# NOVEL 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

*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

## LIST OF TABLES

## LIST OF FIGURES

**NOVEL 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

---

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

# Part I

# Introduction

## 1.1 The Importance of Molecular Simulation

This ref[2] 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!)

# Part II

# Published Work

# Part III

# Unpublished Work

# 5 ab initio force fields using LMO-EDA

## 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 in the event that the SAPT energies themselves are accurate with respect to the exact underlying potential energy surface (PES). practice) Indeed, for systems where SAPT and CCSD(T) (a gold-standard electronic structure theory that closely matches the exact PES) disagree by several  kJ/mol, there is little advantage in developing SAPT-based force fields with sub- kJ/mol precision. 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 might we best modify our typical methodology for ab-initio force field development?

The purpose of this chapter is to partially address these two questions, all within the specific context of force field development for Coordinatively-Unsaturated (CUS) Metal-Organic Frameworks (MOFs). Importantly, the results presented here were gathered from 2012–2014, so some important advances (namely those presented in Chapters 3 and 4) have yet to be incorporated into the force fields presented here, probably to the determinent of the accuracy and transferability that might be possible with the LMO-EDA-based methodology. Should this project be picked up in the future, it will likely prove advantageous to refit the LMO-EDA-based force fields described herein 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, and are defined as porous materials containing 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 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 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 identify (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.[6,8,9] 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) that yield good predictions for these CS-MOFs can frequently (and substantially!) underpredict adsorption in CUS-MOFs.[8–10] Importantly, these underpredictions are especially prominent at low pressures, where metal-adsorbate interactions dominate.[8–10] While CUS-MOFs can sometimes be studied using quantum mechanical means,[9,11,12] clearly new and improved force fields will be required to perform in-depth simulations and large-scale screenings of these materials, and

such studies are already being undertaken.[13–16]

The goal of the present chapter is to present a general methodology for developing 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 $CO_2$ 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 techical 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 the M-MOF-74 series and other similar 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.[17,18] While full details for this force field development methodology can be found in Refs. 18, 19, a short workflow is given here for Mg-MOF-74:

1. Generate a representative cluster model from which interaction parameters can be determined for each pairwise interaction. An example cluster, used to parameterize Mg–$CO_2$ interactions in Mg-MOF-74, is shown in Fig. 5.1.

2. Using DFT-SAPT (a variant of SAPT with monomer densities given by Density Functional Theory (DFT)), compute 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 $CO_2$ with respect to the

MOF cluster, and the corresponding DFT-SAPT total interaction energies are shown for a subset of representative points.

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 partial charges for the overall system.

4. For each component of the DFT-SAPT interaction energy, parameterize the relevant functional forms (as detailed in Ref. 18 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.[19]

In the generation of force fields for CUS-MOFs, we expect that many of the advantages of the development methodology for coordinatively-saturated MOFs (such as the component-by-component based parameterization and protocol for partial charge determination) should also translate well to CUS-MOF materials. Nevertheless, there are two reasons why a SAPT-based methodology is non-ideal for generating CUS-MOF force fields. First, and as shown in Fig. 5.1 for a representative Mg-MOF-74 cluster model, by comparing to benchmark CCSD(T)-f12 calculations we have discovered SAPT to be in error for CUS-MOF-like systems. DFT-SAPT is known to struggle with highly ionic systems (relative to CCSD(T) or DFT methods),[20,21] and so this error is perhaps not surprising. (Possible sources of the discrepancy between SAPT and CCSD(T)-f12 will be discussed in Section 5.8.) 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 $CO_2$ by a full 6 kJ/mol compared to CCSD(T)-f12, we would not expect to see good predictions for the $CO_2$ adsorption isotherm with a SAPT-based methodology. For CUS-MOFs and other similar systems, a new strategy for force field development is 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

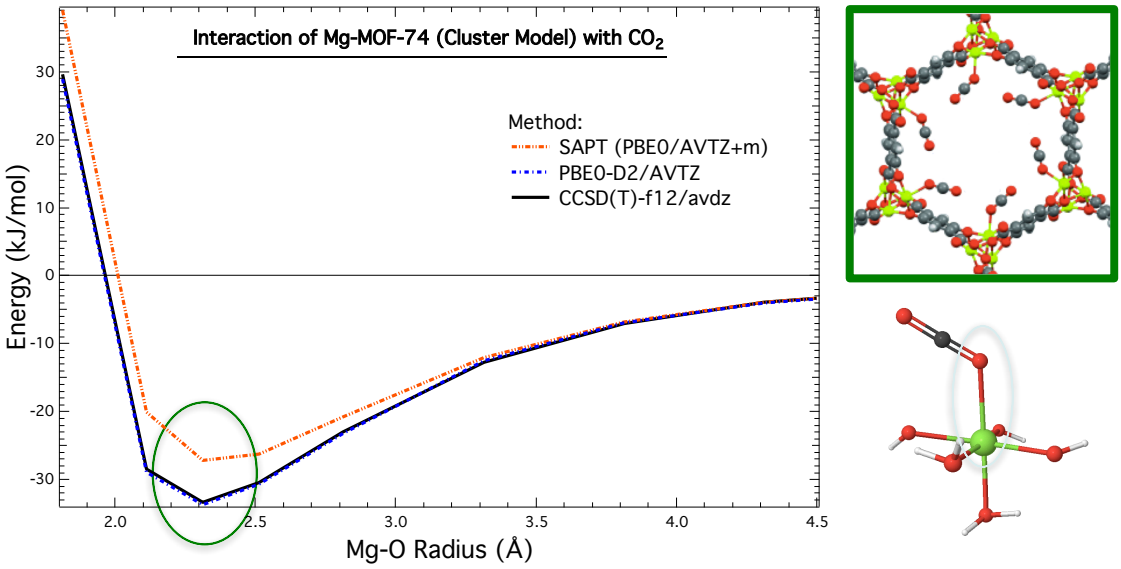



Figure 5.1: Model PES for interactions between $CO_2$ and Mg-MOF-74. (Left) Interaction energies between $CO_2$ and a cluster model of Mg-MOF-74 (shown bottom right), computed at a CCSD(T)-f12 (black), SAPT (orange), and/or PBE0-D2 (blue) level of theory. Discrepancies between SAPT and CCSD(T)-f12 in 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.

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.[22] For these reasons, a new, open-shell-compatible electronic structure benchmark is highly preferable.

## 5.4 Parameterizing CUS-MOF force fields with LMO-EDA

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 development strategy (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 (EDA) with the following qualities:

1. High accuracy with respect to to CCSD(T)-f12 benchmark energies
2. Physically-meaningful energy decomposition into (at least) electrostatics, exchange, induction, and dispersion
3. Quantitative correspondence between the energy decompositions of SAPT and the new method for systems where total energies from SAPT, CCSD(T)-f12, and the new method agree

Assuming these three qualites 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.

A substantial number of EDAs exist in the literature, and the interested reader is referred to Ref. 21 for a review and comparison of various popular methods. Aside from SAPT, which is a perturbative method, most EDAs are 'variational', meaning that the various energy components are calculated in stages from a series of constrained relaxations of the monomer wavefunctions into the optimized dimer wavefunction. For this reason, all variational EDAs are guaranteed to have total energies that match the result from a supermolecular interaction energy calculation. Furthermore, these EDAs are often implemented for wavefunction and DFT methods, thus allowing for significant flexibility (compared to the SAPT EDA) in terms of finding an EDA whose total energy closely matches CCSD(T)-f12. Indeed,

and as shown in Fig. 5.1, PBE0-D2 shows excellent agreement with CCSD(T)-f12 for a Mg-MOF-74 cluster model, and so any DFT-compatible EDA should meet our first criteria from above.

Although all variational EDAs yield the same total interaction energy for a given level of theory, many EDAs can differ substantially in terms of how this total energy is decomposed into chemically-meaningful components. At the time this research was completed, only a handful of variational EDAs distinguished each electrostatics, exchange, induction, and dispersion. (Notably, the recent second-generation ALMO-EDA[23] now separates their 'frozen' energy term into electrostatic, exchange, and dispersion components, and thus might be worth future investigation.) Of the popular EDA methods available in 2014, we found that LMO-EDA,[24,25] GKS-EDA,[26] and PIEDA[27] decompose the total interaction in a manner philosophically similar to SAPT, and include each electrostatic, exchange, induction, and dispersion terms. These three methods thus meet our second criteria for an optimal energy decomposition scheme for CUS-MOFs, and complete formalisms and details for the methods can be found in Refs. 24–27.

As for the last criterion, that of maximum correspondence between SAPT and a variational EDA, we have performed component-by-component analyses to compare SAPT to both LMO-EDA and GKS-EDA. PIEDA is known to overestimate the relative magnitude of the polarization energy, compared to SAPT, and thus was not considered in detail.[21] As for LMO-EDA and GKS-EDA (both of which are based on very similar theories, and tend to yield similar energy decompositions), we have in general found semi-quantitative to quantitative agreement with the SAPT energy decomposition, particularly for the electrostatic and exchange energies. Comparisons between LMO-EDA and SAPT are shown for the $CO_2$ dimer (Fig. 5.2) and for $CO_2$ interacting with a model Mg-MOF-74 compound (Fig. 5.3). GKS-EDA results are not shown, as the LMO-EDA and GKS-EDA results tend to be very similar, with the GKS-EDA results in slightly worse agreement with SAPT. For this reason, and because LMO-EDA does the best job of meeting our three criteria above, we choose in this work to use LMO-EDA as our new benchmark EDA for fitting CUS-MOF force fields.

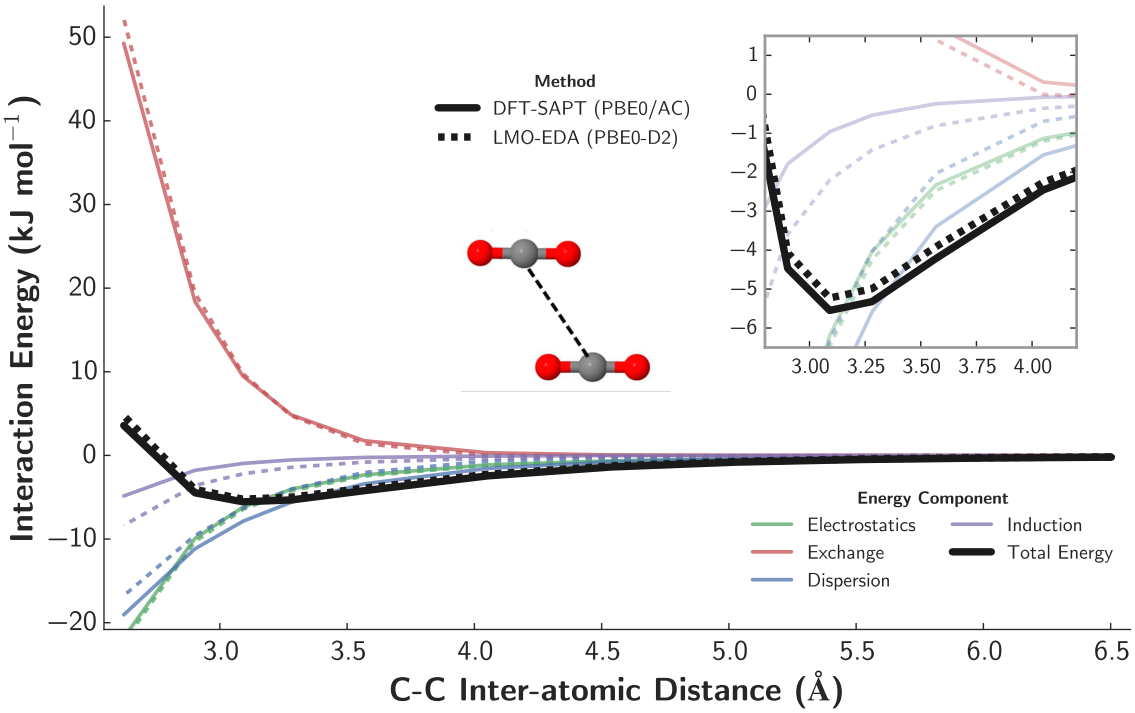



Figure 5.2: PES and associated energy decomposition for the slipped parallel geometry of the $CO_2$ dimer as a function of the C–C interatomic distance. The PES has been computed by both DFT-SAPT (PBE0) (solid lines) and LMO-EDA-PBE0-D2 (dashed lines), and each electrostatics (green), exchange (red), dispersion (blue), induction (purple), and total energy (black) components are displayed. Note that, for the DFT-SAPT energies, the δHF contribution has been incorporated into the induction energy.

In addition to describing the advantages of the LMO-EDA method, it is worthwhile to overview some of its relevant shortcomings and limitations. As with most variational EDA methods,[21] and especially for DFT-based methods, it becomes difficult to precisely assign and separate out the true 'dispersion' energy for a system. This limitation is also true of LMO-EDA, where the dispersion energy is defined as the difference in correlation energy between the monomer and dimer wavefunctions. (For density functionals employing Grimme's –D dispersion correction, this correction is also added to the LMO-EDA dispersion energy.) For functionals that

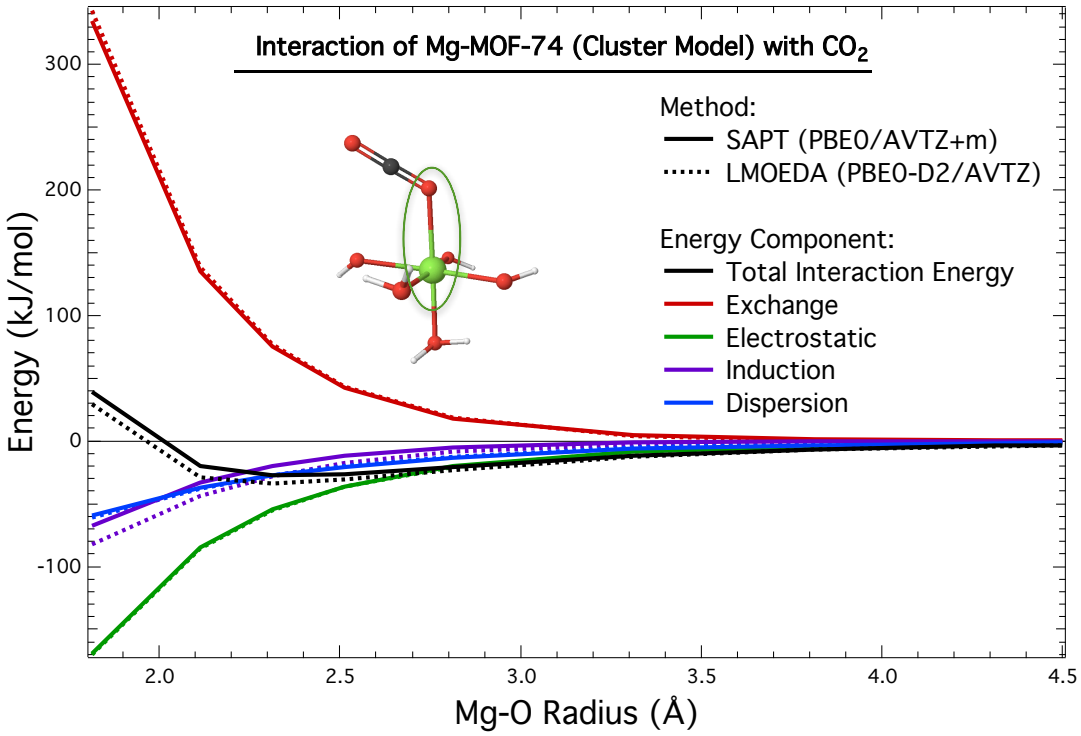



Figure 5.3: PES and associated energy decomposition for a $CO_2$ + $MgO_5$ cluster, as a function of the highlighted Mg–O interatomic distance. The PES has been computed by both DFT-SAPT (PBE0) (solid lines) and LMO-EDA-PBE0-D2 (dashed lines), and colors and labels for the energy decomposition are as in Fig. 5.2.

have a well-defined and theoretically-grounded distinction between the exchange and correlation functionals, the LMO-EDA energies tend to agree well with SAPT, and we have found good agreement (for instance) between SAPT and LMO-EDA-PBE0-D2. With other functionals, such as with our tests using the M06 functional, there is no separation between the exchange and correlation functionals, and LMO-EDA gives unphysical values for both the exchange and dispersion energies in this case. (Notably, GKS-EDA attempts to rectify this issue by changing the LMO-EDA formalism for dispersion. While this leads to qualitative agreement between SAPT and GKS-EDA for a wider variety of functionals, the quantitative agreement for the PBE0-D2 functional is somewhat worsened for the systems studied herein, and

we instead use LMO-EDA-PBE0-D2 for all results in this work.)

A second, and purely practical, limitation of LMO-EDA is its memory-intensive implementation in GAMESS. As will be discussed in detail later, calculations on a large (43 heavy atom) cluster model of Mg-MOF-74 were infeasible for us (using the Phoenix cluster in 2014) in all but the smallest VDZ basis set, and calculations on an identical cluster model of Co-MOF-74 could not be completed at all. For this reason, the LMO-EDA method is practically restricted to studies of smaller systems and/or basis sets.

## 5.5 Computational Methods

### 5.5.1 Partial Charge Determination

Partial charges for Mg-MOF-74 were determined in a manner analagous to Ref. 19 using the $Q_{SBU}$ method. Two cluster models, one a hydrogen-capped DOBDC ligand environment, and one a capped $MgO_5$ inorganic chain, were constructed and analysed using a Distributed Multipole Analysis (DMA). The resulting DMA charges were then used to obtain charge paramters for the ligand and inorganic SBU, respectively. See Section 5.A for final charge parameters.

### 5.5.2 Force Field Fitting

Two types of force field functional forms were considered in this work. The first, a 'single-exponential' functional form, exactly matches that used in Ref. 28, with the exception that δHF parameters were not fit to the Mg atomtype. This fitting choice was due to the fact that LMO-EDA only provides a total induction term (rather than splitting into $2^{nd}$- and higher-order induction energies, as with SAPT).

For the 'double-exponential' functional form used to fit the Mg-MOF-74-Yu cluster model, the same functional form was used as in the single-exponential case, with the exception that two sets of short-range interaction parameters (labeled Mg and Du in Section 5.A) were assigned to the Mg atomic center. This effectively

meant that Mg was described by two separate exponential decays, thus enabling additional parameterization flexibility for the force fields discussed in Section 5.6.

In all cases, force fields were fit using the Fortran code described in the Appendix of Ref. 29.

## 5.6 Results

### 5.6.1 Initial Force Field and Cluster Model Analysis

Originally, we attempted to fit Mg parameters on the basis of a small, 6 heavy atom cluster ('Mg-MOF-74-small', see Fig. 5.4 for chemical structure), which we felt would be representative of the Mg environment in Mg-MOF-74. Using the functional forms discussed in Section 5.5, force field parameters were fit to reproduce LMO-EDA-PBE0-D2 energies for a variety of $CO_2$/Mg-MOF-74-small interactions, with results shown in Fig. 5.4. Though select interaction energies disagree by several kJ/mol between LMO-EDA-PBE0-D2 and the force field energies, overall the agreement is reasonable, and the force field correctly reproduces trends in the interaction energies without significant systematic error.

Based on the agreement between PBE0-D2 and the force field, as well as between PBE0-D2 and CCSD(T)-f12, we expected to obtain good $CO_2$ adsorption isotherm predictions for the Mg-MOF-74 system itself. By contrast, our computed isotherm substantially underpredicts the experimental adsorption at low pressures, where Mg–$CO_2$ interactions are known to dominate. This underprediction strongly suggests that we had originally underestimated the magnitude of the Mg–$CO_2$ binding, a result which we were then able to attribute to our choice of cluster model (vida infra).

Cluster models for the M-MOF-74 series have been investigated by several groups, and it has been found in general that computed binding energies are sensitive both to the size of the cluster model as well as the treatment of geometry relaxation effects.[31,32] Consequently, we calculated the $CO_2$ binding energies and geometries of both our original Mg-MOF-74-small cluster as well as for two larger

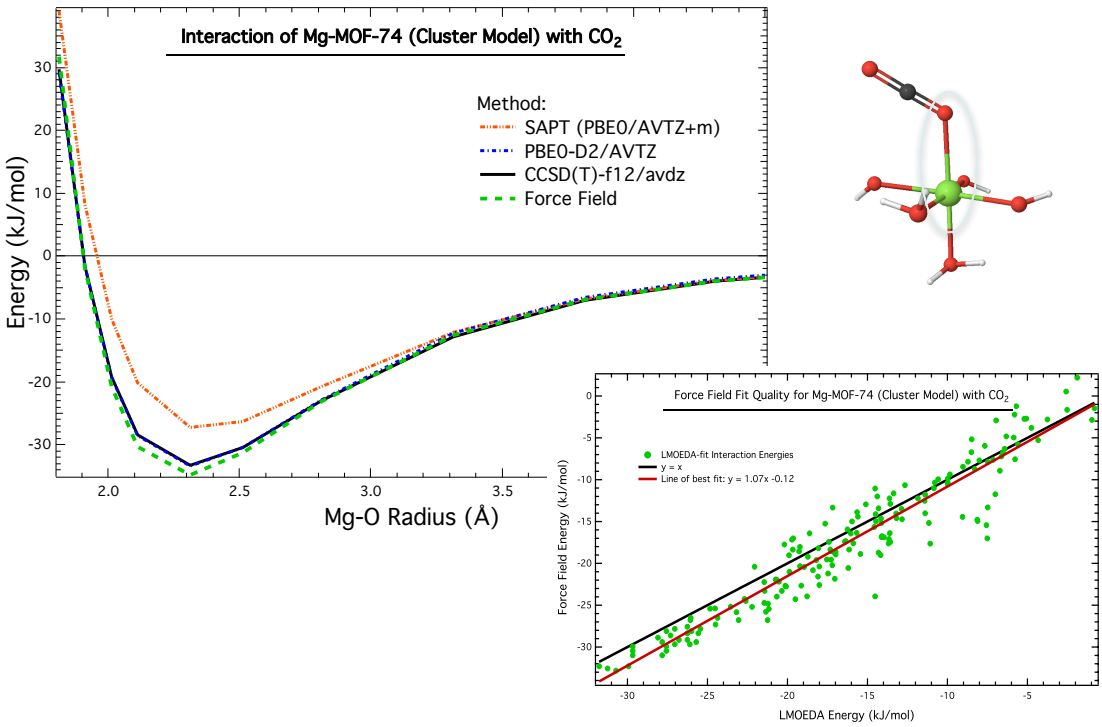
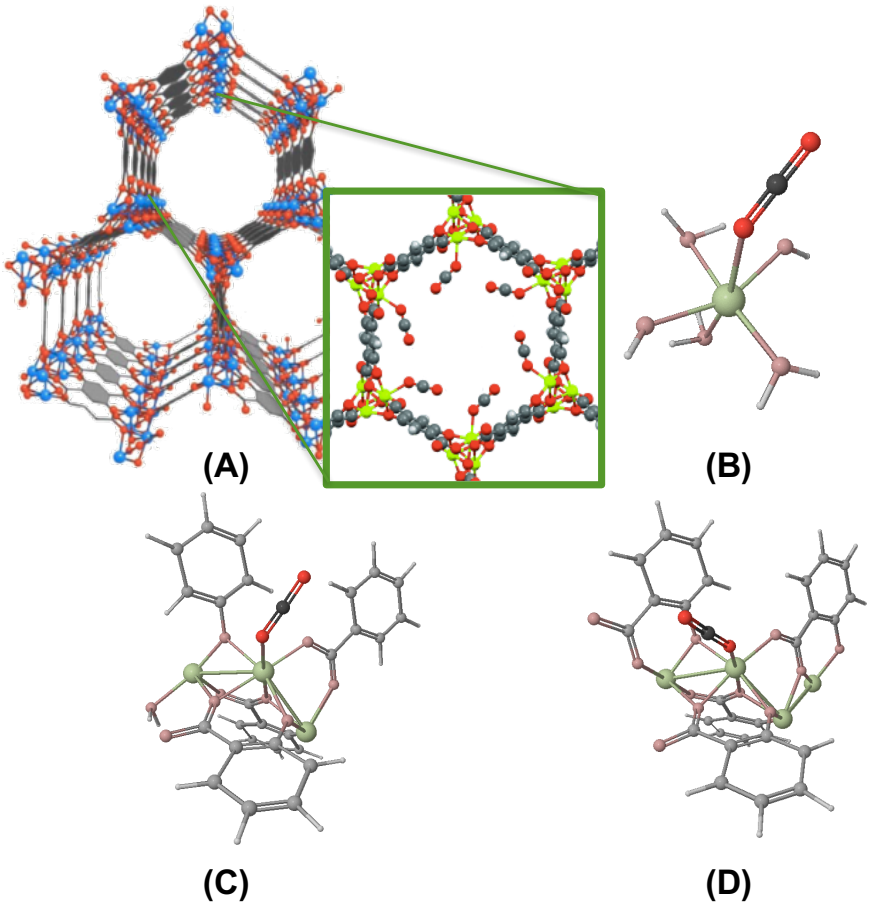



Figure 5.4: Force field fitting quality for the Mg-MOF-74-small cluster. (Top left) Various electronic structure benchmarks for Mg-MOF-74-small along with the classical potential. Each DFT-SAPT (orange dot-dashed), PBE0-D2 (blue dot-dashed), CCSD(T)-f12 (black solid), and the LMO-EDA-based force field (green dashed) are shown as a function of the Mg–O interatomic distance (non-bonding pair highlighted at top right). (Bottom right) Force field fit quality, as benchmarked against LMO-EDA-PBE0-D2, for a semi-random set of dimer configurations of the Mg-MOF-74-small cluster model interacting with $CO_2$. The black line establishes the $y = x$ benchmark, and the red line represents the line of best fit.




| Model | $CO_2$ Binding Energy (kJ/mol) | Mg–O Interatomic Distance (Å) | Mg–O–C Tilt Angle (°) |
|---|---|---|---|
| **A**[11] | **-41.5** | **2.31** | **129** |
| B | -23.3 | 2.31 | 122 |
| C | -31.4 | 2.28 | 123 |
| D | -41.7 | 2.20 | 149 |

Figure 5.5: Various structures and cluster models for Mg-MOF-74 interacting with $CO_2$. (A) Full periodic Mg-MOF-74 structure with inset showing adsorbed $CO_2$ positions. (B) Mg-MOF-74-small cluster, containing 6 heavy atoms (not including $CO_2$). (C) Yu et al. cluster model for Mg-MOF-74, denoted in text as Mg-MOF-74-Yu. (D) Dzubak et al. cluster model for Mg-MOF-74, denoted in text as Mg-MOF-74-Dzubak. All cluster models as shown with optimized $CO_2$ positions, and bond lengths and angles for adsorbed $CO_2$ are given in the bottom table. Data for (A) was taken from Valenzano et al. using a B3LYP-D level of theory,[11] whereas data for (B-D) was computed in this work using PBE0-D2. Finally, note that the binding energy for (A) includes framework geometry relaxation effects, whereas (B-D) were computed using semi-rigid cluster geometries and only optimizing the $CO_2$ position and exposed $MgO_5$ pocket.

clusters developed in Refs. 6, 30. These latter two clusters, respectively denoted Mg-MOF-74-Yu and Mg-MOF-74-Dzubak, are the same size (each with 60 atoms), but have distinct stoichiometries and geometries. To test the influence of model cluster on the $CO_2$ binding energy/geometry, we performed two sets of optimizations of the Mg-MOF-74-Yu and Mg-MOF-74-Dzubak clusters: one in which only the $CO_2$ position was optimized, and one in which the exposed $MgO_5$ pocket was additionally relaxed. Binding geometries were relatively insensitive to the geometry relaxation, though binding energies varied by 2-5 kJ/mol, in agreement with other studies that have tested geometry relaxation effects.[31] Results for the $CO_2 + MgO_5$ relaxation are shown in Fig. 5.5.

Of the three studied cluster models, both Mg-MOF-74-small and Mg-MOF-74-Yu correctly reproduce the Mg–O interatomic distance and Mg–O–C tilt angle. These geometrical parameters arise primarily from electrostatic interactions between $CO_2$ and the $MgO_5$ pocket,[11] suggesting that both of these models capture such important interaction features. By contrast, the Mg-MOF-74-Dzubak model predicts a substantially shorter binding distance and increased tilt angle, both in contrast to results from the periodic system. In part, these deficiencies can be attributed to spurious $CO_2$ interactions with the exposed carbonyl capping groups in the Mg-MOF-74-Dzubak model, as these exposed carbonyls are not present in the periodic system or the other two cluster models. As a second distinction, a Mulliken charge analysis of the Mg-MOF-74-Dzubak cluster yields larger partial charges for the surrounding Mg atoms as compared to the Mg-MOF-74-Yu model, which may help explain the increased binding and shortened Mg–O contact in the Mg-MOF-74-Dzubak model.

There are also substantial differences in binding energies between the various cluster models. Importantly, Mg-MOF-74-small severely underbinds $CO_2$ compared to all other tested systems. These results for the Mg-MOF-74-small cluster indicate the inadequacy of such a small model, and likely explain the underprediction of the $CO_2$ adsorption isotherm from above. The Mg-MOF-74-Dzubak model shows best energetic agreement with the periodic system. Nevertheless, some of the Mg-MOF-74-Dzubak binding energy arises from truncation effects (as described above), and

the energetic agreement is thus due (at least in part) to error cancellation. Indeed, some of the binding energy in the periodic system arises from (attractive) long-range interactions, and thus we should expect to see a cluster model somewhat underpredict the binding energy. Primarily for its good agreement in binding geometries, and reasonable agreement in binding energy, we opt to use the Mg-MOF-74-Yu cluster model for the remainder of this work.

### 5.6.2 Final Mg-MOF-74 $CO_2$ Adsorption Isotherm

Using our new Mg-MOF-74-Yu cluster model, we next attempted to refit force field paramters for Mg. As discussed earlier, and because of the size of this new cluster (60 atoms), LMO-EDA-PBE0-D2 calculations became cost prohibitive in all but the smallest VDZ basis set, and thus could only be carried out for a limited set of points. Starting from the minimum energy configuration shown in Fig. 5.5, we fit Mg parameters to a 12-point scan along the Mg-O bond vector, with fit results shown in Fig. 5.6. Interestingly, though the functional form used in this fit was sufficient to accurately paramterize the interaction energies in the Mg-MOF-74-small cluster, the same force field methodology proved unsuccessful in paramterizing Mg-MOF-74-Yu–$CO_2$ interactions. We knew at the time that this inaccuracy was probably a consequence of uncertainties in correctly parameterizing the Mg short-range exponent. (See Chapter 3 for a full discussion of new methods for parameterizing the short-range potential.) Nevertheless, because the Slater-ISA methodology for short-range interactions had not yet been developed, we opted instead to fit the Mg interactions to a double exponential functional form, with each exponent corresponding to the ionization potential for either $Mg^+$ or $Mg^{2+}$ (the two atomic environments most likely to correctly represent the Mg cation). As shown in Fig. 5.6, this form could excellently reproduce the Mg-MOF-74-Yu model PES.

Using the double exponential functional form from above, we recomputed the adsorption isotherm of $CO_2$ in Mg-MOF-74. Before comparing to experiment, and as recommended by others,[33] we scaled the experimental isotherm in order to account for the pore blocking effects that are common in the M-MOF-74 series.

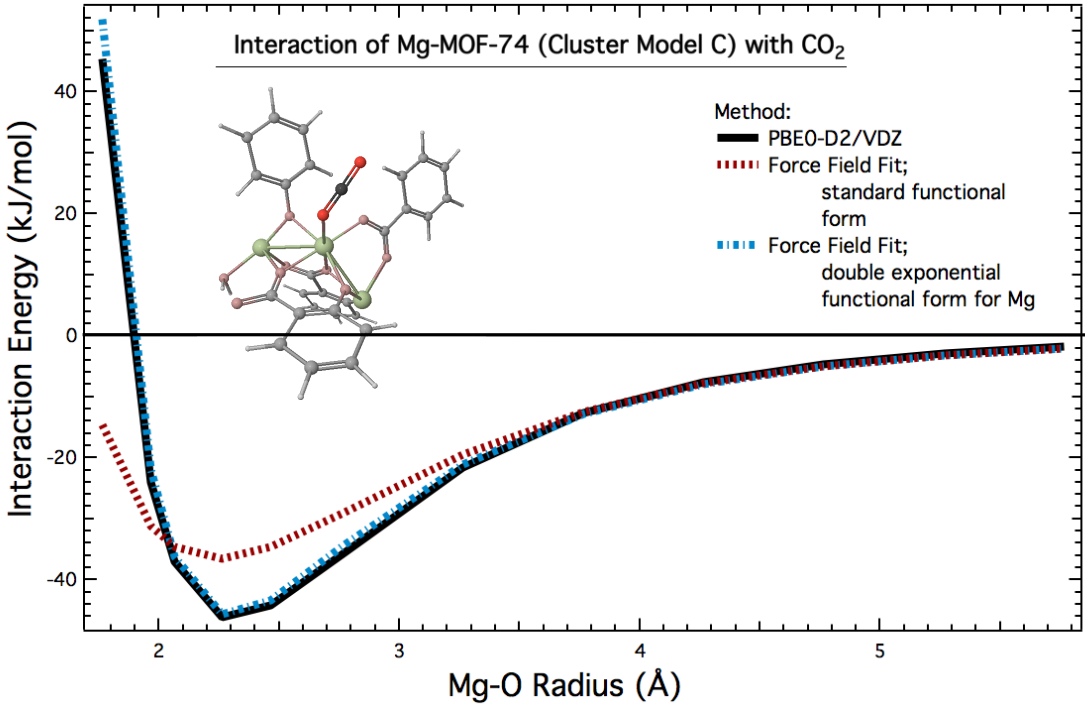



Figure 5.6: Force field fitting quality for the Mg-MOF-74-Yu cluster. A PBE0-D2 benchmark (black solid) is displayed along with two force field fits: single-exponential (red dashed) and double-exponential for Mg (blue dash-dotted). In either case, a cut of the PES is shown along the interatomic distance between the central Mg atom and the closest-contact oxygen atom in $CO_2$.

Using this scaled isotherm, we then obtain excellent agreement between our model potential and experiment (Fig. 5.7). Crucially, this accuracy is seen both at low- and high-pressure ranges, indicating the accuracy of the force field in modeling both the strong Mg–$CO_2$ binding as well as the weaker physisorption regime.

### 5.6.3 Transferability to Other Adsorption Isotherms

In addition to using our Mg paramters to compute the $CO_2$ adsorption isotherm, we also used our Mg force field in conjunction with the $N_2$ parameters developed by Yu et al.[34] to predict the $N_2$ adsorption isotherm. These predictions were generally

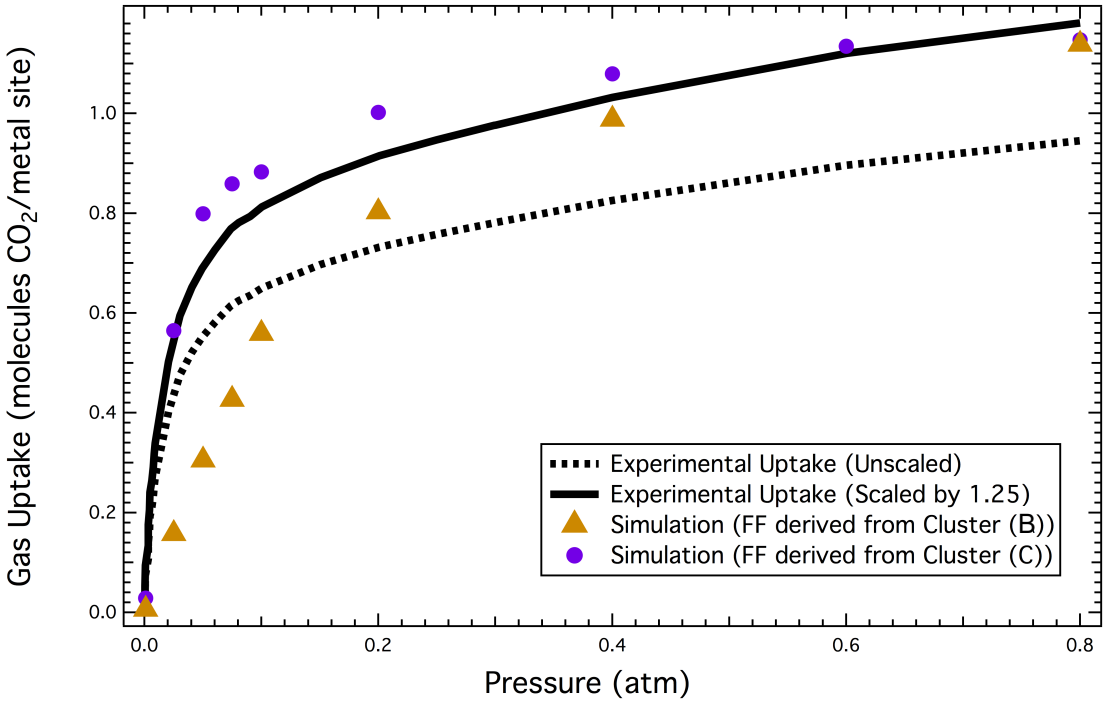



Figure 5.7: Predicted $CO_2$ adsorption isotherm for Mg-MOF-74. Two experimental isotherms are shown, one as directly measured by experiment (dashed line) and one scaled to account for pore block effects (solid line). Predictions from two force fields are also shown, where Mg parameters for each force field were fit either to the Mg-MOF-74-small cluster (gold triangles) or to the Mg-MOF-74-Yu cluster (purple circles). Cluster geometries are given in **??**.

poor, and results are not shown. Nevertheless, the poor $N_2$ results suggest a lack of transferability of our Mg parameters, possibly (and as discussed in the section on Future Work) due to the unphysical double-exponential functional form used to parameterize Mg.

### 5.6.4 Transferability to Other M-MOF-74 systems

As a second test of transferability, we also attempted to develop force fields for other compounds in the M-MOF-74 series, starting with Co-MOF-74. Unfortunately, the open-shell nature and increased electron count of Co-MOF-74 made LMO-EDA

calculations computationally prohibitive for any reasonable basis set, and these systems were not investigated further.

## 5.7 Conclusions

In this unpublished work, we have determined a new methodology for fitting force fields to CUS-MOFs. While largely following our previous methodology for MOF force field development, we have shown how an LMO-EDA Energy Decomposition Analysis can be used in leiu of SAPT to generate accurate ab initio benchmark energies in cases where SAPT itself is inaccurate. Using this new methodology, we have successfully modeled $CO_2$ interactions in Mg-MOF-74, and have simulated the adsorption isotherm for $CO_2$ in Mg-MOF-74 with good accuracy compared to experiment.

Ultimately, the methods presented herein suffered from a number of practical and fundamental issues (vida infra). Until these limitations can be fully addressed, we do not anticipate that the LMO-EDA method can be broadly used to develop transferable force fields for CUS-MOFs or other large systems where SAPT is in error.

## 5.8 Future Work

Throughout this Chapter, we have attempted to highlight some of the key limitations of our force field development methodology for CUS-MOFs. In summary, the following issues would need to be resolved in order to expand the scope and utility of the present research:

1. **Memory Limitations in GAMESS**: As evidenced in this work, relatively large (60+ atom) cluster models are required to correctly parameterize force fields for the M-MOF-74 system. While these cluster sizes do not present difficulties for standard DFT calculations with reasonable basis sets, the corresponding LMO-EDA calculations were, as implemented in the GAMESS software pack-

age, infeasible due to memory requirements. Some time was spent attempting to address these memory issues, particularly for the memory-intensive Edmiston-Ruedenberg localization subroutine that is the source of the problem. However, due to our lack of familiarity with the GAMESS software and the LMO-EDA source code, this pursuit was eventually dropped.

2. **Fundamental Issues with LMO-EDA**: As discussed in Section 5.4, the LMO-EDA method has several theoretical limitations. In particular, and especially for functionals with no defined separation between exchange and correlation functionals, LMO-EDA does not offer a clean separation between the exchange and dispersion energies. Furthermore, and unlike some recent EDA methods,[23?] LMO-EDA cannot separate induction into charge transfer and polarization components.

3. **Transferability of the Force Field Functional Form**: While our final force field for studying $CO_2$ interactions in Mg-MOF-74 is highly accurate (both with respect to ab initio theory and with respect to experiment), it does not appear that this accuracy extends to models for the adsorption of other small molecules, such as $N_2$. This transferability limitation is almost certainly due to the chosen double-exponential functional form and/or the parameterization process used to obtain Mg parameters, and improvements to this methodology will be essential to make our work on the $CO_2$–Mg-MOF-74 system applicable to general force field development for CUS-MOFs. In particular, future work will require a better force field for describing short-range interactions, as the functional forms and paramters used in this work struggled to both accurately and transferably model the Mg-MOF-74 exchange energies.

While several of these issues (particularly practical limitations with the LMO-EDA implementation) have yet to be addressed in a meaningful fashion, several recent theoretical advances may pave the way for continued work on this project. Thus for CUS-MOFs and other systems where DFT-SAPT might be in error, we offer the following recommendations:

1. **Improved SAPT energies**: Recently, it has been proposed that the commonly used single-exchange ('$S^2$') approximation can lead to errors in the description of the induction energy, particularly for ionic systems.[20,35] While it is difficult to attribute errors in SAPT to a particular energy component, it may well be that SAPT poorly describes Mg-MOF-74 due to the $S^2$ treatment of the induction energy. In this case, eliminating the $S^2$ approximation might improve the DFT-SAPT total interaction energies, thus enabling SAPT to be used for (at least) closed-shell CUS-MOFs.

2. **New SAPT Correction Schemes**: As discussed in Chapter 4, deviations between SAPT and CCSD(T) can be rectified by adding a δCCSD(T) correction to the total SAPT energy. As discussed in Chapter 4, we have empirically had good success modeling this δCCSD(T) correction as part of the dispersion energy. Though this partitioning choice may require adjustment for treating Mg-MOF-74, the results in Chapter 4 indicate that simply correcting (rather than entirely ignoring) the DFT-SAPT energies is a promising strategy for transferable force field development.

3. **New EDA schemes**: Since this work was completed, a second-generation ALMO-EDA scheme has been implemented in the Q-Chem software package.[23] Crucially, and unlike its predecessor, this ALMO-EDA scheme now breaks up the interaction energy into electrostatic, exchange, polarization, charge-transfer, and dispersion components. While there is no guarantee that such an EDA could serve as the basis for CUS-MOF force field development (see Section 5.4), these and other recently developed EDAs may be worth investigation, and could eventually replace the (practically problematic) LMO-EDA method.

4. **Improved Force Field Functional Forms**: Since 2014, we have made significant progress in developing more accurate and transferable intermolecular force fields (see Chapters 3 and 4), and many of these advances particularly improve the description of the short-range potential itself. There is a good

chance that either the Slater-ISA FF or MASTIFF methodologies would yield high-quality force fields for Mg-MOF-74 and other systems. In this case, continued work on this project might be an exciting avenue for showcasing the MASTIFF methodology in the context of accurate inorganic/organometallic force field development.

## 5.A Force Field Parameters for $CO_2$ and Mg-MOF-74

Final force field parameters, fit using the double-exponential functional form above and the Mg-MOF-74-Yu cluster model, for $CO_2$ and Mg-MOF-74. These parameters should be read in as input into our group's lattice simulation code, see `http://schmidt.chem.wisc.edu/montecarlosimulationcodes` for details.

Listing 5.1: co2_mof74.pmt

```
lennard_jones_type      1       "1 for buckingham, 2 for lennard jones"

" parameters are listed as charge, A,B,C, polarizability
" units are A:kJ/mol,  B:A^-1,  C: KJ/mol*A^6   epsilon:KJ/mol,   sigma:A,
    polarizability: A^3 "
" polarizability is defined as q^2/k, spring constant is set to .1*1.8897^3 e^2/A^3"


solute_species
atom_type_parameters  ( q, Aexch, Aelec, Aind, Adhf, Adisp, C6, C8, C10, alpha )
2
C0      0.6573800       95510.43        -27846.98       -13425.1        2065.044
        0.0     6.891E02        0.0     0.0     1.1926153
O0      -0.328690       521902.066      -163908.84      -4475.8095      -26042.04
        0.0     1.8341E03       0.0     0.0     0.9009290

solute parameters for framework cross terms  ( Aexch, Aelec, Aind, Adhf, Adisp, B, C6, C8
    , C10, C12 )
C0      74376.65        24130.18     12513.37           795.36
        0.0     3.4384       1147.41867     6329.41038     29659.50100     183546.714
O0      354373.13       108208.5      2544.89          -18178.7
        0.0     3.7795          867.27598     4266.54582     28761.10636   132581.301

solute dhf cross terms (check code for input format if more than one cross term)
-6124.0
```

```
solute-solute exponents ( Bii , Bij , Bjj)
3.5105206        3.6993494        3.9288490


framework_species
atom_type_parameters    ( q, Aexch, Aelec, Aind, Adhf, Adisp, B, C6, C8, C10, C12, alpha )
9
C1   -0.1639350   612892.611   229850.679    -6511.529   -60036.930
         0.000         3.438     1628.820002     6821.530007     44464.989999    193602.980000
                 0.0
H1    0.2637500     8538.651     1678.771      -612.739      -502.639
         0.000         3.778      129.439978      679.640001     4995.299998         0.000000
                 0.0
C2   -0.3191850   612892.611   229850.679    -6511.529   -60036.930
         0.000         3.438     1628.820002     6821.530007     44464.989999    193602.980000
                 0.0
C3    0.4964850   612892.611   229850.679    -6511.529   -60036.930
         0.000         3.438     1628.820002     6821.530007     44464.989999    193602.980000
                 0.0
O3   -1.0339500     3398.424     1965.168      -182.412      -178.025
         0.000         2.457     2237.635879    29956.090890    561056.184030   7451461.601923
                 0.0
C4    0.9468300   263600.161   112896.479      -11.681     -2837.170
         0.000         3.438      772.870024     2349.180008     27539.189998    102366.260000
                 0.0
O4   -0.8903225   656757.170   174054.351    -45410.640    -33954.271
         0.000         3.779     1799.560008    11576.089993     50164.639999         0.000000
                 0.0
Mg    1.5906500      917.037     2417.463    -12542.799         0.000
         0.000         2.834      630.723467         0.000000         0.000000         0.000000
                 0.0
Du    0.0000000    29176.333        0.000   142260.481         0.000
         0.000         3.973         0.000000         0.000000         0.000000         0.000000
                 0.0
```

## 5.B Simulation Parameters $CO_2$ Adsorption in Mg-MOF-74

Lattice simulation parameters for $CO_2$ adsorption in Mg-MOF-74. Of particular importance is the 'orientation_try' keyword, which is necessary to sample the specific binding geometries $CO_2$ adopts when binding to the open-metal site. These simulation parameters should be read in as input into our group's lattice simulation code, see `http://schmidt.chem.wisc.edu/montecarlosimulationcodes` for details.

Listing 5.2: simulation_parameters.pmt

```
Simulation Methodology
energy_decomposition            yes        ! yes for our force fields, no for UFF LJ, etc
    .
solute_cross_parameter_set      yes        ! this should be set to yes if using different
     solute parameters
           ! for solute-solute and solute-framework interactions as in our force fields,
                no otherwise
C8_10_dispersion_terms          yes        ! set to yes if using C8, C10 dispersion terms
     as in our force fields
C12_dispersion                  yes
electrostatic_type              pme        ! either "pme" for particle-mesh ewald, "
    cutoff", or "none"
lj_comb_rule                    ZIFFF      ! "opls" or "standard" for lj, "standard" or "
    ZIFFF" for bkghm

Simulation Parameters
temperature                     296.0           ! temperature in Kelvin
too_close                       1.8             ! reject move if molecules are within
    this separation in Angstroms.
           ! helpful to avoid unnecessary energy calculations and to prevent drude
               oscillator catastrophes
lj_bkghm                        1               ! 1 for bkghm force field, 2 for lj
screen_type                     1               ! screening for coulomb potential: 0 = no
     screening, 1 = Tang-Toennies type screening for our force fields
springconstant                  0.1             ! spring constant for drude oscillators (
    au). set to 0.1 for our CO2/N2 force fields
thole                           2.0             ! thole parameter for intra-molecular
    drude oscillator screening.  Set to 2.0 for our CO2/N2 force fields.
drude_simulation                1               ! set to 1 if drude-oscillators are being
     used, 0 otherwise
pme_grid                        100             ! size of the pme grid
alpha_sqrt                      0.6             ! alpha sqrt for the electrostatic
```

```
    interactions
lj_asqrt                    0.6             ! alpha sqrt for the pme dispersion
lj_cutoff                   7.5             ! cutoff for long range LJ or C6,C8,C10
    dispersion interactions
ewald_cutoff                5.0             ! cutoff for real space pme
cav_grid_a                  30
cav_grid_b                  30
cav_grid_c                  30
na_nslist                   30              ! neighbour list searching grid
nb_nslist                   30              ! neighbour list searching grid
nc_nslist                   30              ! neighbour list searching grid
orientation_try             2000            ! max number of orientation samplings
REL_THRSH                   0.05            ! sampling threshold
ABS_THRSH                   3.0
BZ_CUTOFF                   100.0
```

# Part IV

# Practical Matters

# Part V

# Conclusions and Future Work

# Part VI

# Codes

## 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.; AreaÌ'n, C. O.; Bordiga, S. *J. Phys. Chem. C* **2010**, *114*, 11185–11191.

[12] Poloni, R.; Lee, K.; Berger, R. F.; Smit, B. **2014**,

[13] Lin, L.-c.; Lee, K.; Gagliardi, L.; Smit, B. *J. Chem. Theory Comput.* **2014**, *10*, 1477–1488.

[14] Haldoupis, E.; Borycz, J.; Shi, H.; Vogiatzis, K. D.; Bai, P.; Queen, W. L.; Gagliardi, L.; Siepmann, J. I. *J. Phys. Chem. C* **2015**, *119*, 16058–16071.

[15] Mercado, R.; Vlaisavljevich, B.; Lin, L.-C.; Lee, K.; Lee, Y.; Mason, J. A.; Xiao, D. J.; Gonzalez, M. I.; Kapelewski, M. T.; Neaton, J. B.; Smit, B. *J. Phys. Chem. C* **2016**, acs.jpcc.6b03393.

[16] Becker, T. M.; Heinen, J.; Dubbeldam, D.; Lin, L.-C.; Vlugt, T. J. H. *J. Phys. Chem. C* **2017**, acs.jpcc.6b12052.

[17] McDaniel, J. G.; Schmidt, J. R. *J. Phys. Chem. C* **2012**, *116*, 14031–14039.

[18] McDaniel, J. G.; Yu, K.; Schmidt, J. R. *J. Phys. Chem. C* **2012**, *116*, 1892–1903.

[19] McDaniel, J. G.; Li, S.; Tylianakis, E.; Snurr, R. Q.; Schmidt, J. R. *J. Phys. Chem. C* **2015**, *119*, 3143–3152.

[20] Lao, K. U.; Schaeffer, R.; Jansen, G.; Herbert, J. M. *J. Chem. Theory Comput.* **2015**, 150417132228001.

[21] Pastorczak, E.; Corminboeuf, C. *J. Chem. Phys.* **2017**, *146*, 120901.

[22] Żuchowski, P. *Chem. Phys. Lett.* **2008**, *450*, 203–209.

[23] Horn, P. R.; Head-gordon, M. *Phys. Chem. Chem. Phys.* **2016**, *18*, 23067–23079.

[24] Su, P.; Li, H. *J. Chem. Phys.* **2009**, *131*, 014102.

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

[26] Su, P.; Jiang, Z.; Chen, Z.; Wu, W. *J. Phys. Chem. A* **2014**, *118*, 2531–42.

[27] Fedorov, D. G.; Kitaura, K. **2006**,

[28] McDaniel, J. G.; Schmidt, J. R. *J. Phys. Chem. A* **2013**, *117*, 2053–2066.

[29] McDaniel, J. G. Development and Application of Physically-Motivated First-Principles Force Fields for Complex Chemical Systems. Ph.D. thesis, UW-Madison, 2014.

[30] Yu, K.; Kiesling, K.; Schmidt, J. R. **2012**,

[31] Verma, P.; Xu, X.; Truhlar, D. G. **2013**,

[32] Valenzano, L.; Civalleri, B.; Sillar, K.; Sauer, J. **2011**, 21777–21784.

[33] Haldoupis, E.; Borycz, J.; Shi, H.; Vogiatzis, K. D.; Bai, P.; Queen, W. L.; Gagliardi, L.; Siepmann, J. I. *J. Phys. Chem. C* **2015**, *74*, 150616135429005.

[34] Yu, K.; McDaniel, J. G.; Schmidt, J. R. *J. Phys. Chem. B* **2011**, *115*, 10054–10063.

[35] Jansen, G.; Scha, R. **2012**,

[36] McDaniel, J. G.; Schmidt, J. R. *J. Phys. Chem. B* **2014**, *118*, 8042–8053.

[37] Frenkel, D.; Smit, B. *Acad. Press*; 2002; Vol. New York,; p 638.

[38] Guibas, L. Representing rotations with quaternions. 1992; `graphics.stanford.edu/courses/cs164-09-spring/Handouts/handout12.pdf`.

[39] Shoemake, K. *Graph. Gems 3*; 1992; Chapter 6, pp 124–132.

[40] Jeziorski, B.; Moszynski, R.; Szalewicz, K. *Chem. Rev.* **1994**, *94*, 1887–1930.

[41] Szalewicz, K. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 254–272.

[42] Knizia, G.; Adler, T. B.; Werner, H.-J. *J. Chem. Phys.* **2009**, *130*, 054104.

[43] Misquitta, A. J.; Stone, A. J. CamCASP: a program for studying intermolecular interactions and for the calculation of molecular properties in distributed form, version 5.8. University of Cambridge, 2015.

[44] Stone, A. J. *J. Chem. Theory Comput.* **2005**, *1*, 1128–1132.

[45] Misquitta, A. J.; Stone, A. J.; Fazeli, F. *J. Chem. Theory Comput.* **2014**, *10*, 5405–5418.

[46] Misquitta, A. J.; Stone, A. J. *J. Chem. Phys.* **2006**, *124*, 024111.

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

[48] Cardamone, S.; Hughes, T. J.; Popelier, P. L. a. *Phys. Chem. Chem. Phys.* **2014**, *16*, 10367.

[49] Kramer, C.; Spinn, A.; Liedl, K. R. *J. Chem. Theory Comput.* **2014**, *10*, 4488–4496.

[50] Dixon, R. W.; Kollman, P. a. *J. Comput. Chem.* **1997**, *18*, 1632–1646.

[51] Chaudret, R.; Gresh, N.; Cisneros, G. A.; Scemama, A.; Piquemal, J.-p. *Can. J. Chem.* **2013**, *91*, 804–810.

[52] Unke, O. T.; Devereux, M.; Meuwly, M. *J. Chem. Phys.* **2017**, *147*, 161712.

[53] Ferenczy, G. G.; Winn, P. J.; Reynolds, C. a. *J. Phys. Chem. A* **1997**, *101*, 5446–5455.

[54] Williams, G. J.; Stone, A. J. *J. Chem. Phys.* **2003**, *119*, 4620–4628.

[55] Misquitta, A. J.; Stone, A. J. *Mol. Phys.* **2008**, *106*, 1631–1643.

---

**S**

**SAPT** Symmetry-Adapted Perturbation Theory. v, viii, 9, 11–18, 26–28, *Glossary:* SAPT

# GLOSSARY

**C | I | 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. v, 9, 12, 28

**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. 12–14, 19, 20

## I

**induction** FILL . 16, 18

**ISA** Iterated Stockholder Atoms,FILL . v

## L

**LMO-EDA** FILL . iv, viii, 9, 15–20, 23, 25–28

## S

**SAPT** Symmetry-Adapted Perturbation Theory, a perturbative treatment of intermolecular interactions which is pretty cool. v, 9