roe, S.; Ren, Z. Erratum: "Ultrafast vibrational spectroscopy (2D-IR) of CO<sub>2</sub> in ionic liquids: Carbon capture from carbon dioxide's point of view" [*J. Chem. Phys.* *142*, 212425 (2015)]. *J. Chem. Phys.* **2017**, *147*, 049901, DOI: [10.1063/1.4995447](https://doi.org/10.1063/1.4995447)
  7. Smith, D.; Burns, L. A.; Sirianni, D. A.; Nascimento, D. R.; Kumar, A.; James, A. M.; Schriber, J. B.; Zhang, T.; Zhang, B.; Abbott, A. S.; Berquist, E. J.; Lechner, M. H.; dos Anjos Cunha, L.; Heide, A. G.; Waldrop, J. M.; King, R. A.; Simmonett, A. C.; Turney, J. M.; Schaefer, H. F.; Evangelista, F. A.; DePrince III, A. E.; Crawford, T. D.; Patkowski, K.; Sherrill, C. D. Psi4NumPy: An Interactive Quantum Chemistry Programming Environment for Reference Implementations and Rapid Development. **2018**, DOI: [10.26434/chemrxiv.5746059.v1](https://doi.org/10.26434/chemrxiv.5746059.v1)



8. Berquist, E. J.; Lambrecht, D. S. A First Principles Approach for Partitioning Linear Response Properties into Additive and Cooperative Contributions. **2018**, DOI: [10.26434/chemrxiv.5773968.v1](https://doi.org/10.26434/chemrxiv.5773968.v1)
9. Jakubek, R. S.; Eskandari, A.; Berquist, E. J.; Lambrecht, D. S.; Asher, S. A. UV Resonance Raman Spectroscopy Monitors the  $\chi_2$  Angle of the Glutamine Sidechain. *in preparation* **2018**

## 8.0 BIBLIOGRAPHY

- [1] Gauss, J. In *Modern Methods and Algorithms of Quantum Chemistry*, ed. by Groten-dorst, J., John von Neumann Institute for Computing: 2000; Vol. 3, pp 541–592 (pages [2](#), [6](#), [15](#), [28](#), [29](#), [165](#)).
- [2] Neese, F. Prediction of molecular properties and molecular spectroscopy with density functional theory: From fundamental theory to exchange-coupling. *Coord. Chem. Rev.* **2009**, *253*, 526–563, DOI: [10.1016/j.ccr.2008.05.014](#) (pages [2](#), [22](#), [158](#), [231](#)).
- [3] Norman, P. A perspective on nonresonant and resonant electronic response theory for time-dependent molecular properties. *Phys. Chem. Chem. Phys.* **2011**, *13*, 20519–20535, DOI: [10.1039/C1CP21951K](#) (pages [2](#), [7](#), [28](#)).
- [4] Helgaker, T.; Coriani, S.; Jørgensen, P.; Kristensen, K.; Olsen, J.; Ruud, K. Recent Advances in Wave Function-Based Methods of Molecular-Property Calculations. *Chem. Rev.* **2012**, *112*, 543–631, DOI: [10.1021/cr2002239](#) (pages [2](#), [158](#)).
- [5] Szabo, A.; Ostlund, N. S., *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Dover Books on Chemistry; Dover Publications: 1989 (pages [2](#), [3](#), [18](#), [23](#)).
- [6] McWeeny, R., *Methods of Molecular Quantum Mechanics*; Theoretical chemistry; Academic Press: 1989 (pages [2](#), [28](#)).
- [7] Yamaguchi, Y.; Goddard, J. D.; Osamura, Y.; Schaefer III, H. F., *A New Dimension to Quantum Chemistry: Analytic Derivative Methods in Ab Initio Molecular Electronic Structure Theory*; International Series of Monographs on Chemistry; Oxford University Press: 1994 (pages [2](#), [18](#), [225](#), [231](#)).
- [8] Barron, L. D., *Molecular Light Scattering and Optical Activity*; Cambridge University Press: 2004 (page [2](#)).

- [9] Mulliken, R. S. Electronic Population Analysis on LCAO–MO Molecular Wave Functions. I. *J. Chem. Phys.* **1955**, *23*, 1833–1840, DOI: [10.1063/1.1740588](https://doi.org/10.1063/1.1740588) (page 3).
- [10] Löwdin, P.-O. On the Non-Orthogonality Problem Connected with the Use of Atomic Wave Functions in the Theory of Molecules and Crystals. *J. Chem. Phys.* **1950**, *18*, 365–375, DOI: [10.1063/1.1747632](https://doi.org/10.1063/1.1747632) (page 3).
- [11] Werling, K. A.; Hutchison, G. R.; Lambrecht, D. S. Piezoelectric Effects of Applied Electric Fields on Hydrogen-Bond Interactions: First-Principles Electronic Structure Investigation of Weak Electrostatic Interactions. *J. Phys. Chem. Lett.* **2013**, *4*, 1365–1370, DOI: [10.1021/jz400355v](https://doi.org/10.1021/jz400355v) (page 3).
- [12] Werling, K. A.; Griffin, M.; Hutchison, G. R.; Lambrecht, D. S. Piezoelectric Hydrogen Bonding: Computational Screening for a Design Rationale. *J. Phys. Chem. A* **2014**, *118*, 7404–7410, DOI: [10.1021/jp412740j](https://doi.org/10.1021/jp412740j) (page 3).
- [13] Marvin, C. W.; Grimm, H. M.; Miller, N. C.; Horne, W. S.; Hutchison, G. R. Interplay among Sequence, Folding Propensity, and Bio-Piezoelectric Response in Short Peptides and Peptoids. *J. Phys. Chem. B* **2017**, *121*, 10269–10275, DOI: [10.1021/acs.jpcc.7b10085](https://doi.org/10.1021/acs.jpcc.7b10085) (page 3).
- [14] Ren, X.; Rinke, P.; Joas, C.; Scheffler, M. Random-phase approximation and its applications in computational chemistry and materials science. *J. Mater. Sci.* **2012**, *47*, 7447–7471, DOI: [10.1007/s10853-012-6570-4](https://doi.org/10.1007/s10853-012-6570-4) (page 4).
- [15] Dreuw, A.; Head-Gordon, M. Single-Reference ab Initio Methods for the Calculation of Excited States of Large Molecules. *Chem. Rev.* **2005**, *105*, 4009–4037, DOI: [10.1021/cr0505627](https://doi.org/10.1021/cr0505627) (page 4).
- [16] Oddershede, J.; Jørgensen, P.; Yeager, D. L. Polarization propagator methods in atomic and molecular calculations. *Computer Physics Reports* **1984**, *2*, 33–92, DOI: [10.1016/0167-7977\(84\)90003-0](https://doi.org/10.1016/0167-7977(84)90003-0) (page 4).
- [17] Jensen, F., *Introduction to Computational Chemistry*; Wiley: 2013 (page 6).
- [18] Warnke, I.; Furche, F. Circular dichroism: electronic. *WIREs Comput Mol Sci* **2012**, *2*, 150–166, DOI: [10.1002/wcms.55](https://doi.org/10.1002/wcms.55) (page 12).
- [19] Bogaard, M. P.; Haines, R. Raman intensities and cartesian polarizability derivatives. *Mol. Phys.* **1980**, *41*, 1281–1289, DOI: [10.1080/00268978000103541](https://doi.org/10.1080/00268978000103541) (page 14).

- [20] Pople, J. A.; Raghavachari, K.; Schlegel, H. B.; Binkley, J. S. Derivative studies in Hartree-Fock and Møller-Plesset theories. *Int. J. Quantum Chem.* **1979**, *16*, 225–241, DOI: [10.1002/qua.560160825](https://doi.org/10.1002/qua.560160825) (page 18).
- [21] Epstein, S. T. General Remainder Theorem. *J. Chem. Phys.* **1968**, *48*, 4725–4726, DOI: [10.1063/1.1668053](https://doi.org/10.1063/1.1668053) (pages 19, 225, 231).
- [22] Epstein, S. T. Constraints and the  $V^{2n+1}$  theorem. *Chem. Phys. Lett.* **1980**, *70*, 311–312, DOI: [10.1016/0009-2614\(80\)85340-1](https://doi.org/10.1016/0009-2614(80)85340-1) (pages 19, 225, 231).
- [23] Darling, C. L.; Schlegel, H. B. Dipole Moments, Polarizabilities, and Infrared Intensities Calculated with Electric Field Dependent Functions. *J. Phys. Chem.* **1994**, *98*, 5855–5861, DOI: [10.1021/j100074a008](https://doi.org/10.1021/j100074a008) (page 22).
- [24] Nobes, R. H.; Pople, J. A.; Radom, L.; Handy, N. C.; Knowles, P. J. Slow convergence of the Møller-Plesset perturbation series: the dissociation energy of hydrogen cyanide and the electron affinity of the cyano radical. *Chem. Phys. Lett.* **1987**, *138*, 481–485, DOI: [10.1016/0009-2614\(87\)80545-6](https://doi.org/10.1016/0009-2614(87)80545-6) (page 23).
- [25] Christiansen, O.; Jørgensen, P.; Hättig, C. Response functions from Fourier component variational perturbation theory applied to a time-averaged quasienergy. *Int. J. Quantum Chem.* **1998**, *68*, 1–52, DOI: [10.1002/\(SICI\)1097-461X\(1998\)68:1<1::AID-QUA1>3.0.CO;2-Z](https://doi.org/10.1002/(SICI)1097-461X(1998)68:1<1::AID-QUA1>3.0.CO;2-Z) (pages 28, 29).
- [26] Toulouse, J. Introduction to the calculation of molecular properties by response theory., 2015 (pages 28, 29).
- [27] Brinzer, T.; Berquist, E. J.; Ren, Z.; Dutta, S.; Johnson, C. A.; Krisher, C. S.; Lambrecht, D. S.; Garrett-Roe, S. Ultrafast Vibrational Spectroscopy (2D-IR) of CO<sub>2</sub> in Ionic Liquids: Carbon Capture from Carbon Dioxide’s Point of View. *J. Chem. Phys.* **2015**, *142*, 212425, DOI: [10.1063/1.4917467](https://doi.org/10.1063/1.4917467) (pages 33, 83, 84, 88, 92, 96, 100–102, 107, 111, 127, 128, 134, 135, 145–147, 160, 165, 233).
- [28] Brinzer, T.; Berquist, E. J.; Dutta, S.; Johnson, C. A.; Krisher, C. S.; Lambrecht, D. S.; Garrett-Roe, S.; Ren, Z. Erratum: "Ultrafast vibrational spectroscopy (2D-IR) of CO<sub>2</sub> in ionic liquids: Carbon capture from carbon dioxide’s point of view" [J. Chem. Phys. 142, 212425 (2015)]. *J. Chem. Phys.* **2017**, *147*, 049901, DOI: [10.1063/1.4995447](https://doi.org/10.1063/1.4995447) (pages 33, 233).

- [29] Bhowan, A. S.; Freeman, B. C. Analysis and Status of Post-Combustion Carbon Dioxide Capture Technologies. *Environ. Sci. Technol.* **2011**, *45*, 8624–8632, DOI: [10.1021/es104291d](https://doi.org/10.1021/es104291d) (page 34).
- [30] Millward, A. R.; Yaghi, O. M. Metal-Organic Frameworks with Exceptionally High Capacity for Storage of Carbon Dioxide at Room Temperature. *J. Am. Chem. Soc.* **2005**, *127*, 17998–17999, DOI: [10.1021/ja0570032](https://doi.org/10.1021/ja0570032) (page 34).
- [31] Sumida, K.; Rogow, D. L.; Mason, J. A.; McDonald, T. M.; Bloch, E. D.; Herm, Z. R.; Bae, T.-h.; Long, J. R. Carbon Dioxide Capture in Metal–Organic Frameworks. *Chem. Rev.* **2012**, *112*, 724–781 (pages 34, 84).
- [32] Du, N.; Park, H. B.; Robertson, G. P.; Dal-Cin, M. M.; Visser, T.; Scoles, L.; Guiver, M. D. Polymer nanosieve membranes for CO<sub>2</sub>-capture applications. *Nat. Mater.* **2011**, *10*, 372–375, DOI: [10.1038/nmat2989](https://doi.org/10.1038/nmat2989) (pages 34, 84).
- [33] Dawson, R.; Stöckel, E.; Holst, J. R.; Adams, D. J.; Cooper, A. I. Microporous organic polymers for carbon dioxide capture. *Energy Environ. Sci.* **2011**, *4*, 4239, DOI: [10.1039/c1ee01971f](https://doi.org/10.1039/c1ee01971f) (pages 34, 84).
- [34] Bates, E. D.; Mayton, R. D.; Ntai, I.; Davis, J. H. CO<sub>2</sub> Capture by a Task-Specific Ionic Liquid. *J. Am. Chem. Soc.* **2002**, *124*, 926–927, DOI: [10.1021/ja017593d](https://doi.org/10.1021/ja017593d) (pages 34, 35).
- [35] Karadas, F.; Atilhan, M.; Aparicio, S. Review on the Use of Ionic Liquids (ILs) as Alternative Fluids for CO<sub>2</sub> Capture and Natural Gas Sweetening. *Energy Fuels* **2010**, *24*, 5817–5828, DOI: [10.1021/ef1011337](https://doi.org/10.1021/ef1011337) (pages 34, 133).
- [36] Bara, J. E.; Camper, D. E.; Gin, D. L.; Noble, R. D. Room-Temperature Ionic Liquids and Composite Materials: Platform Technologies for CO<sub>2</sub> Capture. *Acc. Chem. Res.* **2010**, *43*, 152–159, DOI: [10.1021/ar9001747](https://doi.org/10.1021/ar9001747) (pages 34, 84, 133).
- [37] Holbrey, J. D.; Seddon, K. R. Ionic Liquids. *Clean Technol. Environ. Policy* **1999**, *1*, 223–236, DOI: [10.1007/s100980050036](https://doi.org/10.1007/s100980050036) (page 35).
- [38] Bara, J. E.; Carlisle, T. K.; Gabriel, C. J.; Camper, D.; Finotello, A.; Gin, D. L.; Noble, R. D. Guide to CO<sub>2</sub> Separations in Imidazolium-Based Room-Temperature Ionic Liquids. *Ind. Eng. Chem. Res.* **2009**, *48*, 2739–2751, DOI: [10.1021/ie8016237](https://doi.org/10.1021/ie8016237) (page 35).

- [39] Cadena, C.; Anthony, J. L.; Shah, J. K.; Morrow, T. I.; Brennecke, J. F.; Maginn, E. J. Why Is CO<sub>2</sub> So Soluble in Imidazolium-Based Ionic Liquids? *J. Am. Chem. Soc.* **2004**, *126*, 5300–5308, DOI: [10.1021/ja039615x](https://doi.org/10.1021/ja039615x) (pages [35](#), [44](#), [133](#), [145](#)).
- [40] Gurkan, B. E.; de la Fuente, J. C.; Mindrup, E. M.; Ficke, L. E.; Goodrich, B. F.; Price, E. A.; Schneider, W. F.; Brennecke, J. F. Equimolar CO<sub>2</sub> Absorption by Anion-Functionalized Ionic Liquids. *J. Am. Chem. Soc.* **2010**, *132*, 2116–2117, DOI: [10.1021/ja909305t](https://doi.org/10.1021/ja909305t) (pages [35](#), [133](#)).
- [41] Seo, S.; Quiroz-Guzman, M.; Desilva, M. A.; Lee, T. B.; Huang, Y.; Goodrich, B. F.; Schneider, W. F.; Brennecke, J. F. Chemically Tunable Ionic Liquids with Aprotic Heterocyclic Anion (AHA) for CO<sub>2</sub> Capture. *J. Phys. Chem. B* **2014**, *118*, 5740–5751, DOI: [10.1021/jp502279w](https://doi.org/10.1021/jp502279w) (pages [35](#), [84](#), [133](#)).
- [42] Ramdin, M.; de Loos, T. W.; Vlugt, T. J. H. State-of-the-Art of CO<sub>2</sub> Capture with Ionic Liquids. *Ind. Eng. Chem. Res.* **2012**, *51*, 8149–8177, DOI: [10.1021/ie3003705](https://doi.org/10.1021/ie3003705) (pages [35](#), [145](#)).
- [43] Wang, C.; Luo, X.; Luo, H.; Jiang, D.-e.; Li, H.; Dai, S. Tuning the basicity of ionic liquids for equimolar CO<sub>2</sub> capture. *Angew. Chem. Int. Ed. Engl.* **2011**, *50*, 4918–4922, DOI: [10.1002/anie.201008151](https://doi.org/10.1002/anie.201008151) (page [35](#)).
- [44] Wang, C.; Luo, H.; Jiang, D.-e.; Li, H.; Dai, S. Carbon Dioxide Capture by Superbase-Derived Protic Ionic Liquids. *Angew. Chemie* **2010**, *122*, 6114–6117, DOI: [10.1002/ange.201002641](https://doi.org/10.1002/ange.201002641) (page [35](#)).
- [45] Zhang, Y.; Zhang, S.; Lu, X.; Zhou, Q.; Fan, W.; Zhang, X. Dual amino-functionalised phosphonium ionic liquids for CO<sub>2</sub> capture. *Chemistry* **2009**, *15*, 3003–3011, DOI: [10.1002/chem.200801184](https://doi.org/10.1002/chem.200801184) (page [35](#)).
- [46] Padiuszyński, K.; Domańska, U. Viscosity of Ionic Liquids: An Extensive Database and a New Group Contribution Model Based on Feed-forward Artificial Neural Network. *J. Chem. Inf. Model.* **2014**, *54*, 1311–1324, DOI: [10.1021/ci500206u](https://doi.org/10.1021/ci500206u) (page [35](#)).
- [47] Yan, F.; Lartey, M.; Jariwala, K.; Bowser, S.; Damodaran, K.; Albenze, E.; Luebke, D. R.; Nulwala, H. B.; Smit, B.; Haranczyk, M. Toward a Materials Genome Approach for Ionic Liquids : Synthesis Guided by Ab Initio Property Maps. *J. Phys. Chem. B* **2014**, *118*, 13609–13620, DOI: [10.1021/jp506972w](https://doi.org/10.1021/jp506972w) (page [35](#)).

- [48] Chatel, G.; Pereira, J. F. B.; Debbeti, V.; Wang, H.; Rogers, R. D. Mixing Ionic Liquids - “Simple Mixtures” or “Double Salts”? *Green Chem.* **2014**, *16*, 2051–2083, DOI: [10.1039/c3gc41389f](https://doi.org/10.1039/c3gc41389f) (page 35).
- [49] Maximo, G. J.; Santos, R. J. B. N.; Brandão, P.; Esperança, J. M. S. S.; Costa, M. C.; Meirelles, A. J. A.; Freire, M. G.; Coutinho, J. A. P. Generating Ionic Liquids from Ionic Solids: An Investigation of the Melting Behavior of Binary Mixtures of Ionic Liquids. *Cryst. Growth Des.* **2014**, *14*, 4270–4277, DOI: [10.1021/cg500655s](https://doi.org/10.1021/cg500655s) (page 35).
- [50] Romanos, G. E.; Zubeir, L. F.; Likodimos, V.; Falaras, P.; Kroon, M. C.; Iliev, B.; Adamova, G.; Schubert, T. J. S. Enhanced CO<sub>2</sub> Capture in Binary Mixtures of 1-Alkyl-3-methylimidazolium Tricyanomethanide Ionic Liquids with Water. *J. Phys. Chem. B* **2013**, *117*, 12234–12251, DOI: [10.1021/jp407364e](https://doi.org/10.1021/jp407364e) (pages 35, 36).
- [51] Camper, D.; Bara, J. E.; Gin, D. L.; Noble, R. D. Room-Temperature Ionic Liquid–Amine Solutions: Tunable Solvents for Efficient and Reversible Capture of CO<sub>2</sub>. *Ind. Eng. Chem. Res.* **2008**, *47*, 8496–8498, DOI: [10.1021/ie801002m](https://doi.org/10.1021/ie801002m) (page 35).
- [52] Yu, G.; Zhang, S.; Yao, X.; Zhang, J.; Dong, K.; Dai, W.; Mori, R. Design of task-specific ionic liquids for capturing CO<sub>2</sub>: A molecular orbital study. *Ind. Eng. Chem. Res.* **2006**, *45*, 2875–2880, DOI: [10.1021/ie050975y](https://doi.org/10.1021/ie050975y) (page 35).
- [53] Bhargava, B. L.; Balasubramanian, S. Probing anion-carbon dioxide interactions in room temperature ionic liquids: Gas phase cluster calculations. *Chem. Phys. Lett.* **2007**, *444*, 242–246, DOI: [10.1016/j.cplett.2007.07.051](https://doi.org/10.1016/j.cplett.2007.07.051) (pages 35, 44).
- [54] Gonzalez-Miquel, M.; Talreja, M.; Ethier, A. L.; Flack, K.; Switzer, J. R.; Biddinger, E. J.; Pollet, P.; Palomar, J.; Rodriguez, F.; Eckert, C. A.; Liotta, C. L. COSMO-RS Studies: Structure-property relationships for CO<sub>2</sub> capture by reversible ionic liquids. *Ind. Eng. Chem. Res.* **2012**, *51*, 16066–16073, DOI: [10.1021/ie302449c](https://doi.org/10.1021/ie302449c) (page 35).
- [55] Steckel, J. A. Ab initio calculations of the interaction between CO<sub>2</sub> and the acetate ion. *J. Phys. Chem. A* **2012**, *116*, 11643–11650, DOI: [10.1021/jp306446d](https://doi.org/10.1021/jp306446d) (page 35).
- [56] Huang, X.; Margulis, C. J.; Li, Y.; Berne, B. J. Why is the partial molar volume of CO<sub>2</sub> so small when dissolved in a room temperature ionic liquid? Structure and

- dynamics of CO<sub>2</sub> dissolved in [bmim<sup>+</sup>][PF<sub>6</sub><sup>-</sup>]. *J. Am. Chem. Soc.* **2005**, *127*, 17842–17851, DOI: [10.1021/ja055315z](https://doi.org/10.1021/ja055315z) (pages 35, 72).
- [57] Kerlé, D.; Ludwig, R.; Geiger, A.; Paschek, D. Temperature dependence of the solubility of carbon dioxide in imidazolium-based ionic liquids. *J. Phys. Chem. B* **2009**, *113*, 12727–35, DOI: [10.1021/jp9055285](https://doi.org/10.1021/jp9055285) (page 35).
- [58] Palomar, J.; Gonzalez-Miquel, M.; Polo, A.; Rodriguez, F. Understanding the Physical Absorption of CO<sub>2</sub> in Ionic Liquids Using the COSMO-RS Method. *Ind. Eng. Chem. Res.* **2011**, *50*, 3452–3463, DOI: [10.1021/ie101572m](https://doi.org/10.1021/ie101572m) (page 35).
- [59] Wendler, K.; Brehm, M.; Malberg, F.; Kirchner, B.; Delle Site, L. Short Time Dynamics of Ionic Liquids in AIMD-Based Power Spectra. *J. Chem. Theory Comput.* **2012**, *8*, 1570–1579, DOI: [10.1021/ct300152t](https://doi.org/10.1021/ct300152t) (page 35).
- [60] Hollóczki, O.; Kelemen, Z.; Könczöl, L.; Szieberth, D.; Nyulászi, L.; Stark, A.; Kirchner, B. Significant cation effects in carbon dioxide-ionic liquid systems. *ChemPhysChem* **2013**, *14*, 315–20, DOI: [10.1002/cphc.201200970](https://doi.org/10.1002/cphc.201200970) (page 35).
- [61] Shim, Y.; Kim, H. J. MD study of Solvation in the Mixture of a Room-Temperature Ionic Liquid and CO<sub>2</sub>. *J. Phys. Chem. B* **2010**, *114*, 10160–10170, DOI: [10.1021/jp105021b](https://doi.org/10.1021/jp105021b) (page 35).
- [62] Shi, W.; Sorescu, D. C.; Luebke, D. R.; Keller, M. J.; Wickramanayake, S. Molecular Simulations and Experimental Studies of Solubility and Diffusivity for Pure and Mixed Gases of H<sub>2</sub>, CO<sub>2</sub>, and Ar Absorbed in the Ionic Liquid 1-*n*-Hexyl-3-methylimidazolium Bis(Trifluoromethylsulfonyl)amide ([hmim][Tf<sub>2</sub>N]). *J. Phys. Chem. B* **2010**, *114*, 6531–6541, DOI: [10.1021/jp101897b](https://doi.org/10.1021/jp101897b) (page 35).
- [63] Shi, W.; Thompson, R. L.; Albenze, E.; Steckel, J. A.; Nulwala, H. B.; Luebke, D. R. Contribution of the Acetate Anion to CO<sub>2</sub> Solubility in Ionic Liquids: Theoretical Method Development and Experimental Study. *J. Phys. Chem. B* **2014**, *118*, 7383–7394, DOI: [10.1021/jp502425a](https://doi.org/10.1021/jp502425a) (page 35).
- [64] Shah, J. K.; Brennecke, J. F.; Maginn, E. J. Thermodynamic properties of the ionic liquid 1-*n*-butyl-3-methylimidazolium hexafluorophosphate from Monte Carlo simulations. *Green Chem.* **2002**, *4*, 112–118, DOI: [10.1039/b110725a](https://doi.org/10.1039/b110725a) (page 35).



- [65] Kerlé, D.; Ludwig, R.; Paschek, D. The Influence of Water on the Solubility of Carbon Dioxide in Imidazolium Based Ionic Liquids. *Zeitschrift für Phys. Chemie* **2013**, *227*, 167–176, DOI: [10.1524/zpch.2013.0344](https://doi.org/10.1524/zpch.2013.0344) (page 35).
- [66] Ghobadi, A. F.; Taghikhani, V.; Elliott, J. R. Investigation on the Solubility of SO<sub>2</sub> and CO<sub>2</sub> in Imidazolium-Based Ionic Liquids Using NPT Monte Carlo Simulation. *J. Phys. Chem. B* **2011**, *115*, 13599–13607, DOI: [10.1021/jp2051239](https://doi.org/10.1021/jp2051239) (page 35).
- [67] Khalil, M.; Demirdöven, N.; Tokmakoff, A. Coherent 2D IR Spectroscopy: Molecular Structure and Dynamics in Solution. *J. Phys. Chem. A* **2003**, *107*, 5258–5279, DOI: [10.1021/jp0219247](https://doi.org/10.1021/jp0219247) (pages 35, 133).
- [68] Cho, M. Coherent Two-Dimensional Optical Spectroscopy. *Chem. Rev.* **2008**, *108*, 1331–1418, DOI: [10.1021/cr078377b](https://doi.org/10.1021/cr078377b) (page 35).
- [69] Hamm, P.; Zanni, M. T., *Concepts and Methods of 2D Infrared Spectroscopy*; Cambridge University Press: 2015; Vol. 1, DOI: [10.1017/CB09781107415324.004](https://doi.org/10.1017/CB09781107415324.004) (pages 35, 58, 133).
- [70] Zheng, J. R.; Kwak, K.; Asbury, J.; Chen, X.; Piletic, I. R.; Fayer, M. D. Ultrafast dynamics of solute-solvent complexation observed at thermal equilibrium in real time. *Science* **2005**, *309*, 1338–1343, DOI: [10.1126/science.1116213](https://doi.org/10.1126/science.1116213) (pages 35, 36).
- [71] Wong, D. B.; Giammanco, C. H.; Fenn, E. E.; Fayer, M. D. Dynamics of Isolated Water Molecules in a Sea of Ions in a Room Temperature Ionic Liquid. *J. Phys. Chem. B* **2013**, *117*, 623–635, DOI: [10.1021/jp310086s](https://doi.org/10.1021/jp310086s) (pages 36, 37, 72–75).
- [72] Woutersen, S.; Mu, Y.; Stock, G.; Hamm, P. Subpicosecond Conformational Dynamics of Small Peptides Probed by Two-Dimensional Vibrational Spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 11254–11258, DOI: [10.1073/pnas.201169498](https://doi.org/10.1073/pnas.201169498) (page 36).
- [73] Cabaço, M. I.; Besnard, M.; Danten, Y.; Coutinho, J. A. P. Solubility of CO<sub>2</sub> in 1-butyl-3-methyl-imidazolium-trifluoro acetate ionic liquid studied by Raman spectroscopy and DFT investigations. *J. Phys. Chem. B* **2011**, *115*, 3538–3550, DOI: [10.1021/jp111453a](https://doi.org/10.1021/jp111453a) (pages 36, 39).
- [74] Corvo, M. C.; Sardinha, J.; Menezes, S. C.; Einloft, S.; Seferin, M.; Dupont, J.; Casimiro, T.; Cabrita, E. J. Solvation of CO<sub>2</sub> in [C<sub>4</sub>mim][BF<sub>4</sub>] and [C<sub>4</sub>mim][PF<sub>6</sub>]

- Ionic Liquids Revealed by High-Pressure NMR Spectroscopy. *Angew. Chem. Int. Ed. Engl.* **2013**, *52*, 13024–13027, DOI: [10.1002/anie.201305630](https://doi.org/10.1002/anie.201305630) (pages 36, 72).
- [75] Carvalho, P. J.; Alvarez, V. H.; Schröder, B.; Gil, A. M.; Marrucho, I. M.; Aznar, M.; Santos, L. M. N. B. F.; Coutinho, J. A. P. Specific solvation interactions of CO<sub>2</sub> on acetate and trifluoroacetate imidazolium based ionic liquids at high pressures. *J. Phys. Chem. B* **2009**, *113*, 6803–6812, DOI: [10.1021/jp901275b](https://doi.org/10.1021/jp901275b) (page 36).
- [76] Dahl, K.; Sando, G. M.; Fox, D. M.; Sutto, T. E.; Owrutsky, J. C. Vibrational Spectroscopy and Dynamics of Small Anions in Ionic Liquid Solutions. *J. Chem. Phys.* **2005**, *123*, 084504, DOI: [10.1063/1.2000229](https://doi.org/10.1063/1.2000229) (page 37).
- [77] Ren, Z.; Ivanova, A. S.; Couchot-Vore, D.; Garrett-Roe, S. Ultrafast Structure and Dynamics in Ionic Liquids: 2D-IR Spectroscopy Probes the Molecular Origin of Viscosity. *J. Phys. Chem. Lett.* **2014**, *5*, 1541–1546, DOI: [10.1021/jz500372f](https://doi.org/10.1021/jz500372f) (pages 37, 72–74, 133).
- [78] Maroncelli, M.; Zhang, X.-X.; Liang, M.; Roy, D.; Ernsting, N. P. Measurements of the Complete Solvation Response of Coumarin 153 in Ionic Liquids and the Accuracy of Simple Dielectric Continuum Predictions. *Faraday Discuss.* **2012**, *154*, 409, DOI: [10.1039/c1fd00058f](https://doi.org/10.1039/c1fd00058f) (page 37).
- [79] Zhang, X.-X.; Liang, M.; Ernsting, N. P.; Maroncelli, M. Complete Solvation Response of Coumarin 153 in Ionic Liquids. *J. Phys. Chem. B* **2013**, *117*, 4291–4304, DOI: [10.1021/jp305430a](https://doi.org/10.1021/jp305430a) (page 37).
- [80] Turton, D. A.; Hunger, J.; Stoppa, A.; Hefter, G.; Thoman, A.; Walkther, M.; Buchner, R.; Wynne, K. Dynamics of Imidazolium Ionic Liquids from a Combined Dielectric Relaxation and Optical Kerr Effect Study: Evidence for Mesoscopic Aggregation. *J. Am. Chem. Soc.* **2009**, *131*, 11140–11146, DOI: [10.1021/ja903315v](https://doi.org/10.1021/ja903315v) (page 37).
- [81] Xiao, D.; Hines, Jr., L. G.; Bartsch, R. A.; Quitevis, E. L. Intermolecular Vibrational Motions of Solute Molecules Confined in Nonpolar Domains of Ionic Liquids. *J. Phys. Chem. B* **2009**, *113*, 4544–4548, DOI: [10.1021/jp811293n](https://doi.org/10.1021/jp811293n) (page 37).
- [82] Bardak, F.; Xiao, D.; Hines, Jr., L. G.; Son, P.; Bartsch, R. A.; Quitevis, E. L.; Yang, P.; Voth, G. A. Nanostructural Organization in Acetonitrile and Ionic Liq-

- uid Mixtures: Molecular Dynamics Simulations and Optical Kerr Effect Spectroscopy. *ChemPhysChem* **2012**, *13*, 1687–1700, DOI: [10.1002/cphc.201200026](https://doi.org/10.1002/cphc.201200026) (page 37).
- [83] Turton, D. A.; Sonnleitner, T.; Ortner, A.; Walther, M.; Hefter, G.; Seddon, K. R.; Stana, S.; Plechkova, N. V.; Buchner, R.; Wynne, K. Structure and Dynamics in Protic Ionic Liquids: A Combined Optical Kerr-Effect and Dielectric Relaxation Spectroscopy Study. *Faraday Discuss.* **2012**, *154*, 145, DOI: [10.1039/c1fd00054c](https://doi.org/10.1039/c1fd00054c) (page 37).
- [84] Castner, E. W. J.; Wishart, J. F.; Shirota, H. Intermolecular Dynamics, Interactions, and Solvation in Ionic Liquids. *Acc. Chem. Res.* **2007**, *40*, 1217–1227, DOI: [10.1021/ar700169g](https://doi.org/10.1021/ar700169g) (page 37).
- [85] Hamm, P.; Lim, M.; Hochstrasser, R. M. In *Ultrafast Phenom. XI Proc. 11th Int. Conf.* 1998, pp 514–516 (pages 37, 38).
- [86] Garrett-Roe, S.; Hamm, P. Purely Absorptive Three-Dimensional Infrared Spectroscopy. *J. Chem. Phys.* **2009**, *130*, 164510, DOI: [10.1063/1.3122982](https://doi.org/10.1063/1.3122982) (pages 37, 61–64).
- [87] Meredith, J. C.; Johnston, K. P.; Seminario, J. M.; Kazarian, S. G.; Eckert, C. A. Quantitative Equilibrium Constants between CO<sub>2</sub> and Lewis Bases from FTIR Spectroscopy. *J. Phys. Chem.* **1996**, *100*, 10837–10848, DOI: [10.1021/jp953161b](https://doi.org/10.1021/jp953161b) (pages 39, 59).
- [88] Kazarian, S. G.; Briscoe, B. J.; Welton, T. Combining ionic liquids and supercritical fluids: *in situ* ATR-IR study of CO<sub>2</sub> dissolved in two ionic liquids at high pressures. *Chem. Commun.* **2000**, 2047–2048, DOI: [10.1039/b005514j](https://doi.org/10.1039/b005514j) (pages 39, 60).
- [89] Giernoth, R.; Bankmann, D. Transition-metal-free synthesis of perdeuterated imidazolium ionic liquids by alkylation and H/D exchange. *European J. Org. Chem.* **2008**, 2881–2886, DOI: [10.1002/ejoc.200700784](https://doi.org/10.1002/ejoc.200700784) (page 40).
- [90] Hamm, P.; Kaindl, R. A.; Stenger, J. Noise suppression in femtosecond mid-infrared light sources. *Opt. Lett.* **2000**, *25*, 1798–1800 (page 40).
- [91] Helbing, J.; Hamm, P. Compact implementation of Fourier transform two-dimensional IR spectroscopy without phase ambiguity. *J. Opt. Soc. Am. B* **2010**, *28*, 171–178, DOI: [10.1364/JOSAB.28.000171](https://doi.org/10.1364/JOSAB.28.000171) (page 40).

- [92] Press, W. H.; Teukolsky, S. A.; Vetterlin, W. T.; Flannery, B. P., *Numerical Recipes in C*, 2nd; Cambridge University Press: New York, 1992 (page 41).
- [93] Shao, Y.; Gan, Z.; Epifanovsky, E.; Gilbert, A. T. B.; Wormit, M.; Kussmann, J.; Lange, A. W.; Behn, A.; Deng, J.; Feng, X.; Ghosh, D.; Goldey, M.; Horn, P. R.; Jacobson, L. D.; Kaliman, I.; Khaliullin, R. Z.; Kuś, T.; Landau, A.; Liu, J.; Proynov, E. I.; Rhee, Y. M.; Richard, R. M.; Rohrdanz, M. A.; Steele, R. P.; Sundstrom, E. J.; Woodcock, H. L.; Zimmerman, P. M.; Zuev, D.; Albrecht, B.; Alguire, E.; Austin, B.; Beran, G. J. O.; Bernard, Y. A.; Berquist, E.; Brandhorst, K.; Bravaya, K. B.; Brown, S. T.; Casanova, D.; Chang, C.-M.; Chen, Y.; Chien, S. H.; Closser, K. D.; Crittenden, D. L.; Diedenhofen, M.; DiStasio Jr., R. A.; Do, H.; Dutoi, A. D.; Edgar, R. G.; Fatehi, S.; Fusti-Molnar, L.; Ghysels, A.; Golubeva-Zadorozhnaya, A.; Gomes, J.; Hanson-Heine, M. W. D.; Harbach, P. H. P.; Hauser, A. W.; Hohenstein, E. G.; Holden, Z. C.; Jagau, T.-C.; Ji, H.; Kaduk, B.; Khistyayev, K.; Kim, J.; Kim, J.; King, R. A.; Klunzinger, P.; Kosenkov, D.; Kowalczyk, T.; Krauter, C. M.; Lao, K. U.; Laurent, A. D.; Lawler, K. V.; Levchenko, S. V.; Lin, C. Y.; Liu, F.; Livshits, E.; Lochan, R. C.; Luenser, A.; Manohar, P.; Manzer, S. F.; Mao, S.-P.; Mardirossian, N.; Marenich, A. V.; Maurer, S. A.; Mayhall, N. J.; Neuscammann, E.; Oana, C. M.; Olivares-Amaya, R.; O'Neill, D. P.; Parkhill, J. A.; Perrine, T. M.; Peverati, R.; Prociuk, A.; Rehn, D. R.; Rosta, E.; Russ, N. J.; Sharada, S. M.; Sharma, S.; Small, D. W.; Sodt, A.; Stein, T.; Stück, D.; Su, Y.-C.; Thom, A. J. W.; Tsuchimochi, T.; Vanovschi, V.; Vogt, L.; Vydrov, O.; Wang, T.; Watson, M. A.; Wenzel, J.; White, A.; Williams, C. F.; Yang, J.; Yeganeh, S.; Yost, S. R.; You, Z.-Q.; Zhang, I. Y.; Zhang, X.; Zhao, Y.; Brooks, B. R.; Chan, G. K. L.; Chipman, D. M.; Cramer, C. J.; Goddard, W. A.; Gordon, M. S.; Hehre, W. J.; Klamt, A.; Schaefer, H. F.; Schmidt, M. W.; Sherrill, C. D.; Truhlar, D. G.; Warshel, A.; Xu, X.; Aspuru-Guzik, A.; Baer, R.; Bell, A. T.; Besley, N. A.; Chai, J.-D.; Dreuw, A.; Dunietz, B. D.; Furlani, T. R.; Gwaltney, S. R.; Hsu, C.-P.; Jung, Y.; Kong, J.; Lambrecht, D. S.; Liang, W.; Ochsenfeld, C.; Rassolov, V. A.; Slipchenko, L. V.; Subotnik, J. E.; Van Voorhis, T.; Herbert, J. M.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M. *Advances in molecular quantum chemistry*

- contained in the Q-Chem 4 program package. *Mol. Phys.* **2015**, *113*, 184–215, DOI: [10.1080/00268976.2014.952696](https://doi.org/10.1080/00268976.2014.952696) (pages 42, 89, 155, 167, 232).
- [94] Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648, DOI: [10.1063/1.464913](https://doi.org/10.1063/1.464913) (pages 42, 88).
- [95] Stephens, P. J.; Devlin, F. J.; Ashvar, C. S.; Chabalowski, C. F.; Frisch, M. J. Theoretical calculation of vibrational circular dichroism spectra. *Faraday Discuss.* **1994**, *99*, 103, DOI: [10.1039/fd9949900103](https://doi.org/10.1039/fd9949900103) (pages 42, 88).
- [96] Hehre, W. J.; Ditchfield, R.; Pople, J. A. Self-Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian-Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1972**, *56*, 2257–2261, DOI: [10.1063/1.1677527](https://doi.org/10.1063/1.1677527) (pages 42, 88).
- [97] Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. a. Self-consistent molecular orbital methods. XXIII. A polarization-type basis set for second-row elements. *J. Chem. Phys.* **1982**, *77*, 3654–3665, DOI: [10.1063/1.444267](https://doi.org/10.1063/1.444267) (pages 42, 88).
- [98] Pulay, P. Convergence acceleration of iterative sequences. The case of SCF iteration. *Chem. Phys. Lett.* **1980**, *73*, 393–398, DOI: [10.1016/0009-2614\(80\)80396-4](https://doi.org/10.1016/0009-2614(80)80396-4) (pages 42, 201).
- [99] Merrick, J. P.; Moran, D.; Radom, L. An Evaluation of Harmonic Vibrational Frequency Scale Factors. *J. Phys. Chem. A* **2007**, *111*, 11683–11700, DOI: [10.1021/jp073974n](https://doi.org/10.1021/jp073974n) (pages 42, 46).
- [100] Khaliullin, R. Z.; Head-Gordon, M.; Bell, A. T. An Efficient Self-Consistent Field Method for Large Systems of Weakly Interacting Components. *J. Chem. Phys.* **2006**, *124*, 204105, DOI: [10.1063/1.2191500](https://doi.org/10.1063/1.2191500) (pages 42, 46, 85, 88, 159, 167).
- [101] Anthony, J. L.; Anderson, J. L.; Maginn, E. J.; Brennecke, J. F. Anion Effects on Gas Solubility in Ionic Liquids. *J. Phys. Chem. B* **2005**, *109*, 6366–6374, DOI: [10.1021/jp0464041](https://doi.org/10.1021/jp0464041) (pages 44, 145).
- [102] Khaliullin, R. Z.; Cobar, E. A.; Lochan, R. C.; Bell, A. T.; Head-Gordon, M. Unravelling the Origin of Intermolecular Interactions Using Absolutely Localized Molec-

- ular Orbitals. *J. Phys. Chem. A* **2007**, *111*, 8753–8765, DOI: [10.1021/jp073685z](https://doi.org/10.1021/jp073685z) (pages [46](#), [85](#), [88](#), [159](#)).
- [103] Khaliullin, R. Z.; Bell, A. T.; Head-Gordon, M. Analysis of Charge Transfer Effects in Molecular Complexes Based on Absolutely Localized Molecular Orbitals. *J. Chem. Phys.* **2008**, *128*, 184112, DOI: [10.1063/1.2912041](https://doi.org/10.1063/1.2912041) (pages [46](#), [49](#), [85](#), [88](#), [159](#)).
- [104] Ramos-Cordoba, E.; Lambrecht, D. S.; Head-Gordon, M. Charge-transfer and the hydrogen bond: Spectroscopic and structural implications from electronic structure calculations. *Faraday Discuss.* **2011**, *150*, 345–362, DOI: [10.1039/C1FD00004G](https://doi.org/10.1039/C1FD00004G) (pages [47](#), [86](#), [96](#)).
- [105] Bakowies, D.; Thiel, W. Hybrid Models for Combined Quantum Mechanical and Molecular Mechanical Approaches. *J. Phys. Chem.* **1996**, *100*, 10580–10594, DOI: [10.1021/jp9536514](https://doi.org/10.1021/jp9536514) (page [48](#)).
- [106] Hu, L.; Söderhjelm, P.; Ryde, U. On the convergence of QM/MM energies. *J. Chem. Theory Comput.* **2011**, *7*, 761–777, DOI: [10.1021/ct100530r](https://doi.org/10.1021/ct100530r) (page [48](#)).
- [107] Breen, K. J.; DeBlase, A. F.; Guasco, T. L.; Voora, V. K.; Jordan, K. D.; Nagata, T.; Johnson, M. A. Bottom-Up View of Water Network-Mediated CO<sub>2</sub> Reduction Using Cryogenic Cluster Ion Spectroscopy and Direct Dynamics Simulations. *J. Phys. Chem. A* **2012**, *116*, 903–912, DOI: [10.1021/jp209493v](https://doi.org/10.1021/jp209493v) (page [48](#)).
- [108] Muraoka, A.; Inokuchi, Y.; Hammer, N. I.; Shin, J.-W.; Johnson, M. A.; Nagata, T. Structural Evolution of the [(CO<sub>2</sub>)<sub>n</sub>(H<sub>2</sub>O)]<sup>-</sup> Cluster Anions: Quantifying the Effect of Hydration on the Excess Charge Accommodation Motif. *J. Phys. Chem. A* **2009**, *113*, 8942–8948, DOI: [10.1021/jp903578e](https://doi.org/10.1021/jp903578e) (page [48](#)).
- [109] Lee, A. J.; Rick, S. W. The effects of charge transfer on the properties of liquid water. *J. Chem. Phys.* **2011**, *134*, 184507, DOI: [10.1063/1.3589419](https://doi.org/10.1063/1.3589419) (page [49](#)).
- [110] Khaliullin, R. Z.; Bell, A. T.; Head-Gordon, M. Electron Donation in the Water-Water Hydrogen Bond. *Chem. - A Eur. J.* **2009**, *15*, 851–855, DOI: [10.1002/chem.200802107](https://doi.org/10.1002/chem.200802107) (page [49](#)).
- [111] Hochstrasser, R. M. Two-Dimensional IR-spectroscopy: Polarization Anisotropy Effects. *Chem. Phys.* **2001**, *266*, 273–284, DOI: [10.1016/S0301-0104\(01\)00232-4](https://doi.org/10.1016/S0301-0104(01)00232-4) (page [58](#)).

- [112] Botan, V.; Hamm, P. Intramolecular vibrational energy relaxation in nitrous acid (HONO). *J. Chem. Phys.* **2008**, *129*, 164506, DOI: [10.1063/1.2996355](https://doi.org/10.1063/1.2996355) (page 59).
- [113] Martin, J. M. L.; Taylor, P. R.; Lee, T. J. Accurate ab initio quartic force fields for the N<sub>2</sub>O and CO<sub>2</sub> molecules. *Chem. Phys. Lett.* **1993**, *205*, 535–542, DOI: [10.1016/0009-2614\(93\)80009-E](https://doi.org/10.1016/0009-2614(93)80009-E) (page 59).
- [114] Dressier, S.; Thiel, W. Anharmonic force fields from density functional theory. *Chem. Phys. Lett.* **1997**, *273*, 71–78 (page 59).
- [115] Culp, J. T.; Goodman, A. L.; Chirdon, D.; Sankar, S. G.; Matranga, C. Mechanism for the Dynamic Adsorption of CO<sub>2</sub> and CH<sub>4</sub> in a Flexible Linear Chain Coordination Polymer as Determined from In Situ Infrared Spectroscopy. *J. Phys. Chem. C* **2010**, *114*, 2184–2191, DOI: [10.1021/jp908202s](https://doi.org/10.1021/jp908202s) (page 59).
- [116] Cunliffe-Jones, D. B. Perturbation of Some Vibrational Bands in Solution. *Spectrochim. Acta Part A* **1969**, *25*, 779 (page 59).
- [117] Kazarian, S. G.; Vincent, M. F.; Bright, F. V.; Liotta, C. L.; Eckert, C. A. Specific Intermolecular Interaction of Carbon Dioxide with Polymers. *J. Am. Chem. Soc.* **1996**, *118*, 1729–1736 (page 59).
- [118] Dobrowolski, J. C.; Jamróz, M. H. Infrared Evidence for CO<sub>2</sub> Electron Donor–Acceptor Complexes. *J. Mol. Struct.* **1992**, *275*, 211–219, DOI: [10.1016/0022-2860\(92\)80196-0](https://doi.org/10.1016/0022-2860(92)80196-0) (page 59).
- [119] Giammanco, C. H.; Kramer, P. L.; Yamada, S. A.; Nishida, J.; Tamimi, A.; Fayer, M. D. Coupling of Carbon Dioxide Stretch and Bend Vibrations Reveals Thermal Population Dynamics in an Ionic Liquid. *J. Phys. Chem. B* **2016**, *120*, 549–556, DOI: [10.1021/acs.jpccb.5b11454](https://doi.org/10.1021/acs.jpccb.5b11454) (pages 70, 134, 147).
- [120] Kaintz, A.; Baker, G.; Benesi, A.; Maroncelli, M. Solute diffusion in ionic liquids, NMR measurements and comparisons to conventional solvents. *J. Phys. Chem. B* **2013**, *117*, 11697–708, DOI: [10.1021/jp405393d](https://doi.org/10.1021/jp405393d) (page 72).
- [121] Kaintz, A.; Baker, G.; Benesi, A.; Maroncelli, M. Correction to "Solute Diffusion in Ionic Liquids, NMR Measurements and Comparisons to Conventional Solvents". *J. Phys. Chem. B* **2014**, *118*, 5615–5615, DOI: [10.1021/jp504170a](https://doi.org/10.1021/jp504170a) (page 72).



- [122] Li, W.; Zhang, Z.; Han, B.; Hu, S.; Xie, Y.; Yang, G. Effect of water and organic solvents on the ionic dissociation of ionic liquids. *J. Phys. Chem. B* **2007**, *111*, 6452–6456, DOI: [10.1021/jp071051m](https://doi.org/10.1021/jp071051m) (pages 73, 74).
- [123] Tokuda, H.; Tsuzuki, S.; Susan, M. A. B. H.; Hayamizu, K.; Watanabe, M. How ionic are room-temperature ionic liquids? An indicator of the physicochemical properties. *J. Phys. Chem. B* **2006**, *110*, 19593–19600, DOI: [10.1021/jp064159v](https://doi.org/10.1021/jp064159v) (page 74).
- [124] Domańska, U.; Królikowska, M. Density and Viscosity of Binary Mixtures of Thiocyanate Ionic Liquids + Water as a Function of Temperature. *J. Solution Chem.* **2012**, *41*, 1422–1445, DOI: [10.1007/s10953-012-9875-7](https://doi.org/10.1007/s10953-012-9875-7) (page 74).
- [125] Sánchez, L. G.; Espel, J. R.; Onink, F.; Meindersma, G. W.; de Haan, A. B. Density, Viscosity, and Surface Tension of Synthesis Grade Imidazolium, Pyridinium, and Pyrrolidinium Based Room Temperature Ionic Liquids. *J. Chem. Eng. Data* **2009**, *54*, 2803–2812, DOI: [10.1021/je800710p](https://doi.org/10.1021/je800710p) (page 74).
- [126] Zhang, Y.; Maginn, E. J. Direct Correlation between Ionic Liquid Transport Properties and Ion Pair Lifetimes: A Molecular Dynamics Study. *J. Phys. Chem. Lett.* **2015**, *6*, 700–705, DOI: [10.1021/acs.jpcllett.5b00003](https://doi.org/10.1021/acs.jpcllett.5b00003) (page 73).
- [127] Steinel, T.; Asbury, J. B.; Corcelli, S. A.; Lawrence, C. P.; Skinner, J. L.; Fayer, M. D. Water dynamics: Dependence on local structure probed with vibrational echo correlation spectroscopy. *Chem. Phys. Lett.* **2004**, *386*, 295–300, DOI: [10.1016/j.cplett.2004.01.042](https://doi.org/10.1016/j.cplett.2004.01.042) (pages 75, 133, 134).
- [128] Asbury, J. B.; Steinel, T.; Stromberg, C.; Corcelli, S. A.; Lawrence, C. P.; Skinner, J. L.; Fayer, M. D. Water Dynamics: Vibrational Echo Correlation Spectroscopy and Comparison to Molecular Dynamics Simulations. *J. Phys. Chem. A* **2004**, *108*, 1107–1119, DOI: [10.1021/jp036266k](https://doi.org/10.1021/jp036266k) (pages 75, 134).
- [129] Corcelli, S. A.; Lawrence, C. P.; Asbury, J. B.; Steinel, T.; Fayer, M. D.; Skinner, J. L. Spectral Diffusion in a Fluctuating Charge Model of Water. *J. Chem. Phys.* **2004**, *121*, 8897–8900, DOI: [10.1063/1.1803532](https://doi.org/10.1063/1.1803532) (page 75).
- [130] Choi, J.-H.; Oh, K.-I.; Lee, H.; Lee, C.; Cho, M. Nitrile and Thiocyanate IR Probes: Quantum Chemistry Calculation Studies and Multivariate Least-square Fitting Analysis. *J. Chem. Phys.* **2008**, *128*, 134506, DOI: [10.1063/1.2844787](https://doi.org/10.1063/1.2844787) (page 75).



- [131] Lindquist, B. A.; Corcelli, S. A. Nitrile Groups as Vibrational Probes: Calculations of the C≡N Infrared Absorption Line Shape of Acetonitrile in Water and Tetrahydrofuran. *J. Phys. Chem. B* **2008**, *112*, 6301–6303, DOI: [10.1021/jp802039e](https://doi.org/10.1021/jp802039e) (pages 75, 87).
- [132] Li, S.; Schmidt, J. R.; Piryatinski, A.; Lawrence, C. P.; Skinner, J. L. Vibrational Spectral Diffusion of Azide in Water. *J. Phys. Chem. B* **2006**, *110*, 18933–18938, DOI: [10.1021/jp057568k](https://doi.org/10.1021/jp057568k) (pages 75, 87).
- [133] Li, S.; Schmidt, J. R.; Corcelli, S. A.; Lawrence, C. P.; Skinner, J. L. Approaches for the Calculation of Vibrational Frequencies in Liquids: Comparison to Benchmarks for Azide/Water Clusters. *J. Chem. Phys.* **2006**, *124*, 204110, DOI: [10.1063/1.2200690](https://doi.org/10.1063/1.2200690) (pages 75, 134).
- [134] Terranova, Z. L.; Corcelli, S. A. Molecular dynamics investigation of the vibrational spectroscopy of isolated water in an ionic liquid. *J. Phys. Chem. B* **2014**, *118*, 8264–8272, DOI: [10.1021/jp501631m](https://doi.org/10.1021/jp501631m) (pages 75, 87, 134, 136–138).
- [135] Berquist, E. J.; Daly, C. A.; Brinzer, T.; Bullard, K. K.; Campbell, Z. M.; Corcelli, S. A.; Garrett-Roe, S.; Lambrecht, D. S. Modeling Carbon Dioxide Vibrational Frequencies in Ionic Liquids: I. Ab Initio Calculations. *J. Phys. Chem. B* **2017**, *121*, 208–220, DOI: [10.1021/acs.jpccb.6b09489](https://doi.org/10.1021/acs.jpccb.6b09489) (pages 83, 134, 135, 233).
- [136] D’Alessandro, D. M.; Smit, B.; Long, J. R. Carbon dioxide capture: Prospects for new materials. *Angew. Chemie - Int. Ed.* **2010**, *49*, 6058–6082, DOI: [10.1002/anie.201000431](https://doi.org/10.1002/anie.201000431) (page 84).
- [137] Corvo, M. C.; Sardinha, J.; Casimiro, T.; Marin, G.; Seferin, M.; Einloft, S.; Menezes, S. C.; Dupont, J.; Cabrita, E. J. A Rational Approach to CO<sub>2</sub> Capture by Imidazolium Ionic Liquids: Tuning CO<sub>2</sub> Solubility by Cation Alkyl Branching. *ChemSusChem* **2015**, *8*, 1935–1946, DOI: [10.1002/cssc.201500104](https://doi.org/10.1002/cssc.201500104) (page 84).
- [138] Hayes, R.; Warr, G. G.; Atkin, R. Structure and Nanostructure in Ionic Liquids. *Chem. Rev.* **2015**, *115*, 6357–6426, DOI: [10.1021/cr500411q](https://doi.org/10.1021/cr500411q) (page 84).
- [139] Gurkan, B.; Goodrich, B. F.; Mindrup, E. M.; Ficke, L. E.; Massel, M.; Seo, S.; Senftle, T. P.; Wu, H.; Glaser, M. F.; Shah, J. K.; Maginn, E. J.; Brennecke, J. F.; Scheider, W. F. Molecular Design of High Capacity, Low Viscosity, Chemically Tunable Ionic

- Liquids for CO<sub>2</sub> Capture. *J. Phys. Chem. Lett.* **2010**, *1*, 3494–3499, DOI: [10.1021/jz101533k](https://doi.org/10.1021/jz101533k) (page 84).
- [140] Grice, K. A.; Groenenboom, M. C.; Manuel, J. D. A.; Sovereign, M. A.; Keith, J. A. Examining the selectivity of borohydride for carbon dioxide and bicarbonate reduction in protic conditions. *Fuel* **2015**, *150*, 139–145, DOI: [10.1016/j.fuel.2015.02.007](https://doi.org/10.1016/j.fuel.2015.02.007) (page 84).
- [141] Rosen, B. A.; Salehi-Khojin, A.; Thorson, M. R.; Zhu, W.; Whipple, D. T.; Kenis, P. J. A.; Masel, R. I. Ionic Liquid-Mediated Selective Conversion of CO<sub>2</sub> to CO at Low Overpotentials. *Science* **2011**, *334*, 643–644, DOI: [10.1126/science.1209786](https://doi.org/10.1126/science.1209786) (page 84).
- [142] Kumar, B.; Llorente, M.; Froehlich, J.; Dang, T.; Sathrum, A.; Kubiak, C. P. Photochemical and Photoelectrochemical Reduction of CO<sub>2</sub>. *Annu. Rev. Phys. Chem.* **2012**, *63*, 541–569, DOI: [10.1146/annurev-physchem-032511-143759](https://doi.org/10.1146/annurev-physchem-032511-143759) (page 84).
- [143] Daly, C. A.; Berquist, E. J.; Brinzer, T.; Garrett-Roe, S.; Lambrecht, D. S.; Corcelli, S. A. Modeling Carbon Dioxide Vibrational Frequencies in Ionic Liquids: II. Spectroscopic Map. *J. Phys. Chem. B* **2016**, *120*, 12633–12642, DOI: [10.1021/acs.jpcc.6b09509](https://doi.org/10.1021/acs.jpcc.6b09509) (pages 85–88, 96, 99–102, 120, 132, 233).
- [144] Choi, J.-H.; Oh, K.-I.; Lee, H.; Lee, C.; Cho, M. Nitrile and thiocyanate IR Probes: Quantum chemistry calculation studies and multivariate least-square fitting analysis. *J. Chem. Phys.* **2008**, *128*, 134506, DOI: [10.1063/1.2844787](https://doi.org/10.1063/1.2844787) (pages 87, 133).
- [145] Schmidt, J. R.; Corcelli, S. A. Infrared Absorption Line Shapes in the Classical Limit: A Comparison of the Classical Dipole and Fluctuating Frequency Approximations. *J. Chem. Phys.* **2008**, *128*, 184504, DOI: [10.1063/1.2917349](https://doi.org/10.1063/1.2917349) (page 87).
- [146] Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **1988**, *38*, 3098–3100, DOI: [10.1103/PhysRevA.38.3098](https://doi.org/10.1103/PhysRevA.38.3098) (page 88).
- [147] Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785–789, DOI: [10.1103/PhysRevB.37.785](https://doi.org/10.1103/PhysRevB.37.785) (page 88).

- [148] Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions. *J. Chem. Phys.* **1980**, *72*, 650–654, DOI: [10.1063/1.438955](https://doi.org/10.1063/1.438955) (page 88).
- [149] McLean, A. D.; Chandler, G. S. Contracted Gaussian basis sets for molecular calculations. I. Second row atoms, Z=11-18. *J. Chem. Phys.* **1980**, *72*, 5639, DOI: [10.1063/1.438980](https://doi.org/10.1063/1.438980) (page 88).
- [150] Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. V. R. Efficient diffuse function-augmented basis sets for anion calculations. III. The 3-21+G basis set for first-row elements, Li-F. *J. Comput. Chem.* **1983**, *4*, 294–301, DOI: [10.1002/jcc.540040303](https://doi.org/10.1002/jcc.540040303) (page 88).
- [151] Dunning Jr, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007, DOI: [10.1063/1.456153](https://doi.org/10.1063/1.456153) (page 88).
- [152] Dunning T.H., J.; Peterson, K. A.; Wilson, A. K. Gaussian basis sets for use in correlated molecular calculations. X. The atoms aluminum through argon revisited. *J. Chem. Phys.* **2001**, *114*, 9244–9253, DOI: [10.1063/1.1367373](https://doi.org/10.1063/1.1367373) (page 88).
- [153] Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminform.* **2012**, *4*, 17, DOI: [10.1186/1758-2946-4-17](https://doi.org/10.1186/1758-2946-4-17) (page 88).
- [154] Hutchison, G. R. Avogadro: an open-source molecular builder and visualization tool. Version 1.1.1., 2012 (page 88).
- [155] Lao, K. U.; Herbert, J. M. Energy Decomposition Analysis with a Well-Defined Charge-Transfer Term for Interpreting Intermolecular Interactions. *J. Chem. Theory Comput.* **2016**, *acs.jctc.6b00155*, DOI: [10.1021/acs.jctc.6b00155](https://doi.org/10.1021/acs.jctc.6b00155) (pages 89, 97).
- [156] Jeziorski, B.; Moszynski, R.; Szalewicz, K. Perturbation Theory Approach to Intermolecular Potential Energy Surfaces of van der Waals Complexes. *Chem. Rev.* **1994**, *94*, 1887–1930, DOI: [10.1021/cr00031a008](https://doi.org/10.1021/cr00031a008) (pages 89, 135, 192).

- [157] Hohenstein, E. G.; Sherrill, C. D. Density fitting and Cholesky decomposition approximations in symmetry-adapted perturbation theory: Implementation and application to probe the nature of  $\pi - \pi$  interactions in linear acenes. *J. Chem. Phys.* **2010**, *132*, DOI: [10.1063/1.3426316](https://doi.org/10.1063/1.3426316) (pages 89, 135).
- [158] Hohenstein, E. G.; Parrish, R. M.; Sherrill, C. D.; Turney, J. M.; Schaefer, H. F. Large-scale symmetry-adapted perturbation theory computations via density fitting and Laplace transformation techniques: Investigating the fundamental forces of DNA-intercalator interactions. *J. Chem. Phys.* **2011**, *135*, DOI: [10.1063/1.3656681](https://doi.org/10.1063/1.3656681) (pages 89, 135, 144).
- [159] Turney, J. M.; Simmonett, A. C.; Parrish, R. M.; Hohenstein, E. G.; Evangelista, F. A.; Fermann, J. T.; Mintz, B. J.; Burns, L. A.; Wilke, J. J.; Abrams, M. L.; Russ, N. J.; Leininger, M. L.; Janssen, C. L.; Seidl, E. T.; Allen, W. D.; Schaefer, H. F.; King, R. A.; Valeev, E. F.; Sherrill, C. D.; Crawford, T. D. Psi4: An open-source ab initio electronic structure program. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 556–565, DOI: [10.1002/wcms.93](https://doi.org/10.1002/wcms.93) (page 89).
- [160] Papajak, E.; Truhlar, D. G. Convergent partially augmented basis sets for post-Hartree-Fock calculations of molecular properties and reaction barrier heights. *J. Chem. Theory Comput.* **2011**, *7*, 10–18, DOI: [10.1021/ct1005533](https://doi.org/10.1021/ct1005533) (page 89).
- [161] Gianinetti, E.; Raimondi, M.; Tornaghi, E. Modification of the Roothaan equations to exclude BSSE from molecular interaction calculations. *Int. J. Quantum Chem.* *60*, 157–166, DOI: [10.1002/\(SICI\)1097-461X\(1996\)60:1<157::AID-QUA17>3.0.CO;2-C](https://doi.org/10.1002/(SICI)1097-461X(1996)60:1<157::AID-QUA17>3.0.CO;2-C) (page 89).
- [162] Levinson, N. M.; Bolte, E. E.; Miller, C. S.; Corcelli, S. A.; Boxer, S. G. Phosphate Vibrations Probe Local Electric Fields and Hydration in Biomolecules. *J. Am. Chem. Soc.* **2011**, *133*, 13236–13239, DOI: [10.1021/ja2042589](https://doi.org/10.1021/ja2042589) (pages 90, 133).
- [163] Kinnaman, C. S.; Cremeens, M. E.; Romesberg, F. E.; Corcelli, S. A. Infrared Line Shape of an  $\alpha$ -Carbon Deuterium-Labeled Amino Acid. *J. Am. Chem. Soc.* **2006**, *128*, 13334–13335, DOI: [10.1021/ja064468z](https://doi.org/10.1021/ja064468z) (page 90).

- [164] Colbert, D. T.; Miller, W. H. A novel discrete variable representation for quantum mechanical reactive scattering via the S-matrix Kohn method. *J. Chem. Phys.* **1992**, *96*, 1982–1991, DOI: [10.1063/1.462100](https://doi.org/10.1063/1.462100) (page 90).
- [165] Seideman, T.; Miller, W. H. Quantum mechanical reaction probabilities via a discrete variable representation-absorbing boundary condition Green’s function. *J. Chem. Phys.* **1992**, *97*, 2499–2514, DOI: [10.1063/1.463088](https://doi.org/10.1063/1.463088) (page 90).
- [166] Cremer, D. Møller-Plesset perturbation theory: from small molecule methods to methods for thousands of atoms. *WIREs Comput Mol Sci* **2011**, *1*, 509–530, DOI: [10.1002/wcms.58](https://doi.org/10.1002/wcms.58) (pages 96, 120).
- [167] Lambrecht, D. S.; Clark, G. N. I.; Head-Gordon, T.; Head-Gordon, M. Exploring the Rich Energy Landscape of Sulfate-Water clusters  $\text{SO}_4^{2-}(\text{H}_2\text{O})_{n=3-7}$ : An Electronic Structure Approach. *J. Phys. Chem. A* **2011**, *115*, 11438–11454, DOI: [10.1021/jp206064n](https://doi.org/10.1021/jp206064n) (page 96).
- [168] Choi, J.-H.; Cho, M. Vibrational solvatochromism and electrochromism of infrared probe molecules containing  $\text{C}\equiv\text{O}$ ,  $\text{C}\equiv\text{N}$ ,  $\text{C}=\text{O}$ , or  $\text{C}-\text{F}$  vibrational chromophore. *J. Chem. Phys.* **2011**, *134*, 154513, DOI: [10.1063/1.3580776](https://doi.org/10.1063/1.3580776) (pages 101, 133, 135).
- [169] Řezáč, J. Cuby: An integrative framework for computational chemistry. *J. Comput. Chem.* **2016**, *37*, 1230–1237, DOI: [10.1002/jcc.24312](https://doi.org/10.1002/jcc.24312) (page 102).
- [170] Řezáč, J. Cuby – Ruby framework for computational chemistry. (page 102).
- [171] Collins, J. R.; Gallup, G. A. The full versus the virtual counterpoise correction for basis set superposition error in self-consistent field calculations. *Chem. Phys. Lett.* **1986**, *123*, 56–61, DOI: [10.1016/0009-2614\(86\)87014-2](https://doi.org/10.1016/0009-2614(86)87014-2) (page 105).
- [172] Mentel, M.; Baerends, E. J. Can the counterpoise correction for basis set superposition effect be justified? *J. Chem. Theory Comput.* **2014**, *10*, 252–267, DOI: [10.1021/ct400990u](https://doi.org/10.1021/ct400990u) (page 105).
- [173] O’Boyle, N. M.; Tenderholt, A. L.; Langner, K. M. cclib: A library for package-independent computational chemistry algorithms. *J. Comput. Chem.* **2008**, *29*, 839–845, DOI: [10.1002/jcc.20823](https://doi.org/10.1002/jcc.20823) (page 117).
- [174] Berquist, E. J.; Langner, K. M.; O’Boyle, N. M.; Tenderholt, A. L. Release of cclib version 1.3.1., 2015, DOI: [10.5281/zenodo.15108](https://doi.org/10.5281/zenodo.15108) (pages 117, 233).

- [175] Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. Climbing the Density Functional Ladder: Non-Empirical Meta-Generalized Gradient Approximation Designed for Molecules and Solids. *Phys. Rev. Lett.* **2003**, *91*, 146401, DOI: [10.1103/PhysRevLett.91.146401](https://doi.org/10.1103/PhysRevLett.91.146401) (page 121).
- [176] Lin, Y. S.; Li, G. D.; Mao, S. P.; Chai, J. D. Long-range corrected hybrid density functionals with improved dispersion corrections. *J. Chem. Theory Comput.* **2013**, *9*, 263–272, DOI: [10.1021/ct300715s](https://doi.org/10.1021/ct300715s) (pages 121, 134).
- [177] Feyereisen, M.; Fitzgerald, G.; Komornicki, A. Use of approximate integrals in ab initio theory. An application in MP2 energy calculations. *Chem. Phys. Lett.* **1993**, *208*, 359–363, DOI: [10.1016/0009-2614\(93\)87156-W](https://doi.org/10.1016/0009-2614(93)87156-W) (page 121).
- [178] Weigend, F.; Häser, M. RI-MP2: first derivatives and global consistency. *Theor. Chem. Acc.* **1997**, *97*, 331–340, DOI: [10.1007/s002140050269](https://doi.org/10.1007/s002140050269) (page 121).
- [179] DiStasio Jr., R. A.; Steele, R. P.; Rhee, Y. M.; Shao, Y.; Head-Gordon, M. An improved algorithm for analytical gradient evaluation in resolution-of-the-identity second-order Møller-Plesset perturbation theory: Application to alanine tetrapeptide conformational analysis. *J. Comput. Chem.* **2007**, *28*, 839–856, DOI: [10.1002/jcc.20604](https://doi.org/10.1002/jcc.20604) (page 121).
- [180] Weigend, F.; Köhn, A.; Hättig, C. Efficient use of the correlation consistent basis sets in resolution of the identity MP2 calculations. *J. Chem. Phys.* **2002**, *116*, 3175–3183, DOI: [10.1063/1.1445115](https://doi.org/10.1063/1.1445115) (page 121).
- [181] Wishart, J. F. Energy applications of ionic liquids. *Energy Environ. Sci.* **2009**, *2*, 956, DOI: [10.1039/b906273d](https://doi.org/10.1039/b906273d) (page 133).
- [182] Armand, M.; Endres, F.; MacFarlane, D. R.; Ohno, H.; Scrosati, B. Ionic-liquid materials for the electrochemical challenges of the future. *Nat. Mater.* **2009**, *8*, 621–9, DOI: [10.1038/nmat2448](https://doi.org/10.1038/nmat2448) (page 133).
- [183] Patel, D. D.; Lee, J. M. Applications of ionic liquids. *Chem. Rec.* **2012**, *12*, 329–355, DOI: [10.1002/tcr.201100036](https://doi.org/10.1002/tcr.201100036) (page 133).
- [184] Crosthwaite, J. M.; Muldoon, M. J.; Dixon, J. K.; Anderson, J. L.; Brennecke, J. F. Phase transition and decomposition temperatures, heat capacities and viscosities of

- pyridinium ionic liquids. *J. Chem. Thermodyn.* **2005**, *37*, 559–568, DOI: [10.1016/j.jct.2005.03.013](https://doi.org/10.1016/j.jct.2005.03.013) (page 133).
- [185] Seki, S.; Kobayashi, T.; Kobayashi, Y.; Takei, K.; Miyashiro, H.; Hayamizu, K.; Tsuzuki, S.; Mitsugi, T.; Umebayashi, Y. Effects of cation and anion on physical properties of room-temperature ionic liquids. *J. Mol. Liq.* **2010**, *152*, 9–13, DOI: [10.1016/j.molliq.2009.10.008](https://doi.org/10.1016/j.molliq.2009.10.008) (page 133).
- [186] Tokuda, H.; Hayamizu, K.; Ishii, K.; Susan, M. A. B. H.; Watanabe, M. Physicochemical properties and structures of room temperature ionic liquids. 2. variation of alkyl chain length in imidazolium cation. *J. Phys. Chem. B* **2005**, *109*, 6103–6110, DOI: [10.1021/jp044626d](https://doi.org/10.1021/jp044626d) (page 133).
- [187] Anthony, J. L.; Maginn, E. J.; Brennecke, J. F. Solubilities and Thermodynamic Properties of Gases in the Ionic Liquid 1-n-Butyl-3-methylimidazolium Hexafluorophosphate. *J. Phys. Chem. B* **2002**, *106*, 7315–7320, DOI: [10.1021/jp020631a](https://doi.org/10.1021/jp020631a) (page 133).
- [188] Shiflett, M. B.; Yokozeki, A. Solubilities and diffusivities of carbon dioxide in ionic liquids: [bmim][PF<sub>6</sub>] and [bmim][BF<sub>4</sub>]. *Ind. Eng. Chem. Res.* **2005**, *44*, 4453–4464, DOI: [10.1021/ie058003d](https://doi.org/10.1021/ie058003d) (page 133).
- [189] Tamimi, A.; Fayer, M. D. Ionic Liquid Dynamics Measured with 2D IR and IR Pump-Probe Experiments on a Linear Anion and the Influence of Potassium Cations. *J. Phys. Chem. B* **2016**, *120*, 5842–5854, DOI: [10.1021/acs.jpccb.6b00409](https://doi.org/10.1021/acs.jpccb.6b00409) (page 133).
- [190] Lee, K.-K.; Park, K.-H.; Kwon, D.; Choi, J.-H.; Son, H.; Park, S.; Cho, M. Ion-pairing dynamics of Li<sup>+</sup> and SCN<sup>-</sup> in dimethylformamide solution: Chemical exchange two-dimensional infrared spectroscopy. *J. Chem. Phys.* **2011**, *134*, 064506, DOI: [10.1063/1.3552961](https://doi.org/10.1063/1.3552961) (pages 133, 134).
- [191] Giammanco, C. H.; Kramer, P. L.; Yamada, S. A.; Nishida, J.; Tamimi, A.; Fayer, M. D. Carbon dioxide in an ionic liquid: Structural and rotational dynamics. *J. Chem. Phys.* **2016**, *144*, 104506, DOI: [10.1063/1.4943390](https://doi.org/10.1063/1.4943390) (pages 134, 147).
- [192] Asbury, J. B.; Steinel, T.; Kwak, K.; Corcelli, S. A.; Lawrence, C. P.; Skinner, J. L.; Fayer, M. D. Dynamics of water probed with vibrational echo correlation spectroscopy. *J. Chem. Phys.* **2004**, *121*, 12431–12446, DOI: [10.1063/1.1818107](https://doi.org/10.1063/1.1818107) (page 134).



- [193] Bakker, H. J.; Skinner, J. L. Vibrational spectroscopy as a probe of structure and dynamics in liquid water. *Chem. Rev.* **2010**, *110*, 1498–1517, DOI: [10.1021/cr9001879](https://doi.org/10.1021/cr9001879) (page 134).
- [194] Fecko, C. J.; Eaves, J. D.; Loparo, J. J.; Tokmakoff, A.; Geissler, P. L. Ultrafast hydrogen-bond dynamics in the infrared spectroscopy of water. *Science* **2003**, *301*, 1698–1702, DOI: [10.1126/science.1087251](https://doi.org/10.1126/science.1087251) (pages 134, 147).
- [195] Eaves, J. D.; Loparo, J. J.; Fecko, C. J.; Roberts, S. T.; Tokmakoff, A.; Geissler, P. L. Hydrogen Bonds in Liquid Water are Broken Only Fleetingly. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13019–13022, DOI: [10.1073/pnas.0505125102](https://doi.org/10.1073/pnas.0505125102) (page 134).
- [196] Gruenbaum, S. M.; Tainter, C. J.; Shi, L.; Ni, Y.; Skinner, J. L. Robustness of frequency, transition dipole, and coupling maps for water vibrational spectroscopy. *J. Chem. Theory Comput.* **2013**, *9*, 3109–3117, DOI: [10.1021/ct400292q](https://doi.org/10.1021/ct400292q) (pages 134, 153).
- [197] Jansen, T. L.; Auer, B. M.; Yang, M.; Skinner, J. L. Two-dimensional infrared spectroscopy and ultrafast anisotropy decay of water. *J. Chem. Phys.* **2010**, *132*, DOI: [10.1063/1.3454733](https://doi.org/10.1063/1.3454733) (page 134).
- [198] Loparo, J. J.; Roberts, S. T.; Tokmakoff, A. Multidimensional infrared spectroscopy of water. I. Vibrational dynamics in two-dimensional IR line shapes. *J. Chem. Phys.* **2006**, *125*, 194521, DOI: [10.1063/1.2382895](https://doi.org/10.1063/1.2382895) (page 134).
- [199] Loparo, J. J.; Roberts, S. T.; Tokmakoff, A. Multidimensional infrared spectroscopy of water. II. Hydrogen bond switching dynamics. *J. Chem. Phys.* **2006**, *125*, 194522, DOI: [10.1063/1.2382896](https://doi.org/10.1063/1.2382896) (page 134).
- [200] Nibbering, E. T. J.; Elsaesser, T. Ultrafast Vibrational Dynamics of Hydrogen Bonds in the Condensed Phase. *Chem. Rev.* **2004**, *104*, 1887–1914, DOI: [10.1021/cr020694p](https://doi.org/10.1021/cr020694p) (page 134).
- [201] Nicodemus, R. A.; Corcelli, S. A.; Skinner, J. L.; Tokmakoff, A. Collective hydrogen bond reorganization in water studied with temperature-dependent ultrafast infrared spectroscopy. *J. Phys. Chem. B* **2011**, *115*, 5604–5616, DOI: [10.1021/jp111434u](https://doi.org/10.1021/jp111434u) (page 134).



- [202] Nicodemus, R. A.; Ramasesha, K.; Roberts, S. T.; Tokmakoff, A. Hydrogen Bond Rearrangements in Water Probed with Temperature-Dependent 2D IR. *J. Phys. Chem. Lett.* **2010**, *1*, 1068–1072, DOI: [10.1021/jz100138z](https://doi.org/10.1021/jz100138z) (page [134](#)).
- [203] Ramasesha, K.; Roberts, S. T.; Nicodemus, R. A.; Mandal, A.; Tokmakoff, A. Ultrafast 2D IR Anisotropy of Water Reveals Reorientation During Hydrogen-Bond Switching. *J. Chem. Phys.* **2011**, *135*, 54509, DOI: [10.1063/1.3623008](https://doi.org/10.1063/1.3623008) (page [134](#)).
- [204] Roberts, S. T.; Ramasesha, K.; Tokmakoff, A. Structural rearrangements in water viewed through two-dimensional infrared spectroscopy. *Acc. Chem. Res.* **2009**, *42*, 1239–1249, DOI: [10.1021/ar900088g](https://doi.org/10.1021/ar900088g) (page [134](#)).
- [205] Auer, B.; Kumar, R.; Schmidt, J. R.; Skinner, J. L. Hydrogen bonding and Raman, IR, and 2D-IR spectroscopy of dilute HOD in liquid D<sub>2</sub>O. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 14215–14220, DOI: [10.1073/pnas.0701482104](https://doi.org/10.1073/pnas.0701482104) (page [134](#)).
- [206] Auer, B. M.; Skinner, J. L. IR and Raman spectra of liquid water: Theory and interpretation. *J. Chem. Phys.* **2008**, *128*, 224511, DOI: [10.1063/1.2925258](https://doi.org/10.1063/1.2925258) (page [134](#)).
- [207] Corcelli, S. A.; Lawrence, C. P.; Skinner, J. L. Combined electronic structure/molecular dynamics approach for ultrafast infrared spectroscopy of dilute HOD in liquid H<sub>2</sub>O and D<sub>2</sub>O. *J. Chem. Phys.* **2004**, *120*, 8107–8117, DOI: [10.1063/1.1683072](https://doi.org/10.1063/1.1683072) (pages [134](#), [135](#)).
- [208] Hayashi, T.; la Cour Jansen, T.; Zhuang, W.; Mukamel, S. Collective Solvent Coordinates for the Infrared Spectrum of HOD in D<sub>2</sub>O Based on an ab Initio Electrostatic Map. *J. Phys. Chem. A* **2005**, *109*, 64–82, DOI: [10.1021/jp046685x](https://doi.org/10.1021/jp046685x) (page [134](#)).
- [209] Jansen, T. L. C.; Cringus, D.; Pshenichnikov, M. S. Dissimilar dynamics of coupled water vibrations. *J. Phys. Chem. A* **2009**, *113*, 6260–6265, DOI: [10.1021/jp900480r](https://doi.org/10.1021/jp900480r) (page [134](#)).
- [210] Li, F.; Skinner, J. L. Infrared and Raman line shapes for ice Ih. II. H<sub>2</sub>O and D<sub>2</sub>O. *J. Chem. Phys.* **2010**, *133*, 244504–244504, DOI: [10.1063/1.3516460](https://doi.org/10.1063/1.3516460) (page [134](#)).
- [211] Lin, Y. S.; Auer, B. M.; Skinner, J. L. Water structure, dynamics, and vibrational spectroscopy in sodium bromide solutions. *J. Chem. Phys.* **2009**, *131*, 144511, DOI: [10.1063/1.3242083](https://doi.org/10.1063/1.3242083) (pages [134](#), [138](#), [143](#)).

- [212] Paarmann, A.; Hayashi, T.; Mukamel, S.; Miller, R. J. D. Nonlinear response of vibrational excitons: Simulating the two-dimensional infrared spectrum of liquid water. *J. Chem. Phys.* **2009**, *130*, 204110, DOI: [10.1063/1.3139003](https://doi.org/10.1063/1.3139003) (page 134).
- [213] Pieniazek, P. A.; Lin, Y.-S.; Chowdhary, J.; Ladanyi, B. M.; Skinner, J. L. Vibrational spectroscopy and dynamics of water confined inside reverse micelles. *J. Phys. Chem. B* **2009**, *113*, 15017–28, DOI: [10.1021/jp906784t](https://doi.org/10.1021/jp906784t) (page 134).
- [214] Shi, L.; Gruenbaum, S. M.; Skinner, J. L. Interpretation of IR and Raman Line Shapes for H<sub>2</sub>O and D<sub>2</sub>O Ice Ih. *J. Phys. Chem. B* **2012**, *116*, 13821–13830, DOI: [10.1021/jp3059239](https://doi.org/10.1021/jp3059239) (page 134).
- [215] Skinner, J. L.; Auer, B. N.; Lin, Y.-S., Vibrational Line Shapes, Spectral Diffusion, and Hydrogen Bonding in Liquid Water. In *Adv. Chem. Phys.* Rice, S. A., Ed.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2009; Vol. 142, p 59, DOI: [10.1002/9780470475935.ch2](https://doi.org/10.1002/9780470475935.ch2) (page 134).
- [216] Tainter, C. J.; Skinner, J. L. The water hexamer: Three-body interactions, structures, energetics, and OH-stretch spectroscopy at finite temperature. *J. Chem. Phys.* **2012**, *137*, 2351–104304, DOI: [10.1063/1.4746157](https://doi.org/10.1063/1.4746157) (page 134).
- [217] Yang, M.; Skinner, J. L. Time-averaging approximation in the interaction picture: Absorption line shapes for coupled chromophores with application to liquid water. *J. Chem. Phys.* **2011**, *135*, 154114–174108, DOI: [10.1063/1.3654005](https://doi.org/10.1063/1.3654005) (page 134).
- [218] Laage, D.; Hynes, J. T. Do more strongly hydrogen-bonded water molecules reorient more slowly? *Chem. Phys. Lett.* **2006**, *433*, 80–85, DOI: [10.1016/j.cplett.2006.11.035](https://doi.org/10.1016/j.cplett.2006.11.035) (page 134).
- [219] Laage, D.; Hynes, J. T. A molecular jump mechanism of water reorientation. *Science* **2006**, *311*, 832–835, DOI: [10.1126/science.1122154](https://doi.org/10.1126/science.1122154) (page 134).
- [220] Laage, D.; Hynes, J. T. On the molecular mechanism of water reorientation. *J. Phys. Chem. B* **2008**, *112*, 14230–14242, DOI: [10.1021/jp805217u](https://doi.org/10.1021/jp805217u) (page 134).
- [221] Laage, D.; Stirnemann, G.; Hynes, J. T. Water jump reorientation and ultrafast vibrational spectroscopy. *J. Photochem. Photobiol. A Chem.* **2012**, *234*, 75–82, DOI: [10.1016/j.jphotochem.2011.12.022](https://doi.org/10.1016/j.jphotochem.2011.12.022) (page 134).

- [222] Laage, D.; Stirnemann, G.; Sterpone, F.; Hynes, J. T. Water jump reorientation: From theoretical prediction to experimental observation. *Acc. Chem. Res.* **2012**, *45*, 53–62, DOI: [10.1021/ar200075u](https://doi.org/10.1021/ar200075u) (page 134).
- [223] Smith, J. D.; Cappa, C. D.; Wilson, K. R.; Cohen, R. C.; Geissler, P. L.; Saykally, R. J. Unified description of temperature-dependent hydrogen-bond rearrangements in liquid water. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 14171–14174, DOI: [10.1073/pnas.0506899102](https://doi.org/10.1073/pnas.0506899102) (page 134).
- [224] Oh, K.-I.; Choi, J.-H.; Lee, J.-H.; Han, J.-B.; Lee, H.; Cho, M. Nitrile and thiocyanate IR probes: Molecular dynamics simulation studies. *J. Chem. Phys.* **2008**, *128*, 154504, DOI: [10.1063/1.2904558](https://doi.org/10.1063/1.2904558) (pages 135, 143).
- [225] Corcelli, S. A.; Skinner, J. L. Infrared and Raman Line Shapes of Dilute HOD in Liquid H<sub>2</sub>O and D<sub>2</sub>O from 10 to 90 degrees C. *J. Phys. Chem. A* **2005**, *109*, 6154–6165, DOI: [10.1021/jp0506540](https://doi.org/10.1021/jp0506540) (pages 135, 141).
- [226] Schmidt, J. R.; Corcelli, S. A.; Skinner, J. L. Pronounced non-Condon effects in the ultrafast infrared spectroscopy of water. *J. Chem. Phys.* **2005**, *123*, 44513, DOI: [10.1063/1.1961472](https://doi.org/10.1063/1.1961472) (pages 135, 141, 155).
- [227] Błasiak, B.; Lee, H.; Cho, M. Vibrational solvatochromism: Towards systematic approach to modeling solvation phenomena. *J. Chem. Phys.* **2013**, *139*, 44111, DOI: [10.1063/1.4816041](https://doi.org/10.1063/1.4816041) (page 135).
- [228] Błasiak, B.; Cho, M. Vibrational solvatochromism. II. A first-principle theory of solvation-induced vibrational frequency shift based on effective fragment potential method. *J. Chem. Phys.* **2014**, *140*, 164107, DOI: [10.1063/1.4872040](https://doi.org/10.1063/1.4872040) (page 135).
- [229] Błasiak, B.; Cho, M. Vibrational solvatochromism. III. Rigorous treatment of the dispersion interaction contribution. *J. Chem. Phys.* **2015**, *143*, 164111, DOI: [10.1063/1.4934667](https://doi.org/10.1063/1.4934667) (pages 135, 153).
- [230] Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys.* **1995**, *117*, 1–19, DOI: [10.1006/jcph.1995.1039](https://doi.org/10.1006/jcph.1995.1039) (pages 136, 137).
- [231] Andreussi, O.; Marzari, N. Transport properties of room-temperature ionic liquids from classical molecular dynamics. *J. Chem. Phys.* **2012**, *137*, 044508, DOI: [10.1063/1.4737388](https://doi.org/10.1063/1.4737388) (page 136).

- [232] Forester, T. R.; Smith, W. SHAKE, rattle, and roll: Efficient constraint algorithms for linked rigid bodies. *J. Comput. Chem.* **1998**, *19*, 102–111, DOI: [10.1002/\(SICI\)1096-987X\(19980115\)19:1<102::AID-JCC9>3.0.CO;2-T](https://doi.org/10.1002/(SICI)1096-987X(19980115)19:1<102::AID-JCC9>3.0.CO;2-T) (page 137).
- [233] Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341, DOI: [10.1016/0021-9991\(77\)90098-5](https://doi.org/10.1016/0021-9991(77)90098-5) (page 137).
- [234] Sprenger, K. G.; Jaeger, V. W.; Pfaendtner, J. The general AMBER force field (GAFF) can accurately predict thermodynamic and transport properties of many ionic liquids. *J. Phys. Chem. B* **2015**, *119*, 5882–5895, DOI: [10.1021/acs.jpcc.5b00689](https://doi.org/10.1021/acs.jpcc.5b00689) (page 137).
- [235] Wang, J.; Wolf, R. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174, DOI: [10.1002/jcc.20035/pdf](https://doi.org/10.1002/jcc.20035/pdf) (page 137).
- [236] Singh, U. C.; Kollman, P. A. A combined *ab initio* quantum mechanical and molecular mechanical method for carrying out simulations on complex molecular systems: Applications to the CH<sub>3</sub>Cl + Cl<sup>-</sup> exchange reaction and gas phase protonation of polyethers. *J. Comput. Chem.* **1986**, *7*, 718, DOI: [10.1002/jcc.540070604](https://doi.org/10.1002/jcc.540070604) (page 137).
- [237] Liu, Z.; Huang, S.; Wang, W. A refined force field for molecular simulation of imidazolium-based ionic liquids. *J. Phys. Chem. B* **2004**, *108*, 12978–12989, DOI: [10.1021/jp048369o](https://doi.org/10.1021/jp048369o) (page 137).
- [238] Chaban, V. V.; Voroshylova, I. V.; Kalugin, O. N. A New Force Field Model for the Simulation of Transport Properties of Imidazolium-Based Ionic Liquids. *Phys. Chem. Chem. Phys.* **2011**, *13*, 7910–20, DOI: [10.1039/c0cp02778b](https://doi.org/10.1039/c0cp02778b) (page 137).
- [239] Schröder, C. Comparing reduced partial charge models with polarizable simulations of ionic liquids. *Phys. Chem. Chem. Phys.* **2012**, *14*, 3089–3102, DOI: [10.1039/c2cp23329k](https://doi.org/10.1039/c2cp23329k) (page 137).
- [240] Potoff, J. J.; Siepmann, J. I. Vapor-liquid equilibria of mixtures containing alkanes, carbon dioxide, and nitrogen. *AIChE J.* **2001**, *47*, 1676–1682, DOI: [10.1002/aic.690470719](https://doi.org/10.1002/aic.690470719) (page 137).

- [241] Perez-Blanco, M. E.; Maginn, E. J. Molecular dynamics simulations of CO<sub>2</sub> at an ionic liquid interface: Adsorption, ordering, and interfacial crossing. *J. Phys. Chem. B* **2010**, *114*, 11827–11837, DOI: [10.1021/jp103862v](https://doi.org/10.1021/jp103862v) (page 137).
- [242] Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An  $N \cdot \log(N)$  method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089, DOI: [10.1063/1.464397](https://doi.org/10.1063/1.464397) (page 137).
- [243] Małolepsza, E.; Straub, J. E. Empirical maps for the calculation of amide I vibrational spectra of proteins from classical molecular dynamics simulations. *J. Phys. Chem. B* **2014**, *118*, 7848–7855, DOI: [10.1021/jp412827s](https://doi.org/10.1021/jp412827s) (page 138).
- [244] Fennell, C. J.; Gezelter, J. D. Is the Ewald summation still necessary? Pairwise alternatives to the accepted standard for long-range electrostatics. *J. Chem. Phys.* **2006**, *124*, 234104, DOI: [10.1063/1.2206581](https://doi.org/10.1063/1.2206581) (page 139).
- [245] Loparo, J. J.; Roberts, S. T.; Nicodemus, R. A.; Tokmakoff, A. Variation of the transition dipole moment across the OH stretching band of water. *Chem. Phys.* **2007**, *341*, 218–229, DOI: [10.1016/j.chemphys.2007.06.056](https://doi.org/10.1016/j.chemphys.2007.06.056) (page 141).
- [246] Miller, C. S.; Ploetz, E. A.; Cremeens, M. E.; Corcelli, S. A. Carbon-deuterium vibrational probes of peptide conformation: Alanine dipeptide and glycine dipeptide. *J. Chem. Phys.* **2009**, *130*, DOI: [10.1063/1.3100185](https://doi.org/10.1063/1.3100185) (page 143).
- [247] Aki, S. N. V. K.; Mellein, B. R.; Saurer, E. M.; Brennecke, J. F. High-pressure phase behavior of carbon dioxide with imidazolium-based ionic liquids. *J. Phys. Chem. B* **2004**, *108*, 20355–20365, DOI: [10.1021/jp046895](https://doi.org/10.1021/jp046895) (page 145).
- [248] Muldoon, M. J.; Aki, S. N. V. K.; Anderson, J. L.; Dixon, J. K.; Brennecke, J. F. Improving Carbon Dioxide Solubility in Ionic Liquids. *J. Phys. Chem. B* **2007**, *111*, 9001–9009, DOI: [10.1021/jp071897q](https://doi.org/10.1021/jp071897q) (page 145).
- [249] Hou, Y.; Baltus, R. E. Experimental Measurement of the Solubility and Diffusivity of CO<sub>2</sub> in Room-Temperature Ionic Liquids Using a Transient Thin-Liquid-Film Method. *Ind. Eng. Chem. Res.* **2007**, *46*, 8166–8175, DOI: [10.1021/ie070501u](https://doi.org/10.1021/ie070501u) (page 145).
- [250] Stenger, J.; Madsen, D.; Hamm, P.; Nibbering, E. T. J.; Elsaesser, T. Ultrafast Vibrational Dephasing of Liquid Water. *Phys. Rev. Lett.* **2001**, *87*, 027401, DOI: [10.1103/PhysRevLett.87.027401](https://doi.org/10.1103/PhysRevLett.87.027401) (page 147).

- [251] Rosenthal, S. J.; Xie, X.; Du, M.; Fleming, G. R. Femtosecond solvation dynamics in acetonitrile: Observation of the inertial contribution to the solvent response. *J. Chem. Phys.* **1991**, *95*, 4715, DOI: [10.1063/1.461742](https://doi.org/10.1063/1.461742) (page 147).
- [252] Brinzer, T.; Daly, C. A.; Corcelli, S. A.; Garrett-Roe, S. Frequency and rotational dynamics of CO<sub>2</sub> in 1-butyl-3-methylimidazolium bis-(trifluoromethylsulfonyl)imide. **2018** (page 153).
- [253] Błasiak, B.; Ritchie, A. W.; Webb, L. J.; Cho, M. Vibrational solvatochromism of nitrile infrared probes: beyond the vibrational Stark dipole approach. *Phys. Chem. Chem. Phys.* **2016**, *18*, 18094–18111, DOI: [10.1039/C6CP01578F](https://doi.org/10.1039/C6CP01578F) (page 153).
- [254] Carbonnière, P.; Dargelos, A.; Pouchan, C. The VCI-P code: an iterative variation–perturbation scheme for efficient computations of anharmonic vibrational levels and IR intensities of polyatomic molecules. *Theor. Chem. Acc.* **2010**, *125*, 543–554, DOI: [10.1007/s00214-009-0689-7](https://doi.org/10.1007/s00214-009-0689-7) (page 155).
- [255] Downing, H. D.; Krohn, B. J.; Hunt, R. H. Coriolis intensity perturbations in  $\Pi - \Sigma$  bands of CO<sub>2</sub>. *J. Mol. Spectrosc.* **1975**, *55*, 66–80, DOI: [10.1016/0022-2852\(75\)90252-0](https://doi.org/10.1016/0022-2852(75)90252-0) (pages 155, 156).
- [256] Berquist, E. J.; Lambrecht, D. S. A First Principles Approach for Partitioning Linear Response Properties into Additive and Cooperative Contributions. **2018**, DOI: [10.26434/chemrxiv.5773968.v1](https://doi.org/10.26434/chemrxiv.5773968.v1) (pages 158, 234).
- [257] Autschbach, J.; Ziegler, T. Double perturbation theory: a powerful tool in computational coordination chemistry. *Coord. Chem. Rev.* **2003**, *238-239*, 83–126, DOI: [10.1016/S0010-8545\(02\)00287-4](https://doi.org/10.1016/S0010-8545(02)00287-4) (page 158).
- [258] Gagliardi, L.; Lindh, R.; Karlström, G. Local properties of quantum chemical systems: The LoProp approach. *J. Chem. Phys.* **2004**, *121*, 4494–4500, DOI: [10.1063/1.1778131](https://doi.org/10.1063/1.1778131) (page 160).
- [259] Söderhjelm, P.; Krogh, J. W.; Karlström, G.; Ryde, U.; Lindh, R. Accuracy of distributed multipoles and polarizabilities: Comparison between the LoProp and MpProp models. *J. Comput. Chem.* **2007**, *28*, 1083–1090, DOI: [10.1002/jcc.20632](https://doi.org/10.1002/jcc.20632) (page 160).

- [260] Mochizuki, Y.; Ishikawa, T.; Tanaka, K.; Tokiwa, H.; Nakano, T.; Tanaka, S. Dynamic polarizability calculation with fragment molecular orbital scheme. *Chem. Phys. Lett.* **2006**, *418*, 418–422, DOI: [10.1016/j.cplett.2005.11.014](https://doi.org/10.1016/j.cplett.2005.11.014) (page 160).
- [261] Neugebauer, J. On the calculation of general response properties in subsystem density functional theory. *J. Chem. Phys.* **2009**, *131*, 084104, DOI: [10.1063/1.3212883](https://doi.org/10.1063/1.3212883) (page 160).
- [262] Pavanello, M. On the subsystem formulation of linear-response time-dependent DFT. *J. Chem. Phys.* **2013**, *138*, 204118, DOI: [10.1063/1.4807059](https://doi.org/10.1063/1.4807059) (page 160).
- [263] Karna, S. P.; Dupuis, M. Frequency dependent nonlinear optical properties of molecules: Formulation and implementation in the HONDO program. *J. Comput. Chem.* **1991**, *12*, 487–504, DOI: [10.1002/jcc.540120409](https://doi.org/10.1002/jcc.540120409) (pages 160, 231).
- [264] Peng, D.; Li, S.; Peng, L.; Gu, F. L.; Yang, W. Time-Dependent Coupled Perturbed Hartree-Fock and Density-Functional-Theory Approach for Calculating Frequency-Dependent (Hyper)Polarizabilities with Nonorthogonal Localized Molecular Orbitals. *J. Chem. Theory Comput.* **2017**, *13*, 4101–4112, DOI: [10.1021/acs.jctc.7b00321](https://doi.org/10.1021/acs.jctc.7b00321) (page 160).
- [265] Jeziorski, B.; Moszynski, R.; Szalewicz, K. Perturbation Theory Approach to Intermolecular Potential Energy Surfaces of van der Waals Complexes. *Chem. Rev.* **1994**, *94*, 1887–1930, DOI: [10.1021/cr00031a008](https://doi.org/10.1021/cr00031a008) (page 160).
- [266] Closser, K. D.; Ge, Q.; Mao, Y.; Shao, Y.; Head-Gordon, M. Superposition of Fragment Excitations for Excited States of Large Clusters with Application to Helium Clusters. *J. Chem. Theory Comput.* **2015**, *11*, 5791–5803, DOI: [10.1021/acs.jctc.5b00703](https://doi.org/10.1021/acs.jctc.5b00703) (pages 161, 177).
- [267] Ge, Q.; Mao, Y.; White, A. F.; Epifanovsky, E.; Closser, K. D.; Head-Gordon, M. Simulating the absorption spectra of helium clusters ( $N = 70, 150, 231, 300$ ) using a charge transfer correction to superposition of fragment single excitations. *J. Chem. Phys.* **2017**, *146*, 044111, DOI: [10.1063/1.4973611](https://doi.org/10.1063/1.4973611) (pages 161, 164).
- [268] Liu, J.; Herbert, J. M. An efficient and accurate approximation to time-dependent density functional theory for systems of weakly coupled monomers. *J. Chem. Phys.* **2015**, *143*, 034106, DOI: [10.1063/1.4926837](https://doi.org/10.1063/1.4926837) (pages 161, 177).



- [269] Liu, J.; Herbert, J. M. Local Excitation Approximations to Time-Dependent Density Functional Theory for Excitation Energies in Solution. *J. Chem. Theory Comput.* **2016**, *12*, 157–166, DOI: [10.1021/acs.jctc.5b00828](https://doi.org/10.1021/acs.jctc.5b00828) (page 161).
- [270] Aidas, K.; Angeli, C.; Bak, K. L.; Bakken, V.; Bast, R.; Boman, L.; Christiansen, O.; Cimiraglia, R.; Coriani, S.; Dahle, P.; Dalskov, E. K.; Ekström, U.; Enevoldsen, T.; Eriksen, J. J.; Ettenhuber, P.; Fernández, B.; Ferrighi, L.; Fliegl, H.; Frediani, L.; Hald, K.; Halkier, A.; Hättig, C.; Heiberg, H.; Helgaker, T.; Hennum, A. C.; Hettema, H.; Hjertenæs, E.; Høst, S.; Høyvik, I.-M.; Iozzi, M. F.; Jansík, B.; Aa. Jensen, H. J.; Jonsson, D.; Jørgensen, P.; Kauczor, J.; Kirpekar, S.; Kjærgaard, T.; Klopper, W.; Knecht, S.; Kobayashi, R.; Koch, H.; Kongsted, J.; Krapp, A.; Kristensen, K.; Ligabue, A.; Lutnæs, O. B.; Melo, J. I.; Mikkelsen, K. V.; Myhre, R. H.; Neiss, C.; Nielsen, C. B.; Norman, P.; Olsen, J.; Olsen, J. M. H.; Osted, A.; Packer, M. J.; Pawłowski, F.; Pedersen, T. B.; Provasi, P. F.; Reine, S.; Rinkevicius, Z.; Ruden, T. A.; Ruud, K.; Rybkin, V. V.; Sałek, P.; Samson, C. C. M.; de Merás, A. S.; Saue, T.; Sauer, S. P. A.; Schimmelpfennig, B.; Sneskov, K.; Steindal, A. H.; Sylvester-Hvid, K. O.; Taylor, P. R.; Teale, A. M.; Tellgren, E. I.; Tew, D. P.; Thorvaldsen, A. J.; Thøgersen, L.; Vahtras, O.; Watson, M. A.; Wilson, D. J. D.; Ziolkowski, M.; Ågren, H. The Dalton quantum chemistry program system. *WIREs Comput. Mol. Sci.* **2014**, *4*, 269–284, DOI: [10.1002/wcms.1172](https://doi.org/10.1002/wcms.1172) (page 161).
- [271] Jørgensen, P.; Jensen, H. J. A.; Olsen, J. Linear response calculations for large scale multiconfiguration self-consistent field wave functions. *J. Chem. Phys.* **1988**, *89*, 3654–3661, DOI: [10.1063/1.454885](https://doi.org/10.1063/1.454885) (page 161).
- [272] Olsen, J.; Yeager, D. L.; Poul Jørgensen, P. Triplet excitation properties in large scale multiconfiguration linear response calculations. *J. Chem. Phys.* **1989**, *91*, 381–388, DOI: [10.1063/1.457471](https://doi.org/10.1063/1.457471) (page 161).
- [273] Foresman, J. B.; Head-Gordon, M.; Pople, J. A.; Frisch, M. J. Toward a systematic molecular orbital theory for excited states. *J. Phys. Chem.* **1992**, *96*, 135–149, DOI: [10.1021/j100180a030](https://doi.org/10.1021/j100180a030) (page 161).
- [274] Cheng, C. Y.; Ryley, M. S.; Peach, M. J.; Tozer, D. J.; Helgaker, T.; Teale, A. M. Molecular properties in the Tamm–Dancoff approximation: indirect nuclear spin–spin



- coupling constants. *Mol. Phys.* **2015**, *113*, 1937–1951, DOI: [10.1080/00268976.2015.1024182](https://doi.org/10.1080/00268976.2015.1024182) (page 162).
- [275] Stone, A. J.; Misquitta, A. J. Charge-transfer in Symmetry-Adapted Perturbation Theory. *Chem. Phys. Lett.* **2009**, *473*, 201–205, DOI: [10.1016/j.cplett.2009.03.073](https://doi.org/10.1016/j.cplett.2009.03.073) (page 165).
- [276] Berquist, E. J.; Daly, C. A.; Brinzer, T.; Bullard, K. K.; Campbell, Z. M.; Corcelli, S. A.; Garrett-Roe, S.; Lambrecht, D. S. Modeling Carbon Dioxide Vibrational Frequencies in Ionic Liquids: I. Ab Initio Calculations. *J. Phys. Chem. B* **2017**, *121*, 208–220, DOI: [10.1021/acs.jpcc.6b09489](https://doi.org/10.1021/acs.jpcc.6b09489) (page 165).
- [277] Mao, Y.; Horn, P. R.; Head-Gordon, M. Energy decomposition analysis in an adiabatic picture. *Phys. Chem. Chem. Phys.* **2017**, *19*, 5944–5958, DOI: [10.1039/C6CP08039A](https://doi.org/10.1039/C6CP08039A) (pages 166, 168).
- [278] Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305, DOI: [10.1039/B508541A](https://doi.org/10.1039/B508541A) (page 166).
- [279] Rappoport, D.; Furche, F. Property-optimized Gaussian basis sets for molecular response calculations. *J. Chem. Phys.* **2010**, *133*, 134105, DOI: [10.1063/1.3484283](https://doi.org/10.1063/1.3484283) (page 166).
- [280] Azar, R. J.; Horn, P. R.; Sundstrom, E. J.; Head-Gordon, M. Useful lower limits to polarization contributions to intermolecular interactions using a minimal basis of localized orthogonal orbitals: Theory and analysis of the water dimer. *J. Chem. Phys.* **2013**, *138*, 084102, DOI: [10.1063/1.4792434](https://doi.org/10.1063/1.4792434) (page 167).
- [281] Stoll, H.; Wagenblast, G.; Preuß, H. On the use of local basis sets for localized molecular orbitals. *Theor. Chim. Acta* **1980**, *57*, 169–178, DOI: [10.1007/BF00574903](https://doi.org/10.1007/BF00574903) (page 167).
- [282] Horn, P. R.; Head-Gordon, M. Polarization contributions to intermolecular interactions revisited with fragment electric-field response functions. *J. Chem. Phys.* **2015**, *143*, 114111, DOI: [10.1063/1.4930534](https://doi.org/10.1063/1.4930534) (page 172).
- [283] Sanderson, C.; Curtin, R. Armadillo: a template-based C++ library for linear algebra. *Journal of Open Source Software* **2016**, *1*, 26 (page 177).

- [284] Parrish, R. M.; Burns, L. A.; Smith, D. G. A.; Simmonett, A. C.; DePrince, A. E.; Hohenstein, E. G.; Bozkaya, U.; Sokolov, A. Y.; Di Remigio, R.; Richard, R. M.; Gonthier, J. F.; James, A. M.; McAlexander, H. R.; Kumar, A.; Saitow, M.; Wang, X.; Pritchard, B. P.; Verma, P.; Schaefer, H. F.; Patkowski, K.; King, R. A.; Valeev, E. F.; Evangelista, F. A.; Turney, J. M.; Crawford, T. D.; Sherrill, C. D. Psi4 1.1: An Open-Source Electronic Structure Program Emphasizing Automation, Advanced Libraries, and Interoperability. *J. Chem. Theory Comput.* **2017**, *13*, 3185–3197, DOI: [10.1021/acs.jctc.7b00174](https://doi.org/10.1021/acs.jctc.7b00174) (pages [177](#), [191](#), [196](#)).
- [285] O’Boyle, N. M.; Tenderholt, A. L.; Langner, K. M. cclib: A library for package-independent computational chemistry algorithms. *J. Comput. Chem.* **2008**, *29*, 839–845, DOI: [10.1002/jcc.20823](https://doi.org/10.1002/jcc.20823) (page [177](#)).
- [286] Berquist, E.; Hutchison, G.; Langner, K. M.; O’Boyle, N. M.; Tenderholt, A. L. Release of cclib version 1.5., 2016, DOI: [10.5281/zenodo.60670](https://doi.org/10.5281/zenodo.60670) (page [177](#)).
- [287] Smith, D.; Burns, L. A.; Sirianni, D. A.; Nascimento, D. R.; Kumar, A.; James, A. M.; Schriber, J. B.; Zhang, T.; Zhang, B.; Abbott, A. S.; Berquist, E. J.; Lechner, M. H.; dos Anjos Cunha, L.; Heide, A. G.; Waldrop, J. M.; King, R. A.; Simmonett, A. C.; Turney, J. M.; Schaefer, H. F.; Evangelista, F. A.; DePrince III, A. E.; Crawford, T. D.; Patkowski, K.; Sherrill, C. D. Psi4NumPy: An Interactive Quantum Chemistry Programming Environment for Reference Implementations and Rapid Development. **2018**, DOI: [10.26434/chemrxiv.5746059.v1](https://doi.org/10.26434/chemrxiv.5746059.v1) (pages [189](#), [233](#)).
- [288] Briling, K. R. Comment on "A new parametrizable model of molecular electronic structure" [J. Chem. Phys. 135, 134120 (2011)]. *J. Chem. Phys.* **2017**, *147*, 157101, DOI: [10.1063/1.5000525](https://doi.org/10.1063/1.5000525) (page [191](#)).
- [289] Laikov, D. N. A new parametrizable model of molecular electronic structure. *J. Chem. Phys.* **2011**, *135*, 134120, DOI: [10.1063/1.3646498](https://doi.org/10.1063/1.3646498) (page [191](#)).
- [290] Source code accompanying the comment [K. R. Briling, J. Chem. Phys. 147, 157101 (2017)] is available at <https://github.com/briling/qm>. (page [191](#)).
- [291] Density Functional Repository., Quantum Chemistry Group, CCLRC Daresbury Laboratory, Daresbury, Cheshire, WA4 4AD United Kingdom. (accessed September 11, 2017). (page [191](#)).

- [292] Lejaeghere, K.; Bihlmayer, G.; Björkman, T.; Blaha, P.; Blügel, S.; Blum, V.; Caliste, D.; Castelli, I. E.; Clark, S. J.; Dal Corso, A.; de Gironcoli, S.; Deutsch, T.; Dewhurst, J. K.; Di Marco, I.; Draxl, C.; Dułak, M.; Eriksson, O.; Flores-Livas, J. A.; Garrity, K. F.; Genovese, L.; Giannozzi, P.; Giantomassi, M.; Goedecker, S.; Gonze, X.; Grånäs, O.; Gross, E. K. U.; Gulans, A.; Gygi, F.; Hamann, D. R.; Hasnip, P. J.; Holzwarth, N. A. W.; Iuşan, D.; Jochym, D. B.; Jollet, F.; Jones, D.; Kresse, G.; Koepernik, K.; Küçükbenli, E.; Kvashnin, Y. O.; Loch, I. L. M.; Lubeck, S.; Marsman, M.; Marzari, N.; Nitzsche, U.; Nordström, L.; Ozaki, T.; Paulatto, L.; Pickard, C. J.; Poelmans, W.; Probert, M. I. J.; Refson, K.; Richter, M.; Rignanese, G.-M.; Saha, S.; Scheffler, M.; Schlipf, M.; Schwarz, K.; Sharma, S.; Tavazza, F.; Thunström, P.; Tkatchenko, A.; Torrent, M.; Vanderbilt, D.; van Setten, M. J.; Van Speybroeck, V.; Wills, J. M.; Yates, J. R.; Zhang, G.-X.; Cottenier, S. Reproducibility in density functional theory calculations of solids. *Science* **2016**, *351*, DOI: [10.1126/science.aad3000](https://doi.org/10.1126/science.aad3000) (page 191).
- [293] Van der Walt, S.; Colbert, S. C.; Varoquaux, G. The NumPy Array: A Structure for Efficient Numerical Computation. *Computing in Science Engineering* **2011**, *13*, 22–30, DOI: [10.1109/MCSE.2011.37](https://doi.org/10.1109/MCSE.2011.37) (pages 192, 194).
- [294] Blackford, L. S.; Demmel, J.; Dongarra, J.; Duff, I.; Hammarling, S.; Henry, G.; Heroux, M.; Kaufman, L.; Lumsdaine, A.; Petitet, A.; Pozo, R.; Remington, K.; Whaley, R. C. An Updated Set of Basic Linear Algebra Subprograms (BLAS). *ACM Trans. Math. Softw.* **2002**, *28*, 135–151, DOI: [10.1145/567806.567807](https://doi.org/10.1145/567806.567807) (page 192).
- [295] Open Source Initiative. (page 192).
- [296] Parr, R. G.; Yang, W., *Density-Functional Theory of Atoms and Molecules*; International Series of Monographs on Chemistry, Vol. 16; Oxford: New York, 1989 (page 192).
- [297] Bartlett, R. J. Many-Body Perturbation Theory and Coupled Cluster Theory for Electron Correlation in Molecules. *Annu. Rev. Phys. Chem.* **1981**, *32*, 359–401, DOI: [10.1146/annurev.pc.32.100181.002043](https://doi.org/10.1146/annurev.pc.32.100181.002043) (page 192).
- [298] Szalewicz, K. Symmetry-adapted perturbation theory of intermolecular forces. *WIREs Comput Mol Sci* **2012**, *2*, 254–272, DOI: [10.1002/wcms.86](https://doi.org/10.1002/wcms.86) (page 192).

- [299] Purvis, G. D.; Bartlett, R. J. A full coupled-cluster singles and doubles model: The inclusion of disconnected triples. *J. Chem. Phys.* **1982**, *76*, 1910–1918, DOI: [10.1063/1.443164](https://doi.org/10.1063/1.443164) (page 192).
- [300] Shavitt, I., Methods of Electronic Structure Theory. In, Schaefer, H. F., Ed.; Plenum Press: New York, 1977, pp 189–275 (page 192).
- [301] Sherrill, C. D.; Schaefer, H. F., Advances in Quantum Chemistry. In, Löwdin, P.-O., Ed.; Academic Press: New York, 1999, pp 143–269 (page 192).
- [302] Quantum Chemical Program Exchange, University of Indiana, Bloomington, Ind. (page 192).
- [303] Boyd, D. B., Pioneers of Quantum Chemistry. In, 2013; Chapter 8, pp 221–273 (page 192).
- [304] NumPy’s Array Interface. (page 194).
- [305] Raffennetti, R. C. Pre-processing two-electron integrals for efficient utilization in many-electron self-consistent field calculations. *Chem. Phys. Lett.* **1973**, *20*, 335–338, DOI: [10.1016/0009-2614\(73\)80060-0](https://doi.org/10.1016/0009-2614(73)80060-0) (page 197).
- [306] Smith, D. G. A.; Støter, F.-R.; McGibbon, R. T.; Werner, N. Optimized Einsum: v1.0., 2016, DOI: [10.5281/zenodo.160842](https://doi.org/10.5281/zenodo.160842) (page 198).
- [307] Stanton, J. F.; Gauss, J.; Watts, J. D.; Bartlett, R. J. A direct product decomposition approach for symmetry exploitation in many-body methods. I. Energy calculations. *J. Chem. Phys.* **1991**, *94*, 4334–4345, DOI: [10.1063/1.460620](https://doi.org/10.1063/1.460620) (page 199).
- [308] Python Anaconda. (page 200).
- [309] Fortenberry, R. C.; McDonald, A. R.; Shepherd, T. D.; Kennedy, M.; Sherrill, C. D., The Promise of Chemical Education: Addressing our Students Needs. In; Chapter 7, pp 85–98 (page 200).
- [310] Perez, F.; Granger, B. E. IPython: A System for Interactive Scientific Computing. *Computing in Science Engineering* **2007**, *9*, 21–29, DOI: [10.1109/MCSE.2007.53](https://doi.org/10.1109/MCSE.2007.53) (page 202).
- [311] Weiss, C. J. Scientific Computing for Chemists: An Undergraduate Course in Simulations, Data Processing, and Visualization. *J. Chem. Educ.* **2017**, *94*, 592–597, DOI: [10.1021/acs.jchemed.7b00078](https://doi.org/10.1021/acs.jchemed.7b00078) (page 202).

- [312] Jakubek, R. S.; Eskandari, A.; Berquist, E. J.; Lambrecht, D. S.; Asher, S. A. UV Resonance Raman Spectroscopy Monitors the  $\chi_2$  Angle of the Glutamine Sidechain. *in preparation* **2018** (page [234](#)).