

**NEW FUNCTIONAL FORMS AND PARAMETERIZATION METHODS FOR AB
INITIO, INTERMOLECULAR FORCE FIELD DEVELOPMENT**

by

Mary J. Van Vleet

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

(Chemistry)

at the

UNIVERSITY OF WISCONSIN–MADISON

2017

Date of final oral examination: 08/15/17

The dissertation is approved by the following members of the Final Oral Committee:

J.R. Schmidt, Associate Professor, Chemistry

Clark R. Landis, Professor, Chemistry

Qiang Cui, Professor, Chemistry

Arun Yethiraj, Professor, Chemistry

Reid Van Lehn, Assistant Professor, Chemical and Biological Engineering

© Copyright by Mary J. Van Vleet 2017
All Rights Reserved

Soli Deo gloria.

ACKNOWLEDGMENTS

It is customary for authors of academic books to include in their prefaces statements such as this: "I am indebted to ... for their invaluable help; however, any errors which remain are my sole responsibility." Occasionally an author will go further. Rather than say that if there are any mistakes then he is responsible for them, he will say that there will inevitably be some mistakes and he is responsible for them....

Although the shouldering of all responsibility is usually a social ritual, the admission that errors exist is not — it is often a sincere avowal of belief. But this appears to present a living and everyday example of a situation which philosophers have commonly dismissed as absurd; that it is sometimes rational to hold logically incompatible beliefs.

— DAVID C. MAKINSON (1965)

Above is the famous “preface paradox,” which illustrates how to use the `wbepi` environment for epigraphs at the beginning of chapters. You probably also want to thank the Academy.

CONTENTS

Contents iii

List of Tables vi

List of Figures vii

Abstract viii

Published Work and Work in Preparation ix

I Introduction 1

1 Introduction 2

1.1 *The Importance of Molecular Simulation* 2

2 Background 3

2.1 *Molecular Mechanics and the Theory of Intermolecular Forces* 3

2.1.1 The Many-Body Expansion 3

2.1.2 Energy Decomposition Schemes 3

Intramolecular Interactions 3

Electrostatics 3

Exchange 3

Induction 3

Dispersion 4

2.2 *Ab-Initio Force Field Development* 4

2.2.1 Electronic Structure Benchmarks 4

SAPT 4

Coupled-Cluster Methods 4

2.3 *ISA-based methods for force field development* 4

II Published Work 5

- 3 Isotropic Ab Initio Force Fields 6
- 4 Anisotropic Ab Initio Force Fields 7

III Unpublished Work 8

- 5 Ab Initio Force Fields using LMOEDA 9
 - 5.1 *Preface* 9
 - 5.2 *Introduction* 10
 - 5.3 *Background and Motivation* 11
 - 5.4 *Theory* 14
 - 5.5 *Methods* 14
- 6 Benchmark Database for Ab Initio Force Field Development 15

IV Practical Matters 16

- 7 Workflow for Intermolecular Force Field Development 18
 - 7.1 *Overview* 18
 - 7.2 *Codes* 18
 - 7.2.1 Molpro 18
 - 7.2.2 CamCASP 18
 - 7.2.3 Scripts 18
 - 7.3 *Geometry Generation* 18
 - 7.4 *SAPT Calculations* 18
 - 7.5 *Monomer-Based Parameterization* 18
 - 7.5.1 Multipoles 18
 - 7.5.2 ISA Exponents 18
 - 7.5.3 Polarization Charges 18
 - 7.5.4 Dispersion Coefficients 18

Jesse's Method	18
Alston's Method	18
7.6 <i>Dimer-Based Parameterization</i>	18
8 POInter: A Program for Intermolecular Force Field Optimization	19
8.1 <i>Overview</i>	19
8.2 <i>Documentation</i>	19
8.3 <i>Examples</i>	19
 V Conclusions and Future Work	 20
9 Future Work	21
10 Conclusions	22
Bibliography	23
Acronyms	24
Glossary	25

LIST OF TABLES

LIST OF FIGURES

- 5.1 Model potential energy surface (PES) for interactions between CO₂ and Mg-MOF-74. (Left) Interaction energies between CO₂ and a cluster model of Mg-MOF-74 (shown bottom right), computed at a CCSD(T)-f12 (black), Symmetry-Adapted Perturbation Theory (SAPT) (orange), and/or PBE0-D2 (blue) level of theory. Discrepancies between SAPT and CCSD(T)-f12 in the minimum-energy region of the potential have been highlighted. (Top right) The structure of CO₂-bound Mg-MOF-74. (Bottom right) The structure of the cluster model used for Mg-MOF-74, where the circled atom pair indicates the relevant Mg-O radius from the x-axis in the leftmost figure. 12

NEW FUNCTIONAL FORMS AND PARAMETERIZATION METHODS FOR AB INITIO, INTERMOLECULAR FORCE FIELD DEVELOPMENT

Mary J. Van Vleet

Under the supervision of Professor J.R. Schmidt

At the University of Wisconsin-Madison

FIXME: basically a placeholder; do not believe

I did some research, read a bunch of papers, published a couple myself, (pick one):

1. ran some experiments and made some graphs,
2. proved some theorems

and now I have a job. I've assembled this document in the last couple of months so you will let me leave. Thanks!

J.R. Schmidt

ABSTRACT

FIXME: basically a placeholder; do not believe

I did some research, read a bunch of papers, published a couple myself, (pick one):

1. ran some experiments and made some graphs,
2. proved some theorems

and now I have a job. I've assembled this document in the last couple of months so you will let me leave. Thanks!

PUBLISHED WORK AND WORK IN PREPARATION

- [1] Van Vleet, M. J.; Misquitta, A. J.; Stone, A. J.; Schmidt, J. R. *J. Chem. Theory Comput.* **2016**, *12*, 3851–3870.

Part I

Introduction

1 INTRODUCTION

1.1 The Importance of Molecular Simulation

This ref² is super cool!

What is molecular simulation? What types of problems can it solve? How does molecular simulation work? (Be sure to include solving Newton's EQs of motion and relevant details on the partition function and interaction energies!)

2 BACKGROUND

2.1 Molecular Mechanics and the Theory of Intermolecular Forces

What is a force field? What are the important components of a force field, and how do we model them?

2.1.1 The Many-Body Expansion

How do we break apart a force field into manageable pieces? Why does it make sense to break a force field into 2- and many-body components?

2.1.2 Energy Decomposition Schemes

Intramolecular Interactions

Brief commentary on the non-intermolecular portions of a force field

Electrostatics

Conceptual description of electrostatics: long-range multipoles and charge penetration

Exchange

Quantum-mechanically-based Pauli Exclusion. Theoretical grounds for exponential behavior

Induction

Charge transfer. Polarization. Polarization Damping.

Dispersion

Theoretical Formulation. Damping.

2.2 Ab-Initio Force Field Development

2.2.1 Electronic Structure Benchmarks

SAPT

General SAPT methodology. DFT-SAPT.

Coupled-Cluster Methods

CCSD(T). CCSD(T)-f12.

2.3 ISA-based methods for force field development

What is ISA? How can ISA be used to generate parameters for intermolecular force field development? What progress has been made from this approach?

Part II

Published Work

3 ISOTROPIC AB INITIO FORCE FIELDS

4 ANISOTROPIC AB INITIO FORCE FIELDS

Part III

Unpublished Work

5 AB INITIO FORCE FIELDS USING LMOEDA

5.1 Preface

The preceding sections have been devoted to a development of various methodologies for ab initio intermolecular force field development, all generally assuming that Symmetry-Adapted Perturbation Theory (SAPT) can be used as a benchmark electronic structure theory. Critically, and especially given the developments discussed in Chapter 4, we can now usually expect our model force field energies to be within ~ 1 kJ/mol of the SAPT reference values! In spite of this success, this high precision between the model and SAPT energies can only lead to experimentally-accurate molecular simulation provided that the SAPT energies themselves are accurate, either with respect to the exact underlying potential energy surface (PES) or (in practice) with respect to gold-standard CCSD(T) calculations. Indeed, for systems where SAPT and CCSD(T) disagree by several kJ/mol, there is little point in developing SAPT-based force fields with sub- kJ/mol accuracy! This limitation raises to two fundamentally important questions. First, for what types of systems might we expect SAPT to be inaccurate? Second, for the systems where SAPT and the exact PES are in disagreement, how must we modify our typical methodology for ab-initio force field development?

The purpose of this chapter is to partially address these questions, all within the specific context of force field development for Metal-Organic Frameworks (MOFs). Note that the results presented here were gathered from 2013–2015, so some important advances (namely those presented in Chapters 3 and 4) haven't been incorporated into the force fields presented here. This is probably to the detriment of the accuracy and transferability that might be possible with the LMO-EDA-based methodology, and (should this project be picked up in the future) it may be necessary to refit these force fields to the functional forms and monomer-based parameters discussed in Chapter 4.

5.2 Introduction

Metal-Organic Frameworks (MOFs) are an increasingly important class of compounds, fundamentally defined as porous materials comprised of inorganic nodes connected by organic linkers. Within this general motif, more than 20,000 compounds have been reported and studied,³ and this vast diversity of MOF materials shows great promise for chemical customization and optimization. Within the past two decades, a huge body of research has been devoted to the design and study of MOFs, and current applications range from gas separation and storage to catalysis and biomedical imaging.³

Somewhat recently, it has been discovered that so-called Coordinatively-Unsaturated (CUS) MOFs can be created by activation of solvent-coordinated inorganic nodes to yield exposed (or 'open') metal sites.⁴⁻⁶ These CUS-MOFs have been shown to exhibit excellent uptakes and selectivities in a number of gas separation and storage problems,^{4,5,7} making this family of compounds an excellent target for future investigation and materials design. Owing to the vast scope of hypothetical CUS-MOF materials, however, and the number of chemically-distinct targets for gas separation/storage, it is unlikely that experiment alone can be used to screen for new and promising CUS-MOF materials.⁸ Rather, a combination of experiment and computational modeling will be required to find (or possibly even rationally design) optimal CUS-MOFs.⁷⁻⁹

Despite the utility of computational studies, it remains challenging to develop molecular models for CUS-MOFs.⁶ Because the strong binding between metal and adsorbate leads to chemical environments substantially different from typical coordinatively-saturated MOFs, many standard force fields (such as UFF and DREIDING) which yield good predictions for these MOFs frequently (and substantially!) underpredict adsorption in CUS-MOFs, especially at low pressures.⁸⁻¹⁰ While CUS-MOFs can sometimes be studied using quantum mechanical means,^{9,11} clearly new and improved force fields will be required to perform in-depth simulations and large-scale screenings of these materials.

The goal of the present chapter is to present a general methodology for devel-

oping accurate and transferable force fields for CUS-MOFs. The current study is limited to a discussion of the MOF-74 series (a prototypical and well-studied CUS-MOF), however it is expected that the methods presented herein might also be applicable to other systems. After outlining this methodology (Sections 5.3 and 5.4), we next show how our force fields can be applied to accurately predict CO₂ adsorption isotherms in Mg-MOF-74. At the present time, we do not have results for other compounds in the M-MOF-74 series (M = Co, Cr, Cu, Fe, Mn, Ni, Ti, V, and Zn), largely as a result of technical challenges in the force field parameterization itself. We discuss these technical challenges in some detail, and conclude with our perspective on the challenges and opportunities associated with developing transferable force fields for these and other CUS-MOF systems.

5.3 Background and Motivation

Prior work in our group has shown how, at least for coordinatively-saturated MOFs, accurate and transferable force fields can be generated for a wide variety of systems by fitting force field parameters on a component-by-component basis to reproduce an ab initio SAPT energy decomposition.^{12,13} While full details for this force field development methodology can be found in refs. 13, 14, a short overview is as follows:

1. Generate a representative cluster model from which interaction parameters can be determined for each (new) pairwise interaction. An example cluster, used to parameterize Mg-CO₂ interactions in Mg-MOF-74, is shown in Fig. 5.1.
2. Compute, using DFT-SAPT (a variant of SAPT with monomer densities given by Density Functional Theory (DFT)), a series of representative dimer interaction energies for the model cluster. For the cluster model in Fig. 5.1, representative dimers were generated by varying the position of CO₂ with respect to the MOF cluster, and the corresponding DFT-SAPT total interaction energies are shown for a subset of representative points.

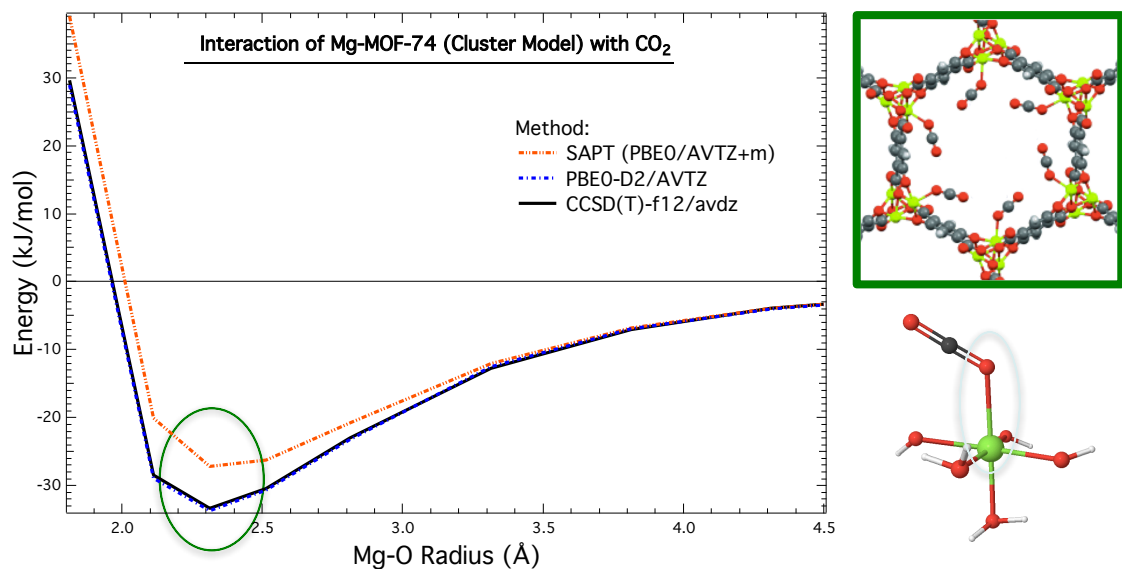


Figure 5.1: Model PES for interactions between CO_2 and Mg-MOF-74. (Left) Interaction energies between CO_2 and a cluster model of Mg-MOF-74 (shown bottom right), computed at a CCSD(T)-f12 (black), SAPT (orange), and/or PBE0-D2 (blue) level of theory. Discrepancies between SAPT and CCSD(T)-f12 in the minimum-energy region of the potential have been highlighted. (Top right) The structure of CO_2 -bound Mg-MOF-74. (Bottom right) The structure of the cluster model used for Mg-MOF-74, where the circled atom pair indicates the relevant Mg-O radius from the x-axis in the leftmost figure.

3. To determine partial charges for the system, generate representative clusters (as described in Section 5.5) for each the organic ligand and the inorganic node, and perform a Distributed Multipole Analysis (DMA) analysis on each cluster to determine charges for the overall periodic system.
4. For each component of the DFT-SAPT interaction energy, parameterize the relevant functional forms (as detailed in ref. 13 and Section 5.4) to reproduce the DFT-SAPT component energy.

Once parameterized, these SAPT-based MOF force fields can be used for calculating individual adsorption isotherms or even for high-throughput screening.¹⁴

In the generation of force fields for CUS-MOFs, we expect that many of the advantages of the above workflow (such as the component-by-component based parameterization and method for partial charge determination) will translate well to force field development for this subclass of MOF materials. Nevertheless, there are two reasons why a SAPT-based methodology cannot be used to generate such force fields. First, and as shown in Fig. 5.1 for a representative Mg-MOF-74 cluster model, we have empirically found SAPT to be in error compared to benchmark CCSD(T)-f12 calculations. SAPT is well-known to struggle with highly-polarizable and/or ionic systems, and so this error is perhaps not surprising. (Possible sources of the discrepancy between SAPT and CCSD(T)-f12 will be discussed in ??.) Nevertheless, and in the absence of fortuitous error cancellation, predictions from an ab initio force field can only be as good as the level of theory that they are parameterized against. Consequently, because SAPT underbinds CO₂ by a full 6 kJ/mol compared to CCSD(T)-f12, we would not expect to see good predictions for the CO₂ adsorption isotherm with a SAPT-based methodology, and a new strategy will be required.

As a second barrier to using a SAPT-based methodology, many of the compounds in the M-MOF-74 series are open-shell. Though this poses no fundamental issue, in practice most implementations of SAPT (aside from the seldom-used SAPT 2012 package developed in Krzysztof Szalewicz’s group at Delaware) do not allow for computations of open-shell systems, and indeed SAPT-based studies of open-shell compounds are very rare.¹⁵ For these reasons, a new electronic structure benchmark is highly preferable.

Based on the results for Mg-MOF-74, it is clear that, at least for CUS-MOFs, a new methodology is required which simultaneously keeps the important advantages of the old workflow (especially the component-by-component based parameterization, which is essential for generating transferable force fields) while overcoming the limitations of SAPT itself. Put differently, for CUS-MOFs we should seek a new electronic structure theory benchmark and associated energy decomposition analysis with the following qualities:

1. High accuracy with respect to CCSD(T)-f12 benchmark energies

2. Physically-meaningful energy decomposition into (at least) electrostatics, exchange, induction, and dispersion
3. For systems where SAPT and CCSD(T)-f12 agree, a quantitative correspondence between the energy decompositions of SAPT and the new method

Assuming these three qualities are met, we expect to be able to generate force fields for CUS-MOFs that are both highly accurate and maximally-compatible with previous force fields developed for coordinatively-saturated MOF systems. For reasons discussed below, and after a thorough comparison of available energy decomposition schemes, the LMO-EDA^{??} has been selected as our energy decomposition of choice, and we turn now to a full description of the method and its benefits.

5.4 Theory

5.5 Methods

Determining Partial Charges

6 BENCHMARK DATABASE FOR AB INITIO FORCE FIELD DEVELOPMENT

Part IV

Practical Matters

7 WORKFLOW FOR INTERMOLECULAR FORCE FIELD DEVELOPMENT

7.1 Overview

7.2 Codes

7.2.1 Molpro

7.2.2 CamCASP

7.2.3 Scripts

7.3 Geometry Generation

7.4 SAPT Calculations

7.5 Monomer-Based Parameterization

7.5.1 Multipoles

7.5.2 ISA Exponents

7.5.3 Polarization Charges

7.5.4 Dispersion Coefficients

Jesse's Method

Alston's Method

7.6 Dimer-Based Parameterization

Refer to POInter Code section

8 POINTER: A PROGRAM FOR INTERMOLECULAR FORCE FIELD OPTIMIZATION

8.1 Overview

8.2 Documentation

8.3 Examples

Part V

Conclusions and Future Work

9 FUTURE WORK

10 CONCLUSIONS

BIBLIOGRAPHY

-
- [2] Stone, A. J. *The Theory of Intermolecular Forces*, 2nd ed.; OUP Oxford, 2013.
- [3] Furukawa, H.; Cordova, K. E.; O’Keeffe, M.; Yaghi, O. M. *Science* (80-.). **2013**, 341, 1230444–1230444.
- [4] Millward, A. R.; Yaghi, O. M. *J. Am. Chem. Soc.* **2005**, 127, 17998–17999.
- [5] Dietzel, P. D. C. et al. *J. Mater. Chem.* **2009**, 19, 7362.
- [6] Dzubak, A. L.; Lin, L.-C.; Kim, J.; Swisher, J. a.; Poloni, R.; Maximoff, S. N.; Smit, B.; Gagliardi, L. *Nat. Chem.* **2012**, 4, 810–816.
- [7] Czaja, A. U.; Trukhan, N.; Müller, U. *Chem. Soc. Rev.* **2009**, 38, 1284.
- [8] Krishna, R.; van Baten, J. M. *Phys. Chem. Chem. Phys.* **2011**, 13, 10593–10616.
- [9] Getman, R. B.; Bae, Y.-s.; Wilmer, C. E.; Snurr, R. Q.; Carlo, M. *Adsorpt. J. Int. Adsorpt. Soc.* **2012**, 703–723.
- [10] Yazaydin, a. O.; Snurr, R. Q.; Park, T.-H.; Koh, K.; Liu, J.; Levan, M. D.; Benin, A. I.; Jakubczak, P.; Lanuza, M.; Galloway, D. B.; Low, J. J.; Willis, R. R. *J. Am. Chem. Soc.* **2009**, 131, 18198–9.
- [11] Valenzano, L.; Civalleri, B.; Chavan, S.; Palomino, G. T.; Areał'n, C. O.; Bordiga, S. *J. Phys. Chem. C* **2010**, 114, 11185–11191.
- [12] McDaniel, J. G.; Schmidt, J. R. *J. Phys. Chem. C* **2012**, 116, 14031–14039.
- [13] McDaniel, J. G.; Yu, K.; Schmidt, J. R. *J. Phys. Chem. C* **2012**, 116, 1892–1903.
- [14] McDaniel, J. G.; Li, S.; Tylianakis, E.; Snurr, R. Q.; Schmidt, J. R. *J. Phys. Chem. C* **2015**, 119, 3143–3152.
- [15] Żuchowski, P. *Chem. Phys. Lett.* **2008**, 450, 203–209.
- [] Su, P.; Li, H. *J. Chem. Phys.* **2009**, 131, 014102.

- [] Chen, Y.; Li, H. *J. Phys. Chem. A* **2010**, *114*, 11719–24.

ACRONYMS

C | D | M | P | S

C

CUS Coordinatively-Unsaturated. 10, 11, 13, 14

D

DFT Density Functional Theory. 11

DFT-SAPT Density Functional Theory Symmetry-Adapted Perturbation Theory.
11, 12

DMA Distributed Multipole Analysis. 12

M

MOF Metal-Organic Framework. 9–14

P

PES potential energy surface. vii, 9, 12, 25

S

SAPT Symmetry-Adapted Perturbation Theory. vii, 9, 11–14, *Glossary*: SAPT

GLOSSARY

C | L | S

C

CCSD(T) Coupled Cluster methods including singles, doubles, and perturbative triples excitations. CCSD(T). Given a sufficiently large (aVQZ or better) basis set, can be used as a ‘gold-standard’ estimate of the exact potential energy surface. 9

CCSD(T)-f12 Explicitly-correlated CCSD(T). Given a sufficiently large (aVDZ or aVTZ) basis set, used throughout this work as a ‘gold-standard’ estimate of the exact potential energy surface. vii, 12–14

L

LMO-EDA FILL . 9, 14

S

SAPT Symmetry-Adapted Perturbation Theory, a perturbative treatment of inter-molecular interactions which is pretty cool. vii, 9