Partial Atomic Charge Methods for Simulating Porous Frameworks with a Net Charge and their Application to Gas Separations in Zeolites.

Christopher Demone

Thesis submitted to the University of Ottawa
In partial fulfillment of the requirements for the
Masters of Science degree in Chemistry

Department of Chemistry and Biomolecular Sciences
Faculty of Science University of Ottawa

Table of Contents

Abstract

Computational simulations using empirical force fields are frequently used to model guest-host interactions in porous periodic systems, where the interaction energy is broken into electrostatic and van der Waals contributions. While simulations such as these have been instrumental in progressing our understanding of neutral periodic systems, limitations in deriving partial atomic charges has largely contributed to the difficulty in modeling charged periodic frameworks. However, many nanoporous materials possess frameworks that have a net charge, which is balanced by counter-ions that intercalate through the pores. For example, virtually all zeolites used in practice contain a proportion of Al, which bestows the framework with a negative charge.

In this respect, we investigate two methods for the generation of partial atomic charges in periodic systems having a net framework charge. First, we examine the validity of generating REPEAT electrostatic potential fitted charges derived from periodic electronic structure calculations, where a constant background charge is added to neutralize the net charge on the framework without adding neutralizing counter-ions. The second method we explore is the split charge equilibration (SQE) method for very rapid charge generation. In its original formulation, the SQE model cannot be applied to systems with a net charge. In this work, we reformulate the SQE method for non-neutral systems to be treated. The new SQE model, which we call SQE_{AB}, was shown to give equivalent results to those of the original SQE model for neutral systems. For charged frameworks, the model was shown to provide partial atomic charges in good agreement with the DFT derived REPEAT method.

Taking advantage of that work, we next focus on the development of a force field for modeling CO_2, N_2, and CH_4 gas adsorption in both neutral and charged zeolites, which we call the AMP (Aluminosilicate MicroPorous) force field. Commonly, the electrostatic potential of zeolites is represented through the use of generic charges, where every atom of the same type in the framework is assigned the same atomic charge. Though this model is fast, it fails to account for structural differences between framework geometries. In this work, we have optimized a set of SQE_{AB} parameters to reproduce the DFT derived electrostatic potentials

(ESPs) of a structurally representative set of both neutral and charged zeolite frameworks. Comparing with other popular models, the SQE_{AB}-AMP charges are shown to better reproduce the QM ESP by more than 30%, on average. Gas uptakes obtained using SQE_{AB} AMP charges were found to be within 5% of those obtained using DFT derived charges. We have further optimized a set of Lennard-Jones parameters to be combined the SQE_{AB}-AMP charges that reproduce experimental uptake data in zeolites.

Table of Figures

Table of Tables

List of Abbreviations

- **ALPO** **AL**umino**P**h**O**sphate
- **AMP** **A**luminosilicate **M**icro**P**orous
- **CBU:** **C**omposite **B**uilding **U**nit
- **CCS:** **C**arbon **C**apture and **S**equestration
- **CIF** **C**rystallographic **I**nformation **F**ile
- **CNG** **C**ompressed **N**atural **G**as
- **COF** **C**ovalent **O**rganic Framework
- **COM** **C**enter **O**f **M**ass
- **CoRE** **C**omputation-**R**eady **E**xperimental
- **CSD** **C**ambridge **S**tructural **D**atabase
- **DFT** **D**ensity **F**unctional **T**heory
- **DOE** **D**epartment **O**f Energy
- **EA** **E**lectron **A**ffinity
- **EEM** **E**lectronegativity **E**qualization **M**ethod
- **ESP** **E**lectro**S**tatic **P**otential
- **FCC** **F**luid **C**atalytic Cracking
- **FTC** **F**ramework **T**ype **C**ode
- **GA** **G**enetic **A**lgorithm
- **GCMC** **G**rand **C**anonical **M**onte **C**arlo
- **GD** **G**radient **D**escent
- **GHG** **G**reen **H**ouse **G**as
- **HF** **H**artree-**F**ock
- **HoA** **H**eat **O**f **A**dsorption
- **IP** **I**onization **P**otential
- **IR** **I**nfra**R**ed
- **IZA-SC** **S**tructural **C**ommission of the **I**nternational **Z**eolite **A**ssociation
- **LHS** **L**atin **H**ypercube **S**ampling
- **LJ** **L**ennard-**J**ones
- **MAD** **M**ean **A**bsolute **D**eviation
- **MC** **M**onte **C**arlo
- **MM** **M**olecular **M**echanics
- **MOF** **M**etal **O**rganic **F**ramework
- **PAW** **P**rojector **A**ugmented **W**ave
- **PBE** **P**erdew-**B**urke-**E**rnzerhof
- **PBU** **P**rimary **B**uilding **U**nit
- **PCP** **P**orous **C**oordination **P**olymers
- **PE** **P**ercent **E**rror
- **PPN** **P**orous **P**olymer **N**etwork
- **QEq** **C**harge **E**quilibration
- **QM** **Q**uantum **M**echanical
- **REPEAT** **R**epeating **E**lectrostatic **P**otential **E**xtracted **AT**omic
- **RMSPE** **R**oot-**M**ean-**S**quared **P**ercent **E**rror
- **SBU** **S**tructural **B**uilding **U**nit
- **SCIBS** **S**ymmetry **C**onstrained **I**ntersite **B**onding **S**earch
- **SQE** **S**plit-**C**harge **E**quilibration
- **SQEAB** **S**plit-**C**harge **E**quilibration **in the** **A**tomic **B**asis
- **SUS** **S**tochastic **U**niversal **S**ampling
- **UA** **U**nited **A**tom
- **UFF** **U**niversal **F**orce **F**ield
- **VdW** **V**an **d**er **W**aal
- **XRD** **X**-**R**ay **D**iffraction

<u>Simulation Packages</u>

- **FA³PS** **F**ully **A**utomated **A**dsorption **A**nalysis in **P**orous **S**olids
- **SIESTA** **S**panish **I**nitiative for **E**lectronic **S**imulations with **T**housands of **A**toms
- **VASP** **V**ienna **A**b **I**nitio **S**oftware **P**ackage

Acknowledgements

Although my name is the one found on this thesis, the entirety of this work would not have been possible without the support of a lot of important people. It has been a long journey that has sometimes made me wonder what exactly I was doing with my life, but standing here now at the finish line I can honestly say that I wouldn't trade this experience for the world. Faced with insurmountable problems that had no easily discernable solutions, has taught me that *persistence* is sometimes more important than intelligence.

First and foremost, I'd like to begin with thanking my family who has been an important source of support and inspiration. I'd like to start by thanking my grandfather, Tony Romeo, for instilling in me a love of nature from a young age. It was at his house that I first began reading about science, from the assortment of books on biology and chemistry in his library. Being Italian, he also instilled a love for food, that goes beyond the simple exquisite tastes of classic Italian cuisine, to having a deep appreciation for the shape and symmetry of the fruits and vegetables we picked from his garden. It was from him that I became intrigued with furthering my understanding of, what seemed to me as a child, a fascinating set of underlying principles that shape the universe in which we live. I'd also like to dedicate this work in part to my grandmother, Barbara Romeo, who during her life taught me a lot about patience – which has come in handy in grad school. I would also like to thank my parents, Melina and Hardy Demone, who over the many years that they have had to endure me (day and night), have never once wavered in the support of my ever-changing dreams and aspirations. Without this support I may never have been willing to take as many chances in life as I have. In this respect I'd also like to thank my two older brothers Dathan and Stephen, who have always been an inspiration for what I do. These are two of my best friends, and I'm proud to be part of a large family, including their wives Julie and Ashley, and my nephew and nieces Jack, Raya, and

Maddalena. To the whole of my family, your love and support have been an important part of every success in my life.

I'd also like to take this opportunity to thank my fellow colleagues (wookies) and friends from the Woo lab. I learned a lot from each any every person present during the time that I spent working in the Woo lab. Firstly, I would like to thank Sean Collins, who I can honestly say I may never have survived to this point without. He taught me a lot about chemistry and computer programming, and I have greatly enjoyed our conversations about hockey (sports disk). Mykhaylo Krykunov is also someone I'd like to thank for his tireless work on developing the SQE_{AB} model, and for helping me understand a lot about the different charge models we worked with. He is a great scientist, and I am grateful for the time I had working with him. Other members of the group who have made a significant impact on me as a person include Burns, Hana, Phil, Jason, Pete, Mo, Bianca and Daff. I will mostly never forget how Phil always found a way to win at Catan. Finally, I would like to thank my research supervisor Dr. Tom Woo who has taught me the importance of communication and 'story-telling' in science. More importantly, Tom has always challenged me to do better. I'd like to express my appreciation for his tireless support throughout this entire process, and my gratitude for his patience through my many mistakes.

1 Introduction

1.1 Summary

Zeolites are three-dimensional microporous crystalline solids with well-defined structures, composed primarily of Si, O and Al. The Si and Al atoms are tetrahedrally coordinated with each other through shared oxygen atoms, to form what is sometimes referred to as a '4-connected net.' In addition to these siliceous and aluminosilicate zeolites, there can also be aluminophosphate materials in which PO_4 tetrahedra are incorporated into the frameworks, as well as the possible incorporation of trace amounts of transition metals such as Mn or Co. However, aluminosilicate materials are by far the most widely used zeolite materials for the industrial processes that give life to the modern world, which relies upon energy efficient manufacturing and chemical purification processes. Whether it is for the petroleum industry, where zeolites are used for catalytic cracking[1], or for the natural gas industry, where they are used to purify methane gas[2], zeolites are an indispensable part of our world.

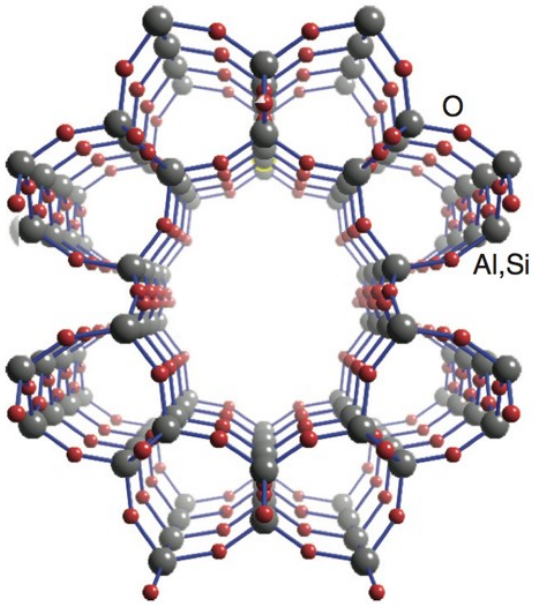

Figure 1.1 *Perspective view down the channel of the channels of a zeolite, with the Al and Si atoms shown in grey, and the O atoms shown in red.*

On the other hand, researchers are continuing to develop new materials with more diverse chemical and structural properties, such as metal-organic frameworks (MOFs). While naturally occurring zeolites were first discovered over a century ago, MOFs are a relatively new class of principally synthetic materials, which are becoming an increasingly popular area of research that spans a wide variety of applications including catalysis[3–5], gas capture and separations[6–8], sensing[9–11], photoluminescence[12–14], electrical conductivity[15,16] and drug delivery.[17,18] These novel microporous materials, formed through self-assembly of inorganic and organic structural building units (SBUs), represent an exciting advancement in porous materials research due to their record breaking internal surface areas, high porosities, and a wide variety of possible chemical compositions. In turn, these features result in an infinite number of diverse topologies with readily tuneable properties.

The work presented in this thesis will be based on two projects related to both MOFs and zeolites, with a greater emphasis given to siliceous and aluminosilicate zeolites. The first of the two projects pertains to the development of a new semi-empirical atomic charge model for fast calculations of the electrostatics in periodic systems with a net charge, as well as the validation of a first-principles based electrostatic potential fitted charge method for frameworks with a net charge. systems. This will be the topic of Chapter 3 of this thesis, where a much more in-depth discussion of the related concepts will be given. The second project, presented in Chapter 4, focuses on the development and validation of a force field for modelling guest-host interactions in aluminosilicate zeolites. In the chapter that follows, an in-depth discussion will be given of the defining characteristics of both MOFs and zeolites. Further, examples of the real-world applications of these materials will be discussed, predominantly in the context of gas adsorption and separations. Next, an overview of recent computational modeling of these materials will be given, in hopes of clearly demonstrating the important role that computer simulations play in driving our understanding of

these porous materials. In the final part of this chapter, the goals of the research presented in this thesis will be outlined.

1.2 Porous Materials: Metal Organic Frameworks and Zeolites

1.2.1 Zeolites

Zeolites constitute an important class of microporous crystalline solids that are the result of corner sharing aluminosilicate or silica polymorphs connected in a tetrahedral, 4-connected framework with uniformly sized pores. Generally, the tetrahedrally coordinated Si and Al atoms are referred to as "T-atoms", or "T-sites". Zeolites can be either naturally occurring or synthetic, with around 40 naturally occurring zeolites, which form in both volcanic and sedimentary rocks, and more than 150 synthetic zeolites. Figure 1.2 gives some examples of both natural (a) and synthetic (b) zeolites. What is most noticeable with about the synthetic zeolites shown in Figure 1.2, is that they have been prepared to have specific macroscopic shapes. Generally, microporous materials will be synthesized as a powder, or in the case of natural zeolites, they will be milled into a powder. They are then structured to have a macroscopic shape which is best suited towards a specific application.

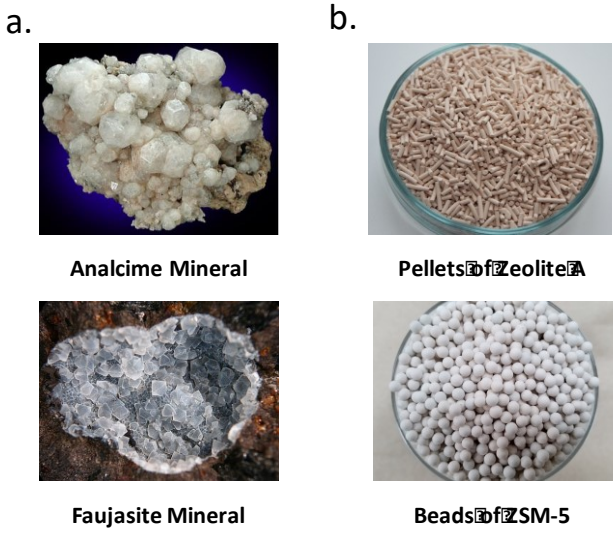

Figure 1.2 *(a) Naturally occurring mineral zeolites (b) synthetic zeolites*

Of the many different inorganic materials that have been researched to date, zeolites are amongst the most widely investigated and discussed. Scientific literature related to zeolites dates back to 1756, when Swedish mineralogist Axel Cronstedt[19,20] reported his observation on a "an unknown species of rock" and coined the name zeolite. Their present day involvement in industry dates to the 1950s when the Union Carbide Corporation developed a process for producing zeolites on a large industrial scale[1]. Still, their synthesis, structures and properties pose fascinating problems even today, while their wide ranging industrial applications has given vitality to this long-energized field of research. The optimization of existing zeolite structures and the search for new materials with new functionalities is continuously driven by the industrial demand for more efficient processes. The long established presence of zeolites in materials science has further resulted in the creation of an extensive human-readable library of chemical and structural data that has been maintained by the Structural Commission of the International Zeolite Association (IZA-SC) since 1996.[21] Within this internet-accessible database there is comprehensible data given for each of the 232 zeolite framework types approved by the IZA-SC. Additionally, there is an abundance of general information regarding zeolites including information about structural classification systems used, different types of structural features present in zeolites, X-ray diffraction data and powder pattern simulation, and references to published works covering a range of topics related to zeolites.[22]

Over the course of the two decades that the database has existed, the number of unique zeolite framework types has increased from 98 to 232, with this number expected to increase further still. In addition to the current 232 unique framework types listed by the IZA-SC database, there are millions of hypothetical zeolite structures that have been predicted through various computational methods that remain to be synthesized and tested.[23–30] Furthermore, there are many unanswered questions regarding zeolites, including the fundamental question regarding the mechanism by which

zeolites form. Thus, the field of zeolite science and technology, although significantly more mature than the MOF field, is still a very active field of research. To facilitate the classification of zeolite topologies as well as allow simple structurally based comparisons between zeolite frameworks, work has been put into breaking the complex networks down into simple representative building blocks. The first and most simple of these are the primary building units (PBUs), which are the most fundamental structural motif in a zeolite framework. These are built from TO_4 tetrahedra ($SiO_4$and, AlO_4 in the case of aluminosilicates) that combine together to form a 3-dimensional network, such as shown in Figure 1.3

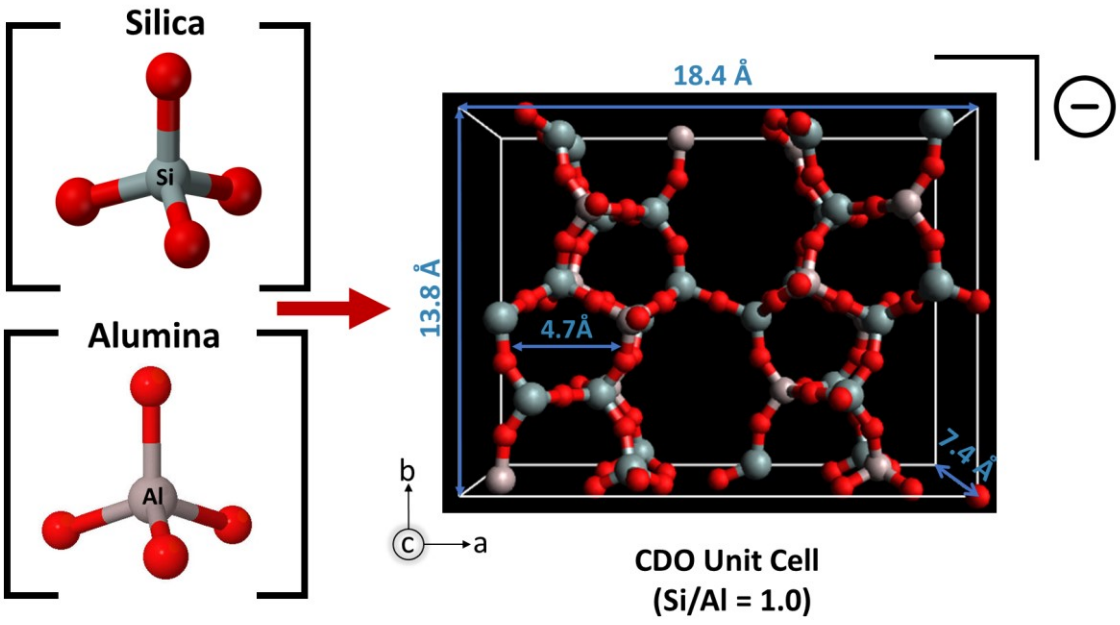

Figure 1.3 *The basic building blocks of aluminosilicate zeolites are tetrahedra of silica and alumina combined in different ratios. Due to the differing oxidation states of Si (4+) and Al (3+), the addition of each Al atom results in a corresponding negative charge on the unit cell.*

From the combination of these basic building blocks, PBUs, the next level of framework substructures is built: structural building units or SBUs. These SBUs can contain a maximum of 16 tetrahedrally coordinated T-atoms, and are derived assuming that an entire framework is composed

of a single type of SBU. However, sometimes more than one can be used to describe the structure. As with all of the sub-structural units used to classify zeolite topologies, there must always be an integral number of SBUs in a unit cell, otherwise the result would be an incomplete structure. Moving up further still in the structural hierarchy, we can see how composite building units (CBUs) are constructed through the combination of SBUs when considering Figure 1.4. The principle difference between SBUs and CBUs is that the latter can contain a much larger number of tetrahedrally coordinated metal centers and, as a result, form much larger and more complex structures. One important feature of CBUs is that some may appear in multiple frameworks and can be useful in identifying relationships between framework types.

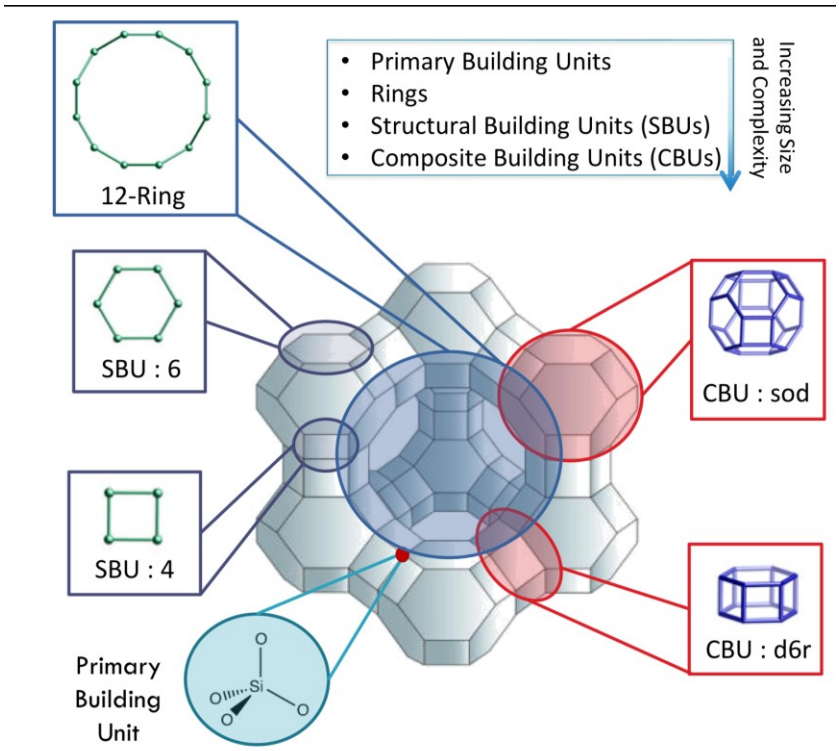

Figure 1.4 *The Hypothetical building units in the faujasite (FAU) supercage.*

Finally combining these together distinct rings and cages are formed. These represent the most important structural motif, as these are the sub-structural units that form the pores and channels that give rise to the physically important and chemically useful properties of zeolites. The

size of the pores and channels will control the diffusion properties of zeolites and act to selectively separate molecules based on shape. Furthermore, it is within the pores and channels that all chemistry will occur, such as catalysis. Figure 1.4 gives examples of each of the structural motifs that we herein discussed. In the cartoon illustration shown, each vertex corresponds to a T-atom, while the lines bridging each vertex would correspond to the bridging O atoms in the zeolite. The ring size is determined by the number of T-atoms, or vertices, present in the ring.

Another system developed by the IZA-SC for classifying zeolite topologies are the framework type codes (FTCs), which are capital three letter codes designated to each unique structure. These FTCs are derived from the name of the type materials, and are not related with the framework composition, the distribution of metal atoms, or the cell dimensions or symmetry. Table 1 gives some examples of FTCs along with the abbreviated name of the type material from which the FTC is derived. There are some special additions to these three letter codes including asterisks (*) and hyphens (-). An asterisk denotes that the framework is disordered or amorphous, while a hyphen indicates an interrupted framework.

Table 1. *Examples of the FTCs of some popular zeolites*

FTC	Type Material
LTA	Linde Type A
FAU	Faujasite
MFI	ZSM-5 (Five)
*STO	SSZ-Thirty-One
-CLO	Cloverite

Disordered, or amorphous, frameworks are a class of zeolites that have defects in the crystal structure. Figure 1.5a shows the representative unit cell of the STO* framework, which is the most representative configuration of the unit cell that would form during synthesis or in nature. In the

case of amorphous frameworks, there are other configurations of the unit cell that are just as likely as the most representative configuration. Figure 1.5b shows an alternative configuration that can arise within the STO* framework. When these configurations of differing symmetry combine, the crystal packing is distorted, resulting in defects in the crystal structure. The IZA-SC database lists 7 unique framework types that are amorphous.

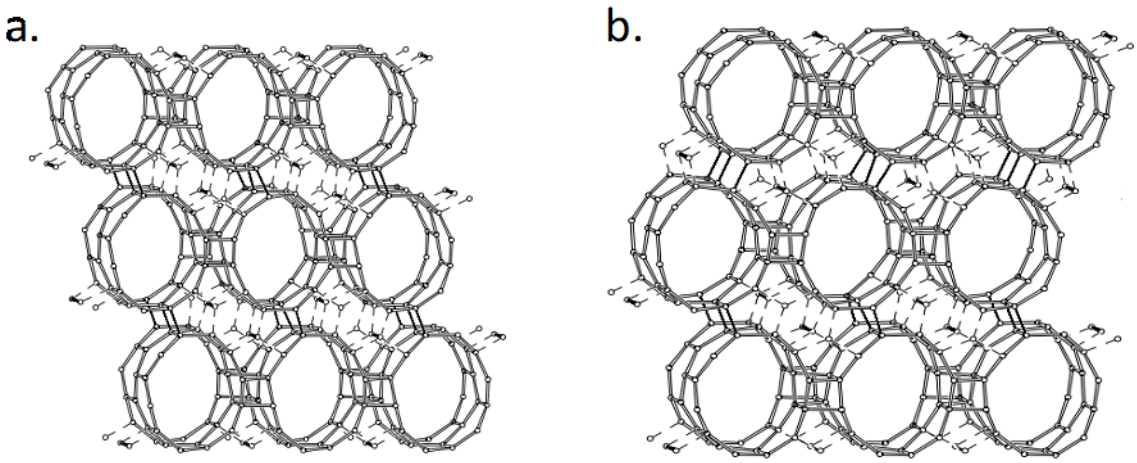

Figure 1.5 *(a) the most representative configuration of the unit cell of the *STO framework, (b) another possible configuration of the *STO framework.*[22]

Conversely, the IZA-SC database lists 14 unique interrupted frameworks, double the number of amorphous frameworks. Within this class of frameworks, not all T-atoms in the framework structure are 4-connected, with –CLO being an example. This can arise when one of the terminal O atoms within a TO_4 cluster is terminated with an H atom. Essentially, these points in the framework represent interruptions in the 4-connected framework that create holes in the net-like structure, as shown in Figure 1.6. In some cases, these interruptions result in defects in the crystal structure and, therefore, an amorphous framework. In general, interrupted frameworks will have larger pore openings as well as increased reactivity, particularly when Al is added. Furthermore, the interruptions in the 4-connected net result in more flexible frameworks.

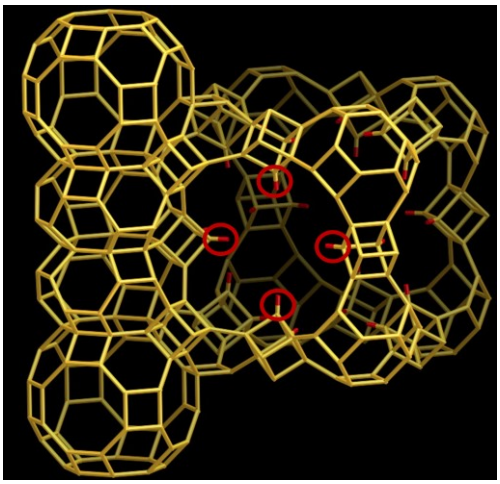

Figure 1.6 *-CLO interupted framework viewed normal to the {001} plane. Some of the interruptions present in the unit cell are circled in red for clarity.*[22]

The addition of aluminum to siliceous frameworks results in some interesting properties that act to increase the reactivity and polar character of the frameworks, as well as bringing about some of the most long-standing issues in the structural characterization of zeolites. When a zeolite is composed solely of Si and O atoms, the +4 oxidation state of the Si atoms is perfectly balanced by the negative charge of the 4 surrounding O atoms. However, when an Si atom is substituted by an Al atom, which is in its +3 oxidation state, there is a resulting imbalance in the charge distribution. Therefore, to compensate for this imbalance in charge, extra-framework cations occupy the pores and channels of the zeolite. Here, the ions increase the polar character of the framework and enable them to act as more effective solid sorbents for chemical separations and catalysis. Additionally, the extra-framework cations can be used to control the size of the pores during synthesis. Yet, it is the inclusion of theses cations that gives rise to the first long-standing issue in the structural characterization of zeolites, that being the identification of the cation positions within the framework. The main issue is that the extra-framework species do not generally follow the symmetry of the framework, and so their positions are disordered.[31]

For example, in the case of the LTA-4A zeolite fully saturated with Al, each 8-ring can accommodate only one Na^+ ion. However, because of the 4-fold symmetry of the 8-ring, there are four equivalent positions for the Na^+ ion to occupy (Figure 1.7). In this case, one Na^+ ion may hop between four equivalent positions (dynamic disorder), or it may be stationary but occupy different positions in different 8-rings throughout the framework (static disorder). Conventional X-ray diffraction (XRD) analysis cannot distinguish between these two possibilities, and an electron-density map generated from the diffraction data would show ¼ of an Na^+ ion (10 electrons) at each equivalent position. When the charge is balanced by H^+ ions instead of Na^+ ions, this is no longer an issue as they will bind to electron deficient O atoms adjacent to Al in the framework. Generally, the position of H-atoms can be determined with infrared spectroscopy (IR) and neutron diffraction studies in conjunction with XRD.[32–35] These acidic protons can also act as reactive centers, which make zeolites especially useful in the petroleum industry as will be discussed later.

The inclusion of Al into zeolite frameworks results in another of the long-standing issues in the structural characterization of zeolites, which is that the Al atoms cannot be differentiated from Si atoms by traditional XRD techniques. XRD relies on the scattering of X-rays by the electrons surrounding a given atom. Because Si is in its +4 oxidation state and Al in its +3 oxidation state, the two ions have exactly the same electronic configuration. Further, both atoms have nearly the same ionic radii at around 54 pm. Because of these two properties, these ions are indistinguishable in their diffraction patterns. It is for this reason that nearly all crystal data for zeolites is given with only silicon in the framework. One particular exception to this is the LTA-4A framework with an Si/Al ratio of 1. At this Si/Al ratio, there are exactly the same number of Si and Al atoms present in a given unit cell. When Si is substituted by Al in a zeolite framework, the substitution occurs in accordance with the empirical Löwenstein rule[37], which states that there can be no Al-O-Al linkages

in the framework. Thus, at such a low Si/Al ratio, Löwenstein's rule allows one to deduce the positions of Al in the LTA-4A framework as there are limited positions for the Al atoms to occupy. The resolution of the Al and Na^+ positions in LTA-4A were carried out by Pluth *et al.*[38] in 1980, with the data readily available within the IZA-SC database.

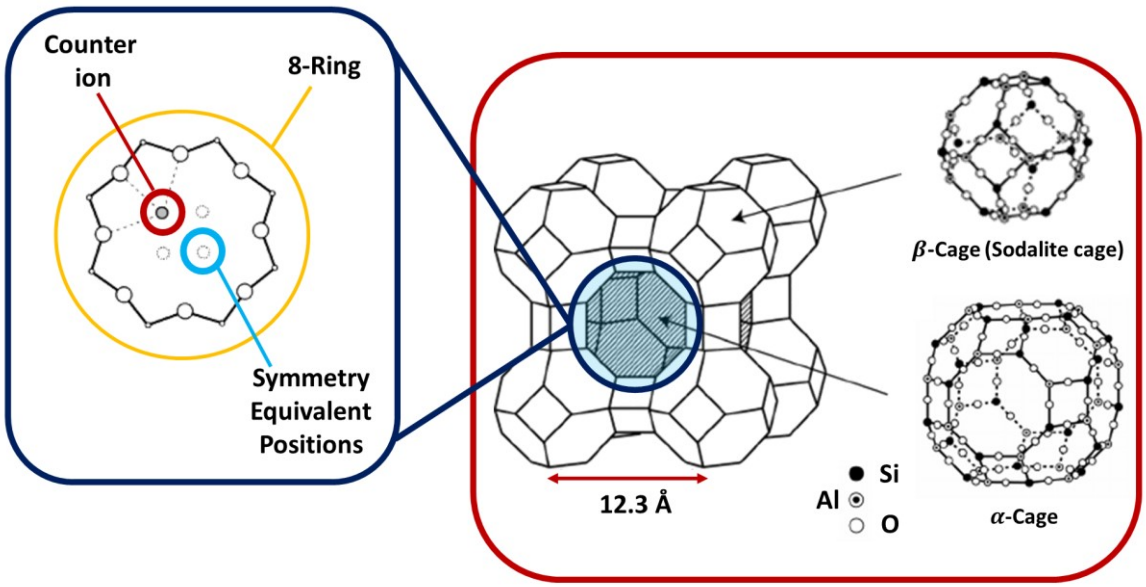

Figure 1.7 *Illustration of disorder in the positions of the extra-framework ions in an LTA-4A aluminosilicate. The disorder arises when there are multiple symmetrically equivalent positions within a pore or channel that a cation can occupy. In such cases the X-ray crystal data will be smeared and, in the example given above, would predict that a quarter of the ion is in each of the symmetry equivalent positions shown. (Figure adapted from another work).*[36]

The physical reasoning behind Löwenstein's rule is intuitive: introducing Al into the framework brings with it a negative charge and having two Al in close proximity would be energetically unfavourable.[39] The principle of aluminum avoidance, as it was called originally by Löwenstein, was met with much controversy to start. However, with time it became clear that that Al-O-Al linkages are not observed in materials prepared by normal hydrothermal routes. Although more recently, some studies have been performed that show that solid-state, high temperature preparations of aluminosilicates can result in violations of the Löwenstein rule. [40,41]

Hence, even with a relatively simple chemical composition, zeolites can be found in a variety of unique geometries, all with their own interesting properties and applications. Though the field of zeolite science and technology is the most long-standing and topical of all inorganic materials, there are still many interesting questions left to be answered, both theoretically and experimentally. Therefore, new models should be developed to help further the understanding of this class of porous materials, including having a better understanding of the kinetic and thermodynamic factors that come into play during synthesis. In later sections of this chapter, specific applications of zeolites will be discussed as well as computational simulations that have been performed to help improve our understanding of the structure of these materials.

1.2.2 Metal Organic Frameworks

Interest in metalorganic frameworks (MOFs) has increased significantly in the past decade, with their currently being over 20,000 published MOFs.[42] This spike in the growth of research related to MOFs can be traced back to the early 2000s when Yaghi *et al.*[43] published their seminal work on MOF-5 (Figure 1.8). In their work, Yaghi and coworkers introduced the use of reticular chemistry to achieve a systematic and controllable process for designing new MOFs. It was also in this paper that the term MOF was coined, before which these materials were known as porous coordination polymers (PCPs), which reflects the involvement of coordination chemistry as well as the periodicity of MOFs. Although the work of Yaghi marked the beginning of a steady increase in research related to MOFs, they had already been investigated for many years by pioneers such as Susum Kitagawa of Kyoto University in Japan.[44,45] Yet, what has really made MOFs such an exciting field of research is their endless tuneability, and therefore the potential problems for which they can be applied to.

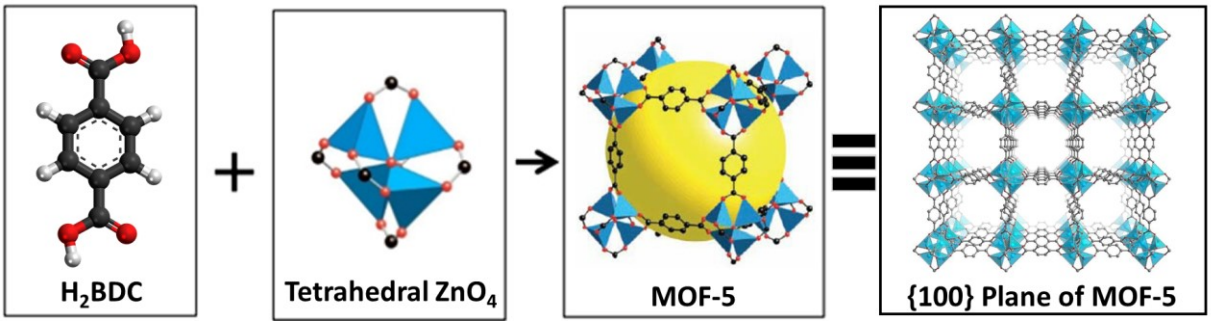

Figure 1.8 *Organic and inorganic SBUs that combine to form MOF-5 via reticular synthesis. The yellow sphere included in the third image from the left gives a visual indication of the pore size of MOF-5.*

The structure and basic function of MOFs is analogous to zeolites, in the sense that both are composed of interlinked, repeating units that form uniform pores. However, unlike zeolites, MOFs are microporous materials that can be synthesized from an infinite combination of organic and inorganic structural building units (SBUs). As a result of this unlimited number of combinations of SBUs there is an infinite number of unique MOFs possible, each with different geometric and chemical properties. Another major difference compared to zeolites is the relative ease with which MOFs can be synthesized. To date there are only a little over 200 zeolite structures reported in the IZA-SC database[22], in contrast with the thousands of MOF structures listed in the Cambridge Structural Database (CSD), with this number doubling every 4 years on average.[42]

Like zeolites, MOFs can also have a net charge, which can be the result of coordinatively unsaturated metal centers, also known as "open-metal sites". Compared with their coordinatively saturated relatives, open-metal site MOFs present certain advantages with respect to adsorption of guest-molecules. The exposed framework metal cation will interact preferentially with guest molecules giving rise to stronger physisorption than for MOFs lacking open-metal sites, while preserving interactions that are well below the strength characteristic of chemisorption (<50 kJ/mol). The ability to physisorb guest-molecules is one of the important features of both MOFs and zeolites, providing ideal conditions for easy desorption of guests, which is an important feature

for many of their applications. A well-known example of an open-metal site MOF is provided by Mg-MOF-74, where unsaturated Mg sites and 2,5-dioxide-1,4-benzenedicarboxylate (DOBDC) provides strong localized binding sites that demonstrate high and selective CO_2 adsorption. At conditions relevant for industrial gas separations, Mg-MOF-74 can adsorb 89 g of CO_2 per kg of material and will release 87% of the captured CO_2 at room temperature.[46,47]

However, depending on the application, open-metal sites can present some disadvantages. For example, having open-metal sites increases the ease with which guests can access the metal, and relative to coordinatively saturated framework metal centers, open-metal sites tend to be much more hydrophilic. Hence, in the presence of humidity, open-metal site MOFs will have a greater tendency to bind water over other guest molecules, reducing the adsorption capacity for target molecules.[48] Furthermore, the water can cause the MOFs to decompose as is the case with the well-known coordinatively unsaturated Mg-MOF-74. Water stability systems such as these still presents a major challenge for MOFs that are to be used as industrial sorbents, since humidity is present in many typical adsorption processes of interest.[49–51] Other gaseous impurities found in industrial process streams, which include H_2S, SO_X, and NO_X, have also been shown to bind strongly and selectively to certain frameworks, resulting in a lowered adsorption capacity for the gas that is meant to be adsorbed.[52–54] Acquiring a better understanding of this issue in order to design better MOFs for specific applications is still an outstanding challenge in the field of porous materials and is currently a highly active field of research.

Originally, MOFs were prepared by diffusion techniques, in which solutions containing metal salts were slowly diffused into solution containing organic ligands and weak bases. This approach was very time-consuming and gave very low yields.[55,56] Currently, solvothermal methods, where metal salts and organic linkers are electrically heated in solvent (such as water or ethanol) at

high temperature for a time up to several days, are used to produce MOFs on large scales by companies such as BASF.[57,58] Such recent synthetic advances have allowed MOFs with targeted pore sizes, internal surface areas, specific topologies and surface chemistry to be synthesized and tested. MOFs can be constructed with essentially any metal center, ranging from uranium-centered MOFs for nerve agent detection[11] to biocompatible Fe-based frameworks.[59] Lanthanides can also be used to construct luminescent MOFs, generally applied to enhance bio-imaging techniques.[60]

The main characteristic feature of MOFs is their porosity, which makes these inorganic-organic hybrids similar to zeolites, with pores that can range from microporous (between 2 - 20 Å) to mesoporous (between 20 – 500 Å). In fact, many MOFs exists with much larger pores compared with zeolites, resulting record internal surface areas. For example, the surface area of 1g (or about a tablespoon) of some MOFs is comparable with that of a square football field (approximately 5351 m^2). On such material is NU-110, which has an exceptional internal surface area of 7140 m^2/g. This MOF also holds the world record for pore volume (4.40 cm^3/g), which is over 4 times greater than that of average zeolites, silicas or porous carbons.[61,62] It is also possible for MOF frameworks to be flexible and exhibit a behavior similar to breathing upon gas adsorption, which enables the pore sizes to change significantly and accommodate larger guest molecules.[63]

Still, even with such control over the desired properties of MOFs, most MOFs have relatively weak thermal and chemical stability compared to zeolites, which restricts their use in large scale industrial applications. This is and cost are the principal reason that zeolites are still the most widely used porous materials within large-scale industrial applications. However, with growing global demands, there is a constant growing need for new technologies in areas such as power generation and storage, industrial chemical production and pharmaceutical drug delivery. Although zeolites currently play a role in some of these areas, MOFs would greatly increase the efficiency of

these processes as well as contribute to areas that zeolites cannot. This outlook on the future of industry and global energy demands is what pushes the discovery of novel, record-breaking MOFs with fine-tuned properties.

1.3 Applications of MOFs and Zeolites

A continuous and vigorous industrial interest and involvement in zeolite research drove initial applications to employ the adsorptive properties of zeolites. The first example of such application of zeolites can be traced back to the end of the 1950's, after the synthesis and discovery of the A-type zeolite by the company Linde (LTA zeolite).[64] Such adsorptive application of zeolites currently account for 10% of the global consumption in terms of weight by ton.[65] Later introduction of zeolites as catalysts to the petrochemical industry further intensified interest into these material and can be considered the most crucial event in the history of zeolites.

With the introduction of X and Y-type zeolites for catalytic cracking at the beginning of the sixties, zeolites became one of the most widely used heterogeneous catalysts in the chemical industry. In recent times, the total global consumption of synthetic zeolites is estimated to be approximately 3×10^5 metric tons per year, with around 95% of them being FAU-type zeolites, which are employed as catalyst for fluid catalytic cracking (FCC)[65]. Other important processes in the petrochemical industry, such as refining, make use of other specific zeolites, for example the MOR-type zeolite that is used in the cumene and ethylbenze processes or the MFI-type zeolite used to catalyze hydrocarbon isomerization reaction as well as alkylation of hydrocarbons. Undoubtedly, the use of zeolites in the petrochemical industry has brought with it significant advancements in efficiency, by both reducing costs and increasing yield by at least 30%.[66]

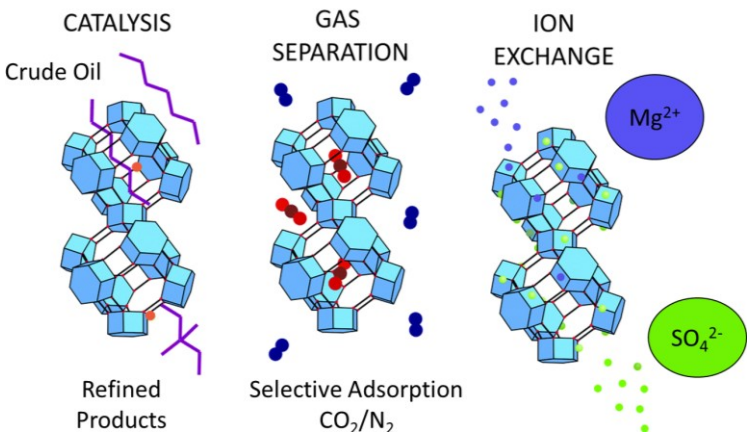

Figure 1.9 *The three most important industrial applications of zeolites which make them an indispensable part of the modern industrial world.*

In terms of the volume of zeolites consumed annually, use as heterogeneous catalysts only accounts for 17% of annual global consumption, while on an economic basis heterogeneous catalysis represents 50% of the total market value.[65] The largest portion of the annual consumption of zeolites results from their use in laundry detergents as ion-exchange agents , accounting for 73%.[65] Use of zeolites as catalysts represents the second largest market for synthetic zeolites in terms of volume, accounting for approximately 17% of annual global consumption. Although currently the detergent industry is still the largest consumer of zeolites today, the use of zeolites as detergent builders is expected decline , while on the other hand the use of zeolites as catalysts is expected to continue increase alongside the ever increasing demand for petroleum derived feedstocks.[65] Notwithstanding, all of these applications, which are of considerable economic importance, arise from the structural and chemical properties of zeolites. In particular, there is no single zeolite that is of any industrial importance that is completely siliceous, and all will contain a proportion of Al atoms with a corresponding number of charge balancing counter-ions. In this regard, aluminosilicates are among the most important material for the modern industrial world. The three main areas of applications for zeolites, are summarized in Figure 1.9.

Similar to zeolites, much of the driving force behind the steadily increasing amount of research related to MOFs, is due in large to the many opportunities to improve current technologies utilized in industry. MOFs too can be applied to each of the same areas of application as zeolites (Figure 1.9). However, there is a multitude of additional applications that are specific to MOFs. Again, this is a result of their greater complexity relative to zeolites, which allows for much more freedom in fine-tuning the properties for a given application. In addition to gas separations, MOFs have been investigated as mediums for gas storage, particularly vehicular hydrogen and methane gas storage.[67–69] MOFs can also be used in catalytic applications, although none to date have been used commercially on an industrial scale.[70] Both the functional groups that are built into the framework and the open metal sites present in MOFs with coordinatively unsaturated metal centers can act as catalytic sites. Due to the increased internal surface area and pore sizes relative to zeolites, there is a significant amount of control over the design of the catalytic environment. For instance, MOFs can be used as synthetic analogues of natural enzymes such as cytochrome P450, chymotrypsin, phosphotriesterase, hydrogenase, and carbonic anhydrase.[71] Functionalization of the organic linkers allows one to shape the pore and add active sites along the framework surface. Additionally, the pore channels within MOFs have been used as reaction hosts for photochemical polymerization reactions.[72,73] Indeed, there is an abundance of reported examples of MOFs being used as heterogeneous catalysts under a variety of conditions, identifying MOFs as excellent candidates as selective and tunable heterogeneous catalysts.[74–77] Furthermore, they may be applied to assymetric catalysis using homochiral MOFs to direct enantioselective reactions.[78] For a more detailed overview of the catalytic applications of MOF, Kitagawa *et al.* have published an extensive review.[79]

A large consideration in selecting MOF materials for a specific application is the thermal and chemical stability. Additionally, the cost of the materials needed to synthesize MOFs is relatively

high , currently making their application toward industrial processes economically unfeasible when compared with the cost of zeolites.[80] However, MOFs are beginning to find their own niche in the world of industrial applications. Many in the scientific community expect that MOFs will become both prolific in industry and ubiquitous in broader everyday consumer applications in the near future.[81,82]

1.3.1 Gas Applications of Zeolites and MOFs

One of the principal applications of both MOFs and zeolites is gaseous separations, which can be used to increase purity of one gas in a mixture, or to capture harmful greenhouse gases. With the rise of societies environmental awareness, climate change has become a familiar household term as well as a global issue that receives a lot of attention.[83–85] Growing food shortages and increased frequency of extreme weather, along with rising sea levels and degrading ecosystems have increased the demand for solutions to this problem that will affect future generations around the globe.[86–88] Among the pollutants responsible for the global shift in climate, anthropogenic CO_2 emissions are the biggest contributor with approximately 32.3 giga tones emitted globally each year.[89] The two largest contributors to global CO_2 emissions are power generation, accounting for 29% of total global greenhouse gas (GHG) emissions, and burning fossil fuels for transportation, accounting for 27% of global GHG emissions.[90]

To help reduce the CO_2 emissions related to power generation, carbon capture and sequestration (CCS) is a prominent solution. CCS is a proposed strategy that encompasses all technologies which aim to selectively filter CO_2 from process streams, which will then be pressurized and transported for permanent storage in either deep sea reservoirs or depleted oil and gas fields.[91–93] Alternatively, the recovered CO_2 can be utilized to retrieve the remaining oil and gas in nearly depleted wells by injecting the CO_2 into the well.[94–96] In regards to reducing GHG emissions related

to transportation, alternative fuel sources are being investigated as well mediums for storing these fuels. In particular, natural gas is one of the most applicable and cleanest fuels in comparison with other fossil fuels such as oil and coal.[97] In all of the cases mentioned, selective adsorption of a gaseous species is required.

In the natural gas market, zeolites already play a very important role in purifying natural gas, which generally consists of a gaseous mixture of different hydrocarbons (consisting mainly of methane), and other impurities such as CO_2. N_2, H_2S, and water. Of these, CO_2 is the major contaminant in the natural gas taken from the Earth, which in addition to reducing the heating value of natural gas it also, in the presence of water and other contaminants, can turn into corrosive gases that can damage pipelines. In presence of CO_2, the heating value of natural gas is reduced by nearly half, from around 11 kWh/ m^3 for pure natural gas to 6.5 kWh/m^3. This loss in energy content is the principal reason that the natural gas upgrading is of such importance.[98] The conventional method for CO_2 removal from natural gas is an amine adsorption process, where the CO_2 is chemisorbed, forming covalent bonds to the amines. Consequently, the energy required to break the CO_2-amine interactions and regenerate the amine solution is exceptionally high when compared to adsorption processes utilized by solid porous sorbents such as MOFs and zeolites. To recover the amines, they must be brought to a boil at high temperatures (~150 C). This not only increases the energetic cost of using liquid amines, but the high temperatures also cause them to deteriorate over time making them inefficient on a large scale. Adsorption in solid sorbents, on the other hands, is the result of transferring some gaseous molecules to the surface of the sorbent. In this process, adsorption takes place based on the affinity between the fluid phase and the surface of the solid sorbent via physisorption, resulting in a much lower binding energy than chemisorption (< 50 kJ/mol). In other words, if there is a mixture of gases in the fluid phases, one of the gases can have a

stronger affinity toward the sorbent surface and will be preferentially adsorbed over other gases in the mixture. These two features are what make solid sorbents such as MOFs and zeolites front runners for these types of gas applications: physisorption allows the adsorption to be reversible and selective at lower energetic costs. Figure 1.10 summarizes the industrial processes by which reversible adsorption of a gaseous species is achieved within porous materials.

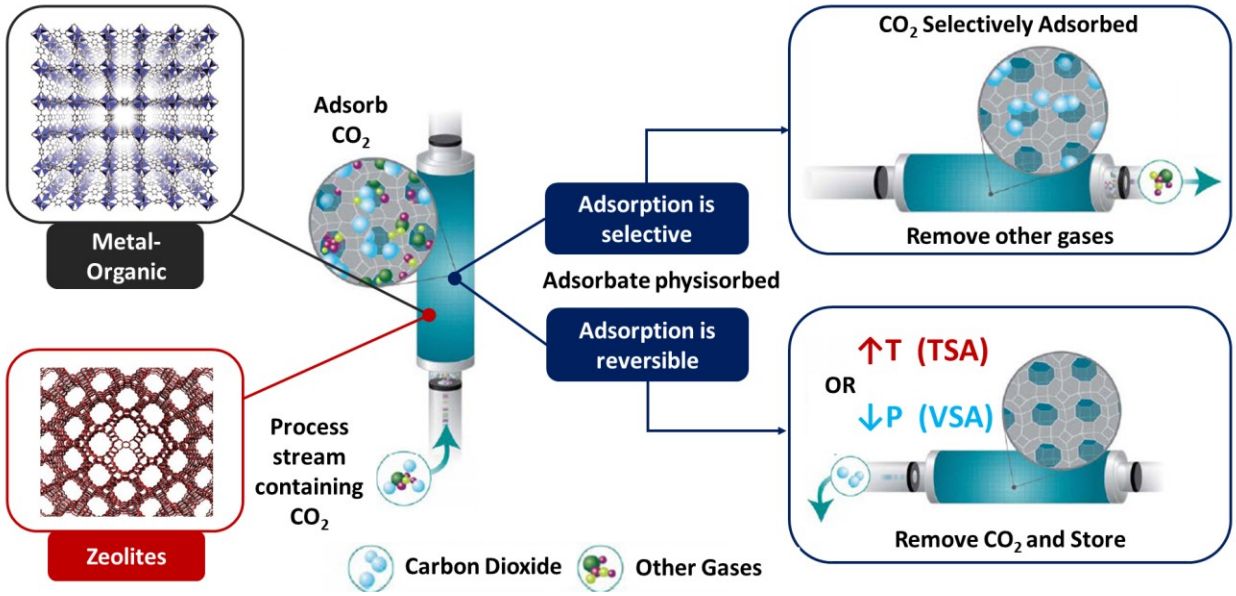

Figure 1.10 *Illustration of the different reversible adsorption processes utilized in industry with porous materials, such as MOFs and zeolites. The use of porous materials for capturing GHGs or purifying natural gas relies on the fact that adsorbates are physisorbed rather than chemisorbed. The lower energy interaction of physisorption (< 50 kJ/mol) is strong enough to selectively bind certain adsorbate molecules but not so strong as to prevent the desorption of the same molecules from the porous framework. Two unique processes are used to desorb molecules from the porous frameworks. (1) TSA, where the temperature is increased and (2) VSA where the pressured is reduced, in order to desorb the guest molecules. The third method used, VTSA, is simply a combination of TSA and VSA.*

Different zeolites are used for processing natural gas including zeolite 4A (LTA with Na^+ ions in pores), but zeolite 13X (FAU topology) is by far the most widely used zeolite for industrial gas applications due to its large capacity in comparison to other zeolites.[99–101] Although MOFs have not yet been applied to natural gas purification, or other gaseous applications, on as large of a scale as zeolites, MOFs represent a future direction of materials research related to these types of applications. The main drawback with zeolites is their smaller internal surface areas, resulting in a

smaller adsorption capacity when compared to MOFs. Additionally, in such gaseous separation processes, their water stability can also pose a problem. In the presence of water the adsorption capacity of zeolite 13X is reduced significantly (Figure 1.11).[102] Though this is also a problem with MOFs, there are fewer possible modifications that can be made to zeolites to improve this property (for example, addition of functional groups to the organic linkers in MOFs to increase hydrophobicity).

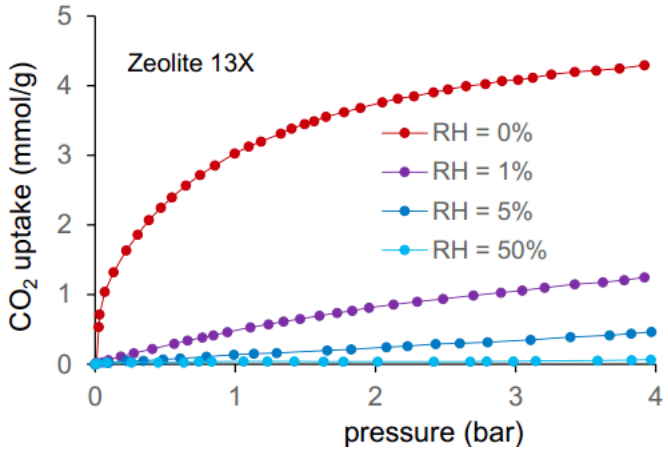

Figure 1.11 *The amount of CO_2 adsorbed by zeolite 13X increases drastically even with a small increase in the relative humidity (RH) of the gas stream: at 1% RH the CO_2 capacity has already dropped by approximately 50.*

Beyond natural gas processing, zeolites have been applied to variety of other gaseous applications such as H_2 production or recovery[103,104], dehydration and CO_2 removal from air for cryogenic processing[105], separations of high-purity oxygen from air via VSA[106–108], and recovery of inert gases.[109] On the other hand, MOFs also represent promising candidates for many of the same applications of zeolites in addition to other MOF specific applications, such as gast storage amd CCS. Gas storage has been extensively investigated in the literature due to the exceptional internal surface areas of MOFs.[69,110,111] The front runners for these applications include MOF-177 (promising material for gas separation and storage at room temperature[112,113], MIL-101 and Mg-MOF-74 (a common benchmark material for post-combustion CO_2 capture)[114], have internal surface areas of

5640 m^2/g, 5900 m^2/g, and 1525 m^2/g respectively. Due to the lower surface area of Mg-MOF-74, it would not be suitable for gas storage but is still the front-runner for post-combustion CO_2 capture due to its high selectivity for CO_2 over N_2.

1.3.2 The Future of Zeolites and MOFs

As demonstrated throughout this section, zeolites and MOFs have comparable characteristics that enable their application to many of the same types of chemical processes. As such, it may seem that with greater structural and chemical diversity, MOFs, will one day replace zeolites when large scale production of MOFs is more economically feasible. However, there are many like examples for zeolites, which demonstrate their importance as well as identify future opportunities within industry. In fact, although MOFs appear to be the bigger, better version of zeolites, there is no indication that they will replace zeolites. For example, as catalysts it is more likely that MOFs and zeolites will be applied to different industrial processes, finding their respective niches within the large global market for heterogeneous catalysis. Furthermore, the literature makes it quite evident that both zeolites and MOFs will continue to be of great interest for industry. This is particularly true due to the potential that both materials present in a diversity of fields that are of vital importance to industrialized society, today and in the future. As will be seen in the following section, computational modeling is aiding in the development of databases of hypothetical materials, with the number of possible framework structures for both zeolites and MOFs being essentially infinite. Therefore, it can be expected that both zeolites and MOFs will play a prolific and synergistic role in various industrial applications such as raw material change, energy storage and environmental pollution control.

1.4 Computational Modeling of MOFs and Zeolites

This section will present some of the computational modeling that has been done with MOFs and zeolites. An important goal of computational chemistry is to determine principles that aid in the design of new materials optimized for specific application. In this regard, this section will be focused mainly on computational modeling related to the design of zeolites and MOFs, and the testing, or *screening*, of these materials for determining what application they are best suited for. The growing sophistication of computer-aided materials design models has promoted a considerable amount of progress in both fields.

1.4.1 Structural Modeling: Development of Hypothetical Databases

Structural modeling of zeolites has contributed greatly to our understanding of many aspects of zeolite synthesis and chemistry. Consequently, there is considerable value in exploring the different possible zeolite topologies in order to understand, at least in a general sense, the relationship between structure and the likely properties so as to generate a library, or database, of potential zeolite structures. These databases can then be screened with relative ease in order to identify candidate materials with properties that one wishes to optimize. In this way, computational modeling can be used as a tool to identify possible materials that would be best suited for specific applications. This information, in turn, could be applied by synthetic chemists as a guide to design new materials.

Examining the structure of zeolite frameworks, they can be considered as a network of corner-sharing tetrahedra. In the 3-dimensional structure, each T-atom forms the vertices of a 4-connected net that uniquely defines each zeolite topology. This, in addition to the fact that zeolites are both periodic and contain well-known SBUs, makes the prediction of new zeolite topologies relatively straightforward. Systematic development of hypothetical databases of zeolite structures

using computational methods first started in the 1990s. At this time, Treacy *et al.*[115] were able to generate a series of structures using a combinatorial method called the Symmetry Constrained Intersite Bonding Search (SCIBS) method. More recently, Deem *et al.*[27,28] used a simulated annealing approach to generate and refine more than 2.6 million zeolite frameworks.[116] A more sophisticated model applies graph theory, making use of the fact that zeolites can be represented as space-filling tilings. Using this approach, Delgado Friedrich *et al.*[117] demonstrated that all possible tilings may be generated from a given number of unique tetrahedral nodes (where the edge of four tiles meet at each vertex), suggesting the possibility of a near complete solution to generating hypothetical zeolite topologies. However, as the number of possible configurations of 3-periodic 4-connected tetrahedra are infinite, this solution would have to be defined within a set of defined boundaries. Once obtained, the graphs are converted into zeolite structures by appropriately placing silicon and oxygen atoms on the vertices and along the tile edge, respectively, with the final geometry optimized using force field methods. In general, force field based simulations are the preferred method in both generating and screening databases of hypothetical material topologies, due to the large number of calculations required.[118–121]

It should be noted that the majority of the work done with modeling of zeolite structures has assumed that they adopt purely siliceous frameworks due to difficulties resolving Al-sitings, as previously discussed. However, methods are being developed for modeling the aluminum distributions in zeolite frameworks. An example of one such method is outlined by Garcia-Perez *et al.*[122], where a procedure is given for identifying the likely Al and extra-framework cation positions within known zeolite structures. Using force field based methods, a series of Al configurations is evaluated to identify the configuration that gives the best match between experimental and simulated adsorption data. As an illustration of the methods, the authors carefully analyzed different zeolite

structures for which experimental data was available for comparison and there are a feasible number of Al distributions per unit cell to consider. Of the frameworks considered, the adsorption of alkanes in Na-MOR was found to be strongly influenced by the Al distribution. For the MOR framework with 8 Al atoms placed in the unit cell, the methane adsorption for 16 unique structures were evaluated at 293 K, and compared to experimental data. In the structure that was found to give the best agreement with experimental uptake data, the Al atoms were located in 4-rings of the framework. The latter result was in complete agreement with the suggested preferential sites of Al found in previous crystallographic, experimental and theoretical investigations.[123,124]

Using this procedure for modeling the Al and extra-framework cation distributions, an extensive study was performed in order to predict CO_2 adsorption in hypothetical zeolites with varying Si/Al ratios.[125] The analysis carried out in this work helped identify two features of zeolites that are related to the CO_2 uptake. The first of these features is the free volume within the pores, and the second is a closely related property, the nearest neighbour distance. The latter describes the maximum distance separating an atom within the pore a zeolite from the framework atoms, while the former describes the total amount of gas that can be held within the pore. Utilizing these new descriptors, the authors evaluated the CO_2 uptake in experimentally realized structures taken from the IZA-SC database. The results of the screening are further elaborated upon in the following section. From this study, we are provided with an excellent illustration of how well-developed force fields can be applied towards obtaining meaningful results, and identifying design principles.

As with zeolites, the generation of hypothetical MOF databases is also possible, again facilitated by their modular nature. MOF building blocks composed of organic and inorganic structural building units, or SBUs, combine together to form 2 or 3-dimensional structures with a high degree of symmetry. Due to the near infinite possible combinations of SBUs and their

arrangement in three-dimensional space, computational methods have proven to be highly useful in generating hypothetical MOFs. Similar approaches to those taken for generating hypothetical zeolite databases are employed in developing hypothetical MOF structures. For instance, Boyd *et al.*[126] recently presented a method for constructing hypothetical MOFs, using a graph theoretical approach. Using this approach, it is possible to build MOFs with 2- and 3-dimensional periodicity by combining together different SBUs in an automated fashion. Methods such as this one can be used to generate very large databases of materials. This large collection of hypothetical structures can then be screened using force field based simulations, to identify high-performing materials as well as possible structure property relationships.[127,128] As a result of the small number of experimentally realized MOFs (~6,000 structure listed in the CoRE database[129]), development of hypothetical frameworks has received considerable attention in recent years, and methods similar to those described for zeolites are employed regularly.[126,128,130,131]

To conclude this section, it should be noted that the field of structural and topological characterization of zeolites and MOFs is a highly active one, having multiple different branches of study simultaneously on-going. The focus of these branches includes topics such as framework flexibility, tilings, and characterization in terms of local geometries. Continued research such as this will help in developing better models of these materials with the hopes of designing materials towards specific applications. In this respect, computational screening of materials has proven to be a powerful tool. The strategy frequently employed in material screening studies is to identify trends in the material activity based on geometric descriptors such as pore size, free volume and internal surface area. There is the additional possibility for non-geometric descriptors that try and capture the chemical properties of the systems, such as the atomic property weighted radial distribution function

introduced by Fernandez *et al.*[132] However, in order to accurately identify such trends, it is first important to have a large number of diverse materials to be sample.

1.4.2 Screening Studies: Search for the Best Materials

As in many branches of science, the use of computer simulations has aided in examining the physical and chemical processes involving nanoporous materials.[133] Since the study of phenomena such as adsorption, separation and diffusion in these materials generally involves explicit consideration of hundreds to thousands of atoms, classical simulations that utilize empirical force fields to evaluate interaction within these systems are in wide use. In most classical simulations, atomic pair potentials are used to model the non-bonded guest-host interactions, which are generally composed of two components: dispersive, and electrostatic. The electrostatic component of the interaction energy is frequently described by the fixed partial atomic charge approximation. Of the two components, the electrostatic interactions will both be the strongest and over the greatest range. Therefore, the assignment of atomic charges is crucial to the accuracy of force field based simulations. This is especially true since, unlike the potentials describing dispersive interactions, the charges tend not to be transferable between frameworks. Presently, the most widely used method for generating atomic charges is to fit them to reproduce the electrostatic potential (ESP) derived from first principles quantum mechanical (QM) calculations. These so-called ESP -charges have been widely used in modeling MOFs[134–137], while zeolites are more frequently modelled using generic charge sets that are part of an optimized force field.[138–140]

Using such a force field, Smit *et al.* set out to assess zeolites for application towards post combustion CO_2 capture. The screening study involved evaluating over 130,000 aluminosilicate zeolite structures with either Na^+ and Ca^{2+} cations intercalated throughout the pores. The framework structures containing Al were generated by randomly replacing Si atoms with Al while

satisfying Löwenstein's rule. To model the guest-host interactions, the atomic charges were taken from the force field developed for aluminosilicates by Calero *et al.*[141] In addition to the use of generic charges, the force field employs a Lennard-Jones potential with parameters optimized specifically towards predicting CO_2 adsorption in aluminosilicates. With the Calero force field, Smit *et al.* were able to obtain simulated adsorption data for thousands of materials efficiently. From their analysis, they found that the structures that had the highest CO_2 uptake were those that possessed both large free volumes and specific configurations of TO_4 units that maximizes the distance between nearest neighbours within a range of 3 and 4.5 Å. From both experimental and modeling points of view, this finding provides specific properties that can be used as predictors for high CO_2 uptake in zeolites at different Si/Al ratios and for different cation types. To apply these new descriptors, the authors applied them towards evaluating the CO_2 uptake in experimentally realized, pure-silica structures, taken from the IZA-SC database. From this analysis, the SAO and RWY frameworks were found to outperform the more commonly used, and more thoroughly studied, FAU and LTA frameworks. Although these descriptors provide important insights into structure property relationships, there are many other important factors that determine whether a particular material is *better* performing in practice. For example, the synthetic feasibility of the structures, cost, stability diffusion coefficients, as well as the adsorption thermodynamics will also come into play when ranking the relative quality of materials. Hence, it is important to continue developing such descriptors to allow for a more robust screening of both hypothetical and experimentally realized structures in the future.

A more sophisticated study, carried out by Hasan *et al.,*[142] introduces a hierarchical computational approach that combines materials selection with process optimization as a means of identifying cost-effective CO_2 capture materials. The first step in this process is an initial screening

of a database of 194 siliceous zeolite frameworks (based principally on the IZA-SC database),[143] based on structural features related shape selectivity and size selectivity, such as pore size, accessible surface area, free volume and limiting pore diameter. In the following step, the Henry coefficients are calculated on the 18 top performers selected from the initial screening, and filtered further based on the CO_2/N_2 selectivities. From this, a small subset of 24 zeolites were selected for the VSA process optimization to obtain the optimal capture and compression cost, purity, recovery, and energy penalty associated with the entire process. Using this approach, a total 15 experimentally realized frameworks were found to capture CO_2 at 90% purity and 90% recovery with the cost of compression for the top ten performers sitting around 30$, when running the VSA process under the optimal conditions associated with each material. From the complete analysis, the top four frameworks in terms of cost-effectiveness were identified as AHT, NAB, MVY, and ABW. When filtering the materials based on the minimum parasitic energy, the AHT and ABW frameworks are also found amongst the top three materials. Of these, the AHT framework offers the most significant improvement over the current standard, zeolite 13X. Under optimal process conditions for cost minimization, the energy penalty associated with zeolite 13X is 206.30 kWh per ton of CO_2 captured, while for AHT the energy penalty was found to be 124.32 kWh per ton CO_2 captured (~66% decrease in the energetic cost). Though this study was carried out on a small database of materials, the general methodology could be readily extended to other industrial gas separation processes and expanded to include a greater number and diversity of zeolites. Future work such as would be of great practical value to the research community, and further demonstrates how fast and accurate force fields are playing a vital role in advancing the identification of application optimized materials.

A recent study carried out by Collins *et al.*[144] looked at optimizing the functional groups in 141 experimentally realized MOFs. In this study, a genetic algorithm was used to optimize the functional groups in these experimentally realized materials with respect to the CO_2 uptake, volumetric surface area and the parasitic energy. Using a set of 28 functional groups, a total search space of more than 1.64 trillion structures was generated and subsequently screened under conditions relevant to post-combustion CO_2 capture. Using empirical force fields to simulate the interactions between the MOF and CO_2, the adsorption properties of these materials were evaluated From this analysis, Collins *et al* identified 1035 derivatives of 23 different parent MOFs as having exceptional CO_2 uptakes (>3 mmol/g) at 0.15 atm and 298 K. Furthermore, the authors were able to identify functional groups that are associated with increased CO_2 uptakes or decreased parasitic energies. In using experimentally realized MOFs as the starting point for functionalization, the authors assert that the optimized materials should be synthetically realizable through pre- or post-synthetic modifications of the parent MOFs, and they have listed the structures along with their work. Using these structures as well as the identified functional groups, synthetic chemists can narrow down their search for carbon capture materials, and identify or design possible targets for synthesis.

Employing computational simulations, many screening studies beyond those presented here have been carried out for both MOFs and zeolites, investigating a diverse set of potential applications for these materials. As any of the studies referenced within this section will demonstrate, beyond the identification of the optimal materials for specific applications, screening studies provide us with new information about structure-property correlations. These relationships then allow us to predict and rationally design materials with properties that are optimized for a given application. The screening of vast numbers of materials required for such modeling has, for the

most part, been driven by well-developed force fields, such as those advanced by Calero *et al.*[141,145,146] In this context, then, the need for fast and accurate force fields is seemingly apparent.

1.5 Thesis Goals and Outline

With the growth in porous materials research, there is a growing need to be able to study materials that have a net charge balanced by counter-ions percolated throughout the pores. Previously, there were no charge derivation methods that were capable of deriving partial atomic charges directly from the QM ESP of charged periodic systems. Since electrostatic interactions are strong and long-ranged, having physically meaningful atomic charges is critical for obtaining an accurate representation of the guest-host interactions within a given system. In Chapter 3, two charge derivation methods presented are suited for in depth studies of charged periodic materials as well as high-throughput screening of a large number of materials, respectively The first of these methods allows for atomic charges to be derived directly from QM ESP of charged periodic frameworks. The second project presented in Chapter 3 focuses on the development of an extension to an empirical charge derivation method that was previously limited to charge neutral systems. The two charge derivation methods presented in Chapter 3 are suited for in depth studies of charged periodic systems as well as high-throughput screening of a large number of materials, respectively. As evidenced through the literature, high-throughput screening is an important part of computational chemistry's contribution to materials research, as it allows for the identification of key chemical and structural properties that guide synthetic chemists in designing high-performing materials geared towards specific applications.

In the second part of this thesis, a full empirical force field is developed for siliceous and aluminum containing zeolites. In the process of performing a thorough literature search, we observed a large variation in the atomic charges assigned to the atoms in zeolites, as well as a

multitude of different force field parameters. The use of these so-called generic charges, as well as the variation observed amongst the different generic charge sets available, are in part due to the fact that previously there were no methods to obtain atomic charges from the QM ESP for charged frameworks. Although generic charges are less computationally demanding during simulations, they has the draw-back of ignoring any structural differences among a given material, and it has been shown that accounting for these differences can result in improved quantitative accuracy.[140] Taking advantage of the work presented in Chapter 3, the final project in Chapter 4 focuses on the development of a force field using physically relevant charges, derived from the periodic QM ESP of charged aluminosilicate zeolites, for the first time. The force field presented has been optimized for the most important gaseous separations employing zeolites, involving CO_2, N_2, and CH_4. The final chapter of this thesis, Chapter 5, will summarize the results of the different projects presented, while also proposing future directions for research related to this work.

1.6 References

(1) Flanigen, E. M. *Chapter 2 Zeolites and Molecular Sieves: An Historical Perspective*, Studies in.; H. van Bekkum, E.M. Flanigen, P. A. J. and J. C. J., Ed.; Elsevier, 2001.

(2) Ackley, M. W.; Rege, S. U.; Saxena, H. Application of Natural Zeolites in the Purification and Separation of Gases. *Microporous Mesoporous Mater.* **2003**, *61* (1–3), 25–42.

(3) Lee, J.; Farha, O. K.; Roberts, J.; Scheidt, K. A.; Nguyen, S. T.; Hupp, J. T. Metal–organic Framework Materials as Catalysts. *Chem. Soc. Rev.* **2009**, *38* (5), 1450.

(4) Corma, A.; García, H.; Llabrés i Xamena, F. X. Engineering Metal Organic Frameworks for Heterogeneous Catalysis. *Chem. Rev.* **2010**, *110* (8), 4606–4655.

(5) Huang, Y.-B.; Liang, J.; Wang, X.-S.; Cao, R. Multifunctional Metal–organic Framework Catalysts: Synergistic Catalysis and Tandem Reactions. *Chem. Soc. Rev.* **2017**, *46* (1), 126–157.

(6) Li, J.-R.; Kuppler, R. J.; Zhou, H.-C. Selective Gas Adsorption and Separation in Metal–organic Frameworks. *Chem. Soc. Rev.* **2009**, *38* (5), 1477.

(7) Farha, O. K.; Özgür Yazaydın, A.; Eryazici, I.; Malliakas, C. D.; Hauser, B. G.; Kanatzidis, M. G.; Nguyen, S. T.; Snurr, R. Q.; Hupp, J. T. De Novo Synthesis of a Metal–organic Framework Material Featuring Ultrahigh Surface Area and Gas Storage Capacities. *Nat. Chem.* **2010**, *2* (11), 944–948.

(8) Shevlin, S. Molecular Separation: Flexing with the Flow. *Nat. Mater.* **2017**, *16* (8), 785–785.

(9) Xie, Z.; Ma, L.; deKrafft, K. E.; Jin, A.; Lin, W. Porous Phosphorescent Coordination Polymers for Oxygen Sensing. *J. Am. Chem. Soc.* **2010**, *132* (3), 922–923.

(10) Kreno, L. E.; Leong, K.; Farha, O. K.; Allendorf, M.; Van Duyne, R. P.; Hupp, J. T. Metal–Organic Framework Materials as Chemical Sensors. *Chem. Rev.* **2012**, *112* (2), 1105–1125.

(11) Stassen, I.; Bueken, B.; Reinsch, H.; Oudenhoven, J. F. M.; Wouters, D.; Hajek, J.; Van Speybroeck, V.; Stock, N.; Vereecken, P. M.; Van Schaijk, R.; De Vos, D.; Ameloot, R. Towards Metal–organic Framework Based Field Effect Chemical Sensors: UiO-66-NH $_2$ for Nerve Agent Detection. *Chem. Sci.* **2016**, *7* (9), 5827–5832.

(12) Quah, H. S.; Chen, W.; Schreyer, M. K.; Yang, H.; Wong, M. W.; Ji, W.; Vittal, J. J. Multiphoton Harvesting Metal–organic Frameworks. *Nat. Commun.* **2015**, *6*, 7954.

(13) Lin, R.-B.; Liu, S.-Y.; Ye, J.-W.; Li, X.-Y.; Zhang, J.-P. Photoluminescent Metal-Organic Frameworks for Gas Sensing. *Adv. Sci.* **2016**, *3* (7), 1500434.

(14) Chen, L.; Ye, J.-W.; Wang, H.-P.; Pan, M.; Yin, S.-Y.; Wei, Z.-W.; Zhang, L.-Y.; Wu, K.; Fan, Y.-N.; Su, C.-Y. Ultrafast Water Sensing and Thermal Imaging by a Metal-Organic Framework with Switchable Luminescence. *Nat. Commun.* **2017**, *8*, 15985.

(15) Givaja, G.; Amo-Ochoa, P.; Gómez-García, C. J.; Zamora, F. Electrical Conductive Coordination Polymers. *Chem. Soc. Rev.* **2012**, *41* (1), 115–147.

(16) Li, W.-J.; Liu, J.; Sun, Z.-H.; Liu, T.-F.; Lü, J.; Gao, S.-Y.; He, C.; Cao, R.; Luo, J.-H. Integration of Metal-Organic Frameworks into an Electrochemical Dielectric Thin Film for Electronic Applications. *Nat. Commun.* **2016**, *7*, 11830.

(17) Horcajada, P.; Gref, R.; Baati, T.; Allan, P. K.; Maurin, G.; Couvreur, P.; Férey, G.; Morris, R. E.; Serre, C. Metal–Organic Frameworks in Biomedicine. *Chem. Rev.* **2012**, *112* (2), 1232–1268.

(18) Ibrahim, M.; Sabouni, R.; Husseini, G. Anti-Cancer Drug Delivery Using Metal Organic Frameworks (MOFs). *Curr. Med. Chem.* **2017**, *24* (2), 193–214.

(19) Cronstedt, A. F.; Handl., A. *No Title*; Stockholm, 1756.

(20) Bekkum, H. van. *Introduction to Zeolite Science and Practice*; Elsevier, 2001.

(21) History of the Database of Zeolite Structures http://www.iza-structure.org/DatabaseHistory.htm (accessed Dec 6, 2017).

(22) (IZA-SC), S. C. of the I. Z. A. Database of Zeolite Structures http://www.iza-structure.org/databases/ (accessed May 13, 2016).

(23) Deem, M. W.; Newsam, J. M. Determination of 4-Connected Framework Crystal Structures by Simulated Annealing. *Nature* **1989**, *342* (6247), 260–262.

(24) Mellot Draznieks, C.; Newsam, J. M.; Gorman, A. M.; Freeman, C. M.; Férey, G. De Novo Prediction of Inorganic Structures Developed through Automated Assembly of Secondary Building Units (AASBU Method). *Angew. Chemie Int. Ed.* **2000**, *39* (13), 2270–2275.

(25) Yi Li, †; Jihong Yu, *,†; Donghan Liu, †; Wenfu Yan, †; Ruren Xu, *,† and; Xu‡, Y. Design of Zeolite Frameworks with Defined Pore Geometry through Constrained Assembly of Atoms. **2003**.

(26) Woodley, S. M.; Catlow, C. R. A.; Battle, P. D.; Gale, J. D. The Prediction of Inorganic Crystal

Framework Structures Using Excluded Regions within a Genetic Algorithm Approach. *Chem. Commun.* **2004**, *0* (1), 22.

(27) David J. Earl†, ‡ and; Michael W. Deem*, †. Toward a Database of Hypothetical Zeolite Structures. **2006**.

(28) Pophale, R.; Cheeseman, P. A.; Deem, M. W. A Database of New Zeolite-like Materials. *Phys. Chem. Chem. Phys.* **2011**, *13* (27), 12407.

(29) Li, Y.; Li, X.; Liu, J.; Duan, F.; Yu, J. In Silico Prediction and Screening of Modular Crystal Structures via a High-Throughput Genomic Approach. *Nat. Commun.* **2015**, *6*, 8328.

(30) Li, Y.; Yu, J. Genetic Engineering of Inorganic Functional Modular Materials. *Chem. Sci.* **2016**, *7* (6), 3472–3481.

(31) McCusker, L. B.; Baerlocher, C. *Zeolite Structures*; H. van Bekkum, E.M. Flanigen, P. A. J. and J. C. J., Ed.; Elsevier, 2001; Vol. Volume 137.

(32) Flanigen, E. M.; Khatami, H.; Szymanksi, H. A. Infrared Structural Studies of Zeolite Frameworks; 1974; pp 201–229.

(33) Czjzek, M.; Jobic, H.; Fitch, A. N.; Vogt, T. Direct Determination of Proton Positions in D-Y and H-Y Zeolite Samples by Neutron Powder Diffraction. *J. Phys. Chem.* **1992**, *96* (4), 1535–1540.

(34) Smith, J. V. Faujasite-Type Structures: Aluminosilicate Framework: Positions of Cations and Molecules: Nomenclature; 1974; pp 171–200.

(35) Borade, R.; Sayari, A.; Adnot, A.; Kaliaguine, S. Characterization of Acidity in ZSM-5 Zeolites: An X-Ray Photoelectron and IR Spectroscopy Study. *J. Phys. Chem.* **1990**, *94* (15), 5989–5994.

(36) Noroozi, R.; Al-Musawi, T. J.; Kazemian, H.; Kalhori, E. M.; Zarrabi, M. Removal of Cyanide Using Surface-Modified Linde Type-A Zeolite Nanoparticles as an Efficient and Eco-Friendly Material. *J. Water Process Eng.* **2018**, *21*, 44–51.

(37) Löwenstein, W. The Distribution of Aluminum in the Tetrahedra of Silicates and Aluminates. *Am. Mineral.* **1954**, *39*, 92–96.

(38) Pluth, J. J.; Smith, J. V. Accurate Redetermination of Crystal Structure of Dehydrated Zeolite A. Absence of near Zero Coordination of Sodium. Refinement of Silicon,aluminum-Ordered Superstructure. *J. Am. Chem. Soc.* **1980**, *102* (14), 4704–4708.

(39) Bell, R. G.; Jackson, R. A.; Catlow, C. R. A. Löwenstein's Rule in Zeolite A: A Computational Study. *Zeolites* **1992**, *12* (7), 870–871.

(40) Tarling, S. E.; Barnes, P.; Klinowski, J. The Structure and Si,Al Distribution of the Ultramarines. *Acta Crystallogr. Sect. B* **1988**, *44* (2), 128–135.

(41) Fletcher, R. E.; Ling, S.; Slater, B. Violations of Löwenstein's Rule in Zeolites. *Chem. Sci.* **2017**, *8* (11), 7483–7491.

(42) Furukawa, H.; Cordova, K. E.; O'Keeffe, M.; Yaghi, O. M. The Chemistry and Applications of Metal-Organic Frameworks. *Science.* 2013, p 974.

(43) Yaghi, O. M.; Li, H.; Eddaoudi, M.; O'Keeffe, M. Design and Synthesis of an Exceptionally Stable and Highly Porous Metal-Organic Framework. *Nature* **1999**, *402* (6759), 276–279.

(44) Kondo, M.; Yoshitomi, T.; Seki, K.; Matsuzaka, H.; Kitagawa, S. Three-Dimensional Framework with Channeling Cavities for Small Molecules: {[M2(4,4'-bpy)3(NO3) 4]·xH2O}n (M = Co, Ni, Zn). **1997**, *36* (16), 1725–1727.

(45) Kondo, M.; Asami, A.; Fujimoto, K.; Noro, S.; Kitagawa, S.; Ishii, T.; Matsuzaka, H. New Coordination Network of [Cd2(bpob)3(NO3)4]n (bpob=1,4-bis(4- Pyridoxy)benzene) Constructed from Two Structural Isomers of the Ligand. *Int. J. Inorg. Mater.* **1999**, *1* (1), 73–75.

(46) Dietzel, P. D. C.; Blom, R.; Fjellvåg, H. Base-Induced Formation of Two Magnesium Metal-Organic Framework Compounds with a Bifunctional Tetratopic Ligand. *Eur. J. Inorg. Chem.* **2008**, *2008* (23), 3624–3632.

(47) Britt, D.; Furukawa, H.; Wang, B.; Glover, T. G.; Yaghi, O. M. Highly Efficient Separation of Carbon Dioxide by a Metal-Organic Framework Replete with Open Metal Sites. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (49), 20637–20640.

(48) Schoenecker, P. M.; Carson, C. G.; Jasuja, H.; Flemming, C. J. J.; Walton, K. S. Effect of Water Adsorption on Retention of Structure and Surface Area of Metal–Organic Frameworks. *Ind. Eng. Chem. Res.* **2012**, *51* (18), 6513–6519.

(49) Burtch, N. C.; Jasuja, H.; Walton, K. S. Water Stability and Adsorption in Metal–Organic Frameworks. *Chem. Rev.* **2014**, *114* (20), 10575–10612.

(50) Furukawa, H.; Gándara, F.; Zhang, Y.-B.; Jiang, J.; Queen, W. L.; Hudson, M. R.; Yaghi, O. M. Water Adsorption in Porous Metal–Organic Frameworks and Related Materials. *J. Am. Chem. Soc.* **2014**, *136* (11), 4369–4381.

(51) Taylor, J. M.; Vaidhyanathan, R.; Iremonger, S. S.; Shimizu, G. K. H. Enhancing Water Stability of Metal−Organic Frameworks via Phosphonate Monoester Linkers.

(52) Hamon, L.; Serre, C.; Devic, T.; Loiseau, T.; Millange, F.; Férey, G.; Weireld, G. De. Comparative Study of Hydrogen Sulfide Adsorption in the MIL-53(Al, Cr, Fe), MIL-47(V), MIL-100(Cr), and MIL-101(Cr) Metal−Organic Frameworks at Room Temperature. *J. Am. Chem. Soc.* **2009**, *131* (25), 8775–8777.

(53) Ding, L.; Yazaydin, A. Ö. How Well Do Metal–Organic Frameworks Tolerate Flue Gas Impurities? *J. Phys. Chem. C* **2012**, *116* (43), 22987–22991.

(54) Mangano, E.; Kahr, J.; Wright, P. A.; Brandani, S. Accelerated Degradation of MOFs under Flue Gas Conditions. *Faraday Discuss.* **2016**, *192* (0), 181–195.

(55) Tranchemontagne, D. J.; Hunt, J. R.; Yaghi, O. M. Room Temperature Synthesis of Metal-Organic Frameworks: MOF-5, MOF-74, MOF-177, MOF-199, and IRMOF-0. *Tetrahedron* **2008**, *64* (36), 8553–8557.

(56) Dey, C.; Kundu, T.; Biswal, B. P.; Mallick, A.; Banerjee, R. Crystalline Metal-Organic Frameworks (MOFs): Synthesis, Structure and Function. *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater.* **2014**, *70* (1), 3–10.

(57) Wang, C.-C.; Ying, J. Y. Sol−Gel Synthesis and Hydrothermal Processing of Anatase and Rutile Titania Nanocrystals. *Chem. Mater.* **1999**, *11* (11), 3113–3120.

(58) Stock, N.; Biswas, S. Synthesis of Metal-Organic Frameworks (MOFs): Routes to Various MOF Topologies, Morphologies, and Composites. *Chem. Rev.* **2012**, *112* (2), 933–969.

(59) Hu, Q.; Yu, J.; Liu, M.; Liu, A.; Dou, Z.; Yang, Y. A Low Cytotoxic Cationic Metal–Organic Framework Carrier for Controllable Drug Release. *J. Med. Chem.* **2014**, *57* (13), 5679–5685.

(60) Xu, H.; Cao, C.-S.; Kang, X.-M.; Zhao, B. Lanthanide-Based Metal–organic Frameworks as Luminescent Probes. *Dalt. Trans.* **2016**, *45* (45), 18003–18017.

(61) Farha, O. K.; Eryazici, I.; Jeong, N. C.; Hauser, B. G.; Wilmer, C. E.; Sarjeant, A. A.; Snurr, R. Q.; Nguyen, S. T.; Yazaydın, A. Ö.; Hupp, J. T. Metal–Organic Framework Materials with Ultrahigh Surface Areas: Is the Sky the Limit? *J. Am. Chem. Soc.* **2012**, *134* (36), 15016–15021.

(62) Wang, T. C.; Bury, W.; Gómez-Gualdrón, D. A.; Vermeulen, N. A.; Mondloch, J. E.; Deria, P.; Zhang, K.; Moghadam, P. Z.; Sarjeant, A. A.; Snurr, R. Q.; Stoddart, J. F.; Hupp, J. T.; Farha, O. K. Ultrahigh Surface Area Zirconium MOFs and Insights into the Applicability of the BET Theory. *J. Am. Chem. Soc.* **2015**, *137* (10), 3585–3591.

(63) Schneemann, A.; Bon, V.; Schwedler, I.; Senkovska, I.; Kaskel, S.; Fischer, R. A. Flexible Metal–organic Frameworks. *Chem. Soc. Rev.* **2014**, *43* (16), 6062–6096.

(64) Breck, D. W.; Eversole, W. G.; Milton, R. M.; Reed, T. B.; Thomas, T. L. Crystalline Zeolites. I. The Properties of a New Synthetic Zeolite, Type A. *J. Am. Chem. Soc.* **1956**, *78* (23), 5963–5972.

(65) Davis, S.; Y, I. *Chemical Economics Handbook: Zeolites*; 2009.

(66) Yilmaz, B.; Müller, U. Catalytic Applications of Zeolites in Chemical Industry. *Top. Catal.* **2009**, *52* (6–7), 888–895.

(67) Eddaoudi, M.; Kim, J.; Rosi, N.; Vodak, D.; Wachter, J.; O'Keeffe, M.; Yaghi, O. M. Systematic Design of Pore Size and Functionality in Isoreticular MOFs and Their Application in Methane Storage. *Science* **2002**, *295* (5554), 469–472.

(68) Wu, H.; Zhou, W.; Yildirim, T. High-Capacity Methane Storage in Metal−Organic Frameworks M_2(Dhtp): The Important Role of Open Metal Sites. *J. Am. Chem. Soc.* **2009**, *131* (13), 4995–5000.

(69) Murray, L. J.; Dincă, M.; Long, J. R. Hydrogen Storage in Metal-Organic Frameworks. *Chem. Soc. Rev.* **2009**, *38* (5), 1294–1314.

(70) Liu, J.; Chen, L.; Cui, H.; Zhang, J.; Zhang, L.; Su, C.-Y. Applications of Metal-Organic Frameworks in Heterogeneous Supramolecular Catalysis. *Chem. Soc. Rev.* **2014**, *43* (16), 6011–6061.

(71) Nath, I.; Chakraborty, J.; Verpoort, F. Metal Organic Frameworks Mimicking Natural Enzymes: A Structural and Functional Analogy. *Chem. Soc. Rev.* **2016**, *45* (15), 4127–4170.

(72) Yu, J.; Cui, Y.; Wu, C.-D.; Yang, Y.; Chen, B.; Qian, G. Two-Photon Responsive Metal–Organic Framework. *J. Am. Chem. Soc.* **2015**, *137* (12), 4026–4029.

(73) Medishetty, R.; Park, I.-H.; Lee, S. S.; Vittal, J. J. Solid-State Polymerisation via [2+2] Cycloaddition Reaction Involving Coordination Polymers. *Chem. Commun.* **2016**, *52* (21), 3989–4001.

(74) Na, K.; Choi, K. M.; Yaghi, O. M.; Somorjai, G. A. Metal Nanocrystals Embedded in Single Nanocrystals of MOFs Give Unusual Selectivity as Heterogeneous Catalysts. *Nano Lett.* **2014**, *14* (10), 5979–5983.

(75) Ou, S.; Wu, C.-D. Rational Construction of Metal–organic Frameworks for Heterogeneous Catalysis. *Inorg. Chem. Front.* **2014**, *1* (10), 721–734.

(76) Aguirre-Díaz, L. M.; Gándara, F.; Iglesias, M.; Snejko, N.; Gutiérrez-Puebla, E.; Monge, M. Á. Tunable Catalytic Activity of Solid Solution Metal–Organic Frameworks in One-Pot Multicomponent Reactions. *J. Am. Chem. Soc.* **2015**, *137* (19), 6132–6135.

(77) Dhakshinamoorthy, A.; Asiri, A. M.; Garcia, H. Tuneable Nature of Metal Organic Frameworks as Heterogeneous Solid Catalysts for Alcohol Oxidation. *Chem. Commun.* **2017**, *53* (79), 10851–10869.

(78) Chuan-De Wu; Aiguo Hu; Lin Zhang, and; Lin*, W. A Homochiral Porous Metal−Organic Framework for Highly Enantioselective Heterogeneous Asymmetric Catalysis. **2005**.

(79) Kitagawa, S.; Kitaura, R.; Noro, S. Functional Porous Coordination Polymers. *Angew. Chemie Int. Ed.* **2004**, *43* (18), 2334–2375.

(80) Vega, L. F.; Bahamon, D. Comparative Study of MOFs and Zeolites For CO₂ Capture and Separation at Process Conditions. In *Abu Dhabi International Petroleum Exhibition & Conference*; Society of Petroleum Engineers, 2016.

(81) Faust, T. MOFs Move to Market. *Nat. Publ. Gr.* **2016**, *8*.

(82) Hendon, C. H.; Rieth, A. J.; Korzyński, M. D.; Dincă, M. Grand Challenges and Future Opportunities for Metal-Organic Frameworks. *ACS Cent. Sci.* **2017**, *3* (6), 554–563.

(83) Parmesan, C.; Yohe, G. A Globally Coherent Fingerprint of Climate Change Impacts across Natural Systems. *Nature* **2003**, *421* (6918), 37–42.

(84) Oreskes, N. Beyond the Ivory Tower. The Scientific Consensus on Climate Change. *Science* **2004**, *306* (5702), 1686.

(85) Harsch, M. A.; Phillips, A.; Zhou, Y.; Leung, M.-R.; Rinnan, D. S.; Kot, M. Moving Forward: Insights and Applications of Moving-Habitat Models for Climate Change Ecology. *J. Ecol.* **2017**, *105* (5), 1169–1181.

(86) Ramanathan, V.; Feng, Y. On Avoiding Dangerous Anthropogenic Interference with the Climate System: Formidable Challenges Ahead. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (38), 14245–14250.

(87) Solomon, S.; Plattner, G.-K.; Knutti, R.; Friedlingstein, P. Irreversible Climate Change due to Carbon Dioxide Emissions. *Proc. Natl. Acad. Sci.* **2009**, *106* (6), 1704–1709.

(88) Aalst, M.; Adger, N.; Arent, D.; Barnett, K.; Betts, R.; Bilir, E.; Birkmann, J.; Carmin, J.; Chadee, D.; Challinor, A. *Climate Change 2014 – Impacts, Adaptation and Vulnerability: Regional Aspects - Intergovernmental Panel on Climate Change - Google Books*; IPCC, 2014.

(89) Environment - U.S. Energy Information Administration (EIA) https://www.eia.gov/environment/ (accessed Dec 29, 2017).

(90) U.S. Environmental Protection Agency. Sources of Greenhouse Gas Emissions https://www.epa.gov/ghgemissions/sources-greenhouse-gas-emissions (accessed Dec 20, 2017).

(91) Stuart Haszeldine, R. Carbon Capture and Storage: How Green Can Black Be? *Science (80-.).* **2009**, *325* (5948), 1647–1652.

(92) Aydin, G.; Karakurt, I.; Aydiner, K. Evaluation of Geologic Storage Options of CO2: Applicability, Cost, Storage Capacity and Safety. *Energy Policy* **2010**, *38* (9), 5072–5080.

(93) Barker, R.; Hua, Y.; Neville, A. Internal Corrosion of Carbon Steel Pipelines for Dense-Phase CO2

Transport in Carbon Capture and Storage (CCS) – a Review. *Int. Mater. Rev.* **2017**, *62* (1), 1–31.

(94) Aycaguer, A. C.; Lev-On, M.; Winer, A. M. Reducing Carbon Dioxide Emissions with Enhanced Oil Recovery Projects: A Life Cycle Assessment Approach. *Energy and Fuels* **2001**, *15* (2), 303–308.

(95) Alvarado, V.; Manrique, E. Enhanced Oil Recovery: An Update Review. *Energies* **2010**, *3* (9), 1529–1575.

(96) Takht Ravanchi, M.; Sahebdelfar, S. Carbon Dioxide Capture and Utilization in Petrochemical Industry: Potentials and Challenges. *Appl. Petrochemical Res.* **2014**, *4* (1), 63–77.

(97) Avila, A. M.; Yang, F.; Shi, M.; Kuznicki, S. M. Extraction of Ethane from Natural Gas at High Pressure by Adsorption on Na-ETS-10. *Chem. Eng. Sci.* **2011**, *66* (13), 2991–2996.

(98) Arefi Pour, A.; Sharifnia, S.; Neishabori Salehi, R.; Ghodrati, M. Adsorption Separation of CO2/CH4 on the Synthesized NaA Zeolite Shaped with Montmorillonite Clay in Natural Gas Purification Process. *J. Nat. Gas Sci. Eng.* **2016**, *36*, 630–643.

(99) Su, F.; Lu, C. CO2 Capture from Gas Stream by Zeolite 13X Using a Dual-Column Temperature/vacuum Swing Adsorption. *Energy Environ. Sci.* **2012**, *5* (10), 9021.

(100) Chen, S. J.; Tao, Z. C.; Fu, Y.; Zhu, M.; Li, W. L.; Li, X. D. CO2 Separation from Offshore Natural Gas in Quiescent and Flowing States Using 13X Zeolite. *Appl. Energy* **2017**, *205*, 1435–1446.

(101) Ojuva, A. Processing and Performance of Zeolites for Efficient Carbon Dioxide Separation.

(102) Baboolal, J. D. Phd Thesis, University of Alberta, 2015.

(103) Sircar, S.; Golden, T. C. Purification of Hydrogen by Pressure Swing Adsorption. *Sep. Sci. Technol.* **2000**, *35* (5), 667–687.

(104) Delgado, J. A.; Agueda, V. I.; Uguina, M. A.; Sotelo, J. L.; Brea, P. Hydrogen Recovery from off-Gases with Nitrogen-Rich Impurity by Pressure Swing Adsorption Using CaX and 5A Zeolites. *Adsorption* **2015**, *21* (1–2), 107–123.

(105) Smith, A. .; Klosek, J. A Review of Air Separation Technologies and Their Integration with Energy Conversion Processes. *Fuel Process. Technol.* **2001**, *70* (2), 115–134.

(106) Zhaohui Yang, †; Y. S. Lin, *,† and; Zeng‡, Y. High-Temperature Sorption Process for Air Separation and Oxygen Removal. **2002**.

(107) Tian, C.; Fu, Q.; Ding, Z.; Han, Z.; Zhang, D. Experiment and Simulation Study of a Dual-Reflux Pressure Swing Adsorption Process for Separating N2/O2. *Sep. Purif. Technol.* **2017**, *189*, 54–65.

(108) Koh, D.-Y.; Pimentel, B. R.; Babu, V. P.; Stephenson, N.; Chai, S. W.; Rosinski, A.; Lively, R. P. Sub-Ambient Air Separation via Li+ Exchanged Zeolite. *Microporous Mesoporous Mater.* **2018**, *256*, 140–146.

(109) Riedmann, R. A.; Purtschert, R. Separation of Argon from Environmental Samples for Ar-37 and Ar-39 Analyses. *Sep. Purif. Technol.* **2016**, *170*, 217–223.

(110) Suh, M. P.; Park, H. J.; Prasad, T. K.; Lim, D.-W. Hydrogen Storage in Metal–Organic Frameworks. *Chem. Rev.* **2012**, *112* (2), 782–835.

(111) Kitagawa, S. Porous Materials and the Age of Gas. *Angew. Chemie Int. Ed.* **2015**, *54* (37), 10686–10687.

(112) and, Y. L.; Yang*, R. T. Gas Adsorption and Storage in Metal−Organic Framework MOF-177. **2007**.

(113) Gándara, F.; Furukawa, H.; Lee, S.; Yaghi, O. M. High Methane Storage Capacity in Aluminum Metal-Organic Frameworks. *J. Am. Chem. Soc.* **2014**, *136* (14), 5271–5274.

(114) Zhang, Z.; Yao, Z.-Z.; Xiang, S.; Chen, B. Perspective of Microporous Metal–organic Frameworks for CO 2 Capture and Separation. *Energy Environ. Sci.* **2014**, *7* (9), 2868.

(115) Treacy, M. M. J.; Randall, K. H.; Rao, S.; Perry, J. A.; Chadi, D. J. Enumeration of Periodic Tetrahedral Frameworks. *Zeitschrift für Krist. - Cryst. Mater.* **1997**, *212* (11), 768–791.

(116) Li, Y.; Yu, J. New Stories of Zeolite Structures: Their Descriptions, Determinations, Predictions, and Evaluations. *Chem. Rev.* **2014**, *114* (14), 7268–7316.

(117) Friedrichs, O. D.; Dress, A. W. M.; Huson, D. H.; Klinowski, J.; Mackay, A. L. Systematic Enumeration of Crystalline Networks. *Nature* **1999**, *400* (6745), 644–647.

(118) Foster, M. D.; Friedrichs, O. D.; Bell, R. G.; Paz, F. A. A.; Klinowski, J. Chemical Evaluation of Hypothetical Uninodal Zeolites. *J. Am. Chem. Soc.* **2004**, *126* (31), 9769–9775.

(119) Foster, M. D.; Simperler, A.; Bell, R. G.; Friedrichs, O. D.; Paz, F. A. A.; Klinowski, J. Chemically Feasible Hypothetical Crystalline Networks. *Nat. Mater.* **2004**, *3* (4), 234–238.

(120) Simperler, A.; Foster, M. D.; Friedrichs, O. D.; Bell, R. G.; Almeida Paz, F. A.; Klinowski, J. Hypothetical Binodal Zeolitic Frameworks. *Acta Cryst* **2005**, *61*, 263–279.

(121) Brouwer, D. H.; Horvath, M. A Simulated Annealing Approach for Solving Zeolite Crystal Structures from Two-Dimensional NMR Correlation Spectra. *Solid State Nucl. Magn. Reson.* **2015**, *65*, 89–98.

(122) García-Pérez, E.; Parra, J. B.; Ania, C. O.; García-Sánchez, a.; van Baten, J. M.; Krishna, R.; Dubbeldam, D.; Calero, S. A Computational Study of CO2, N2, and CH4 Adsorption in Zeolites. *Adsorption* **2007**, *13* (5–6), 469–476.

(123) Bodart, P.; Nagy, J. B.; Debras, G.; Gabelica, Z.; Jacobs, P. A. Aluminum Siting in Mordenite and Dealumination Mechanism. *J. Phys. Chem.* **1986**, *90* (21), 5183–5190.

(124) Tamás I. Korányi*, †,‡ and; Nagy‡, J. B. Distribution of Aluminum in Different Periodical Building Units of MOR and BEA Zeolites. **2005**.

(125) Kim, J.; Lin, L.-C.; Swisher, J. A.; Haranczyk, M.; Smit, B. Predicting Large CO $_2$ Adsorption in Aluminosilicate Zeolites for Postcombustion Carbon Dioxide Capture. *J. Am. Chem. Soc.* **2012**, *134* (46), 18940–18943.

(126) Boyd, P. G.; Woo, T. K. A Generalized Method for Constructing Hypothetical Nanoporous Materials of Any Net Topology from Graph Theory. *CrystEngComm* **2016**, *18* (21), 3777–3792.

(127) Wilmer, C. E.; Leaf, M.; Lee, C. Y.; Farha, O. K.; Hauser, B. G.; Hupp, J. T.; Snurr, R. Q. Large-Scale Screening of Hypothetical Metal–organic Frameworks. *Nat. Chem.* **2012**, *4* (2), 83–89.

(128) Martin, R. L.; Simon, C. M.; Smit, B.; Haranczyk, M. *In Silico* Design of Porous Polymer Networks: High-Throughput Screening for Methane Storage Materials. *J. Am. Chem. Soc.* **2014**, *136* (13), 5006–5022.

(129) Chung, Y. G.; Camp, J.; Haranczyk, M.; Sikora, B. J.; Bury, W.; Krungleviciute, V.; Yildirim, T.; Farha, O. K.; Sholl, D. S.; Snurr, R. Q. Computation-Ready, Experimental Metal–Organic Frameworks: A

Tool To Enable High-Throughput Screening of Nanoporous Crystals. *Chem. Mater.* **2014**, *26* (21), 6185–6192.

(130) Lin, L.; Berger, A. H.; Martin, R. L.; Kim, J.; Swisher, J. A.; Jariwala, K.; Rycroft, C. H.; Bhown, A. S.; Deem, M. W.; Haranczyk, M. In Silico Screening of Carbon-Capture Materials. *Nat. Mater.* **2012**, *11* (7), 633–641.

(131) Wilmer, C. E.; Snurr, R. Q. Large-Scale Generation and Screening of Hypothetical Metal-Organic Frameworks for Applications in Gas Storage and Separations; Springer, Cham, 2013; pp 257–289.

(132) Fernandez, M.; Trefiak, N. R.; Woo, T. K. Atomic Property Weighted Radial Distribution Functions Descriptors of Metal–Organic Frameworks for the Prediction of Gas Uptake Capacity. *J. Phys. Chem. C* **2013**, *117* (27), 14095–14105.

(133) Düren, T.; Bae, Y.-S.; Snurr, R. Q. Using Molecular Simulation to Characterise Metal–organic Frameworks for Adsorption Applications. *Chem. Soc. Rev.* **2009**, *38* (5), 1237.

(134) Ramsahye, N. a; Maurin, G.; Bourrelly, S.; Llewellyn, P.; Loiseau, T.; Ferey, G. Charge Distribution in Metal Organic Framework Materials: Transferability to a Preliminary Molecular Simulation Study of the CO(2) Adsorption in the MIL-53 (Al) System. *Phys. Chem. Chem. Phys.* **2007**, *9* (9), 1059–1063.

(135) Xiong, R.; Keffer, D. J.; Fuentes-Cabrera, M.; Nicholson, D. M.; Michalkova, A.; Petrova, T.; Leszczynski, J.; Odbadrakh, K.; Doss, B. L.; Lewis, J. P. Effect of Charge Distribution on RDX Adsorption in IRMOF-10. *Langmuir* **2010**, *26* (8), 5942–5950.

(136) Chen, Y.; Jiang, J. A Bio-Metal-Organic Framework for Highly Selective CO2 Capture: A Molecular Simulation Study. *ChemSusChem* **2010**, *3* (8), 982–988.

(137) Grosch, J. S.; Paesani, F. Molecular-Level Characterization of the Breathing Behavior of the Jungle-Gym-Type DMOF-1 Metal–Organic Framework. *J. Am. Chem. Soc.* **2012**, *134* (9), 4207–4215.

(138) Jackson, R. A.; Catlow, C. R. A. Computer Simulation Studies of Zeolite Structure. *Mol. Simul.* **1988**, *1* (4), 207–224.

(139) Zhang, K.; Zhang, L.; Jiang, J. Adsorption of C_1–C_4 Alcohols in Zeolitic Imidazolate Framework-8: Effects of Force Fields, Atomic Charges, and Framework Flexibility. *J. Phys. Chem. C* **2013**, *117* (48), 25628–25635.

(140) Fischer, M.; Bell, R. G. Modeling CO 2 Adsorption in Zeolites Using DFT-Derived Charges : Comparing System-Speci Fi c and Generic Models. *J. Phys. Chem. C* **2013**, *117* (46), 24446–24454.

(141) García-Sánchez, A.; Ania, C. O.; Parra, J. B.; Dubbeldam, D.; Vlugt, T. J. H.; Krishna, R.; Calero, S. Transferable Force Field for Carbon Dioxide Adsorption in Zeolites. *J. Phys. Chem. C* **2009**, *113* (20), 8814–8820.

(142) Hasan, M. M. F.; First, E. L.; Floudas, C. A. Cost-Effective CO2 Capture Based on in Silico Screening of Zeolites and Process Optimization. *Phys. Chem. Chem. Phys.* **2013**, *15* (40), 17601.

(143) First, E. L.; Gournais, E. C.; Wei, J.; Floudas, C. A. ZEOMICS http://helios.princeton.edu/zeomics/ (accessed Oct 5, 2018).

(144) Collins, S. P.; Daff, T. D.; Piotrkowski, S. S.; Woo, T. K. Materials Design by Evolutionary Optimization of Functional Groups in Metal-Organic Frameworks. *Sci. Adv.* **2016**, *2* (11), e1600954–e1600954.

(145) Sofía Calero, *,†,‡; David Dubbeldam, †; Rajamani Krishna, †; Berend Smit, †; Thijs J. H. Vlugt, §; Joeri F. M. Denayer, ∥; Johan A. Martens, ⊥ and; Maesen#, T. L. M. Understanding the Role of Sodium during Adsorption: A Force Field for Alkanes in Sodium-Exchanged Faujasites. **2004**.

(146) Martin-Calvo, A.; Gutiérrez-Sevillano, J. J.; Parra, J. B.; Ania, C. O.; Calero, S. Transferable Force Fields for Adsorption of Small Gases in Zeolites. *Phys. Chem. Chem. Phys.* **2015**, *17* (37), 24048–24055.

2 Methods

2.1 Summary

In this chapter, all of the major components involved in the optimization of empirical force field potentials for aluminosilicate zeolites will be reviewed. All of the methods that will be discussed in this chapter are used throughout the entirety of the research presented in this thesis. It should be noted, however, that this chapter does not contain an exhaustive list of all the methods and theories used in the projects presented in Chapter's 3 and 4. As aluminosilicate zeolites are the main focus of the research presented in this thesis, many of the methods presented in this section will be discussed primarily in the context of these material. However, these methods are applicable to a wide range of both molecular and periodic systems.

2.2 Fully Automated Adsorption Analysis in Porous Solids (FA^{3}PS)

FA^{3}PS is a software package that was developed in the Woo lab in order to automate the calculations necessary for simulating gas uptake and other molecular-level data in porous materials, such as MOFs, COFs, and zeolites. Many steps are required in order to compute gas uptake isotherms and other gas sorption properties, and each step in the process requires specialized inputs containing the instructions needed to run a calculation. In essence, the FA^{3}PS software package will generate all the required inputs and streamline the different calculations that must be performed with minimal input from the user. A schematic for the FA^{3}PS code is provided in Figure 2.1, showing the steps that were followed for many of the simulations performed for this thesis. Each step in the schematic shown in Figure 2.1 will be highlighted and discussed in further detail in sequential order in the following sections.

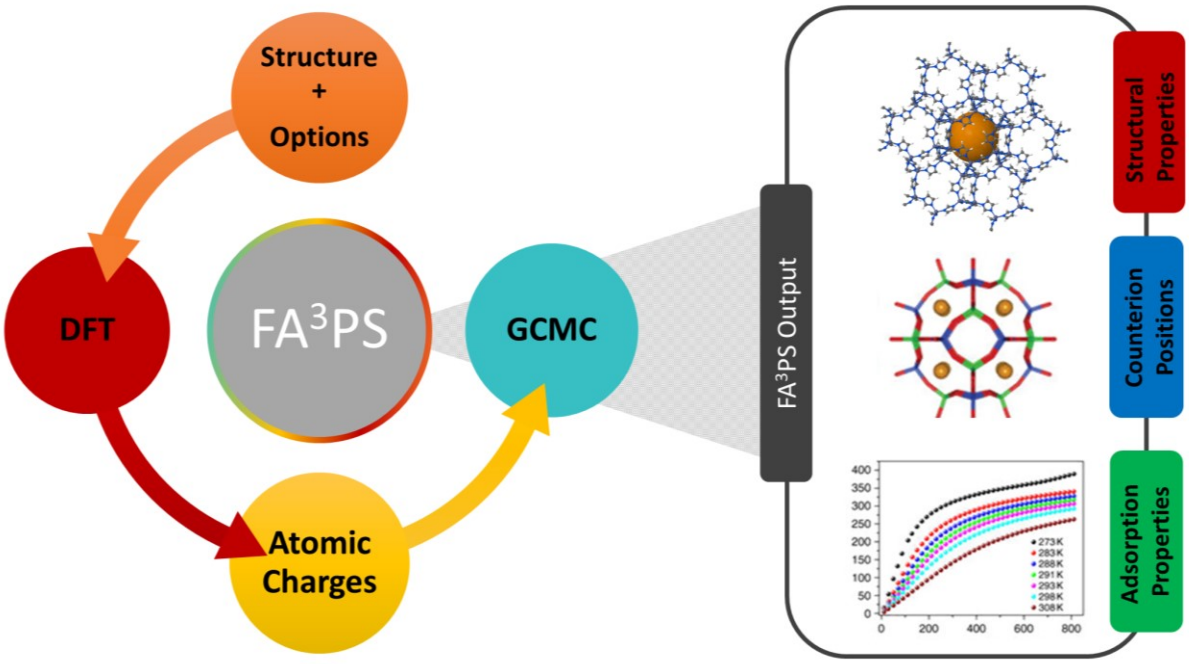

Figure 2.1 *Schematic of the FA³PS process for GCMC simulations of porous materials.*

2.2.1 Periodic Density Functional Theory (DFT)

DFT is a quantum mechanical (QM) method which has made its way from a peripheral position in quantum chemistry to center stage over the past decade, making many significant contributions to the field.[1] This is in part due to recent increases in computing power as well as further development of accurate and efficient DFT methods, making simulations involving hundreds of atoms using DFT feasible. The primary difference between DFT and wave function methods, is that the total electron density is used to calculate the total energy of the system as well to describe the complex self-interactions between electrons.[2] With competing QM methods, such as Hartree-Fock (HF) or more expensive correlated methods, where the electron-electron interactions are calculated by brute force from a complex many-body wave function that grows in dimensionality as the system size increases. Conversely, with DFT, the dimensionality of the total electronic density functional does not increase with system size and, rather, apriori knowledge of how the electrons interact with each other based on electron density is used. Therefore, DFT typically offers significant

savings in computational costs in comparison with wave function methods. Modern DFT methods offer a variety of ways to evaluate the energy from the electron density, or rather, different approaches to expressing the total energy as a functional of the electron density of the system of interest. For example, so-called DFT exchange-correlation functionals, such as the B3LYP DFT method, incorporates a fraction of the HF exchange energy into the calculation. In particular, the B3LYP method is one of the most popular functionals used by the computational chemistry community as it offers a favorable balance between cost and accuracy for many organic and inorganic systems.[2–5]

DFT has not only proven itself as a valuable computational method for molecular systems, but it has also provided valuable insights in solid state chemistry and physics where it has been an accepted method for longer than in quantum chemistry.[5, 6] The most substantial difference in the application of DFT to solid systems relative to molecular systems, is the nature of the basis sets used to expand the orbitals and electron density. For molecular systems, atom-centered basis functions are used, with Gaussian basis sets being the most popular and well-known example.[7] For solid-state systems, delocalized and periodic plane wave basis functions are used in the DFT calculations, which require the use of pseudopotentials. The pseudopotentials are used to represent the core electrons, which interact via electrostatic interactions but are unchanged by bonding. Therefore, using pseudopotentials to represent the core electrons is a valid approach for significantly reducing the computational cost of DFT calculations in periodic systems. Furthermore, as solids can be considered as a collection of infinitely repeating units, it would be computationally unfeasible to analyze such infinite systems using DFT. Thus, the translational symmetry of solid systems is used to reduce the computational cost of using DFT in solid systems by invoking periodic boundary conditions to simplify calculations.

In Figure 2.1, the first step in the automated FA^{3}PS calculation is the periodic DFT calculation, which serves two different functions. Firstly, and most importantly with respect to the work presented in this thesis, it is used to calculate the periodic QM electrostatic potential (ESP), which is needed for determining electrostatic potential fitted partial atomic charges (see section 2.1.2 for more details regarding the relationship between the ESP and the REPEAT charge calculation). Secondly, and less relevant to this work, it will refine the positions of the H atoms in the framework structure, as well as relax other framework atoms as necessary. As there are no hydrogen atoms in the zeolites structures used in this work and the structures are taken from X-ray crystal data, there is no need to optimize or relax atomic positions. As shown in the FA^{3}PS schematic in Figure 2.1, the main input is the structure of the system being studied, which is generally taken from X-ray crystallographic studies, and contained in a so-called crystallographic information file (CIF). This file contains all structural information of a periodic systems, including the atomic positions, unit cell dimensions, geometrical information, and the symmetry of the system.

The Vienna Ab Initio Software Pack (VASP)[8–11] is one of the periodic DFT implementations included in the FA^{3}PS package, and which was used for the work presented in this thesis. There are other DFT implementations that are included in the FA^{3}PS package, such as the Spanish Initiative for Electronic Simulations with Thousands of Atoms (SIESTA),[12] but as VASP was used exclusively for the work carried out in this thesis, the other methodologies will not be discussed. For all periodic DFT calculations in VASP, the Perdew-Burke-Ernzerhof (PBE) exchange correlation functional and the default projector augmented wave (PAW) pseudopotentials were used.[13–15] The reason for using VASP rather than any of the other periodic DFT implementations included in the FA^{3}PS software package is due to accuracy. Comparing with SIESTA, for example, it was stated earlier in this section that plane wave basis sets are used for DFT calculations in solid-state systems(VASP), however,

localized atom-centered basis sets can also be used in solid systems (SIESTA). Although comparable calculations using SIESTA are generally faster than those performed with VASP, highly accurate pseudopotentials are available for nearly all elements of the periodic table for VASP, while the pseudopotentials available in SIESTA are limited, and less accurate. Additionally, plane wave basis sets (VASP) are free of the basis set superposition errors that occur when using localized atomic basis sets (SIESTA), which further contributes to the increased accuracy of VASP calculations relative to SIESTA.

2.2.2 Force Fields

As discussed, using QM there are various strategies that can be employed in solving the Schrödinger equation, thereby providing a complete description of a chemical system. However, because of the elevated computational cost of these approaches, Molecular Mechanics (MM) or force field methods have been regularly used to model large systems. The improved efficiency of force field methods stems from bypassing solving the Schrödinger equation. Rather, the system is described by parameterized atom-pair potentials that have been fit to experimental data and/or high level computational simulations. These parametrized potentials describe both bonded and non-bonded interactions, however, the latter is the primary focus of the work presented in this thesis. These non-bonded interactions included in force fields can be broadly divided into two contributions: electrostatic interactions and Van der Waals (VdW) interactions.

The electrostatic interactions represent the effect of internal redistribution electrons, for which the resulting force is quite strong and can also act over long distances. The total electrostatic energy will be composed of all distinct interactions in a chemical system. Typically, electrostatic interactions can be broken into four distinct categories: ion-ion interactions, ion-dipole interactions, dipole-dipole interactions, and higher multipole moment interactions. In order to represent the

electrostatic characteristics of a system, partial atomic charges are assigned to the atoms composing that system. For small systems, such as small molecule gases (e.g. CO_2, N_2, CH_4, etc.), charges are assigned to correctly reproduce the molecule's net charge and multipole moments. In larger more complex systems, the atomic charges are typically derived through fitting to the QM derived ESP. Regardless of the system, however, a partial atomic charge is not a physical observable such as a permanent dipole, for example. Therefore, there exists various methods for assigning charges to the atoms in a given system, however, only those relevant to this thesis will be discussed.

Van der Waals interactions represent the short-ranged steric repulsion and mid-ranged dispersive attractions between atoms that are not bonded directly to each other. These are an equally important component of the interactions energy since they serve to capture induced dipole-dipole interactions as well as London dispersion forces. In non-polar molecules, such as alkanes, these are the main type of interactions, while for noble gases it is the only means for the atoms to interact. In contrast to electrostatic interactions, there is no single, unique parameter that governs this behavior. Rather, atom-pair or atom-specific parameters are available that have been optimized to model specific chemical systems. More robust, general-purpose force fields are also available, such as the universal force field (UFF) which encompass all atoms in the periodic table.

2.2.3 REPEAT Charge Calculation

The second step in calculating a gas adsorption isotherm with FA^{3}PS is the calculations of the partial atomic charges. There are many methods that have been used for calculating the atomic charges in porous materials such as MOFs and zeolites, but the default method in FA^{3}PS is the Repeating Electrostatic Potential Extracted ATomic (REPEAT) method developed by the Woo group in 2009.[16] The charges generated by this method are classified as electrostatic potential (ESP) fitted partial atomic charges, which have gained in popularity when modeling periodic systems.[17,18]

This method derives charges for each of the atoms in a periodic system based on a DFT-derived ESP grid surrounding each atomic site between 1 -2 VdW radii. The ESP is defined as the energy required to bring a unit charge from an infinite distance away from a point r. Beyond the cutoff distance of 2 VdW radii around each atom in the system, the ESP begins to drop off rapidly and is left out of the least-squares fitting procedure utilized by the REPEAT method in order to reduce the computational expense of the charge calculation. The REPEAT method represents an important advance in the derivation of partial atomic charges in periodic systems, as it was not until its introduction that charges had been derived directly from the periodic QM ESP of a solid system. The problem that was overcome by the REPEAT method arises from the ill-defined nature of the QM ESP in periodic systems, making it difficult to define a reference state for the ESP. This issue emerges due to the fact that the absolute energy of an atom in an infinitely repeating periodic system is not an intrinsic bulk property. To overcome this, Campañá *et al.* introduced a new error functional which acts on the relative differences in the potential rather than on absolute values. A more in depth mathematical description of the REPEAT method will be given in the Computational Details section of Chapter 3 of this thesis.

Since its introduction in 2009, the REPEAT method has become very popular in modelling the electrostatic properties in periodic systems such as MOFs and zeolites. In fact, many researchers have successfully used REPEAT to predict the adsorption properties as well as other physical properties for periodic materials.[19–24] Furthermore, the work of Watanabe *et al.* [25] provides an excellent example of the improved quality of the partial atomic charges generated by the REPEAT method relative to other charge derivation methods used for periodic systems. In this work, Watanabe *et al.* compared many different charge calculation methods when modeling the

thermodynamic properties of the MOF ZIF-8, and amongst the methods compared, REPEAT showed the best agreement with experimental data.

2.2.4 The Lennard-Jones Potential

Once the QM ESP has been calculated and the resulting partial atomic charges on the framework atoms have been determined via the REPEAT method, FA^{3}PS can proceed with the force field based GCMC simulations. To accurately capture the VdW interactions in a given chemical system, it is necessary that the potential behaves in a physically meaningful manner. With respect to the distance separating an interacting atom-pair, the chosen function should have the following characteristics:

1) At long interatomic distances, the interaction energy tends towards zero.
2) At intermediate distances, the interaction energy should tend to a slightly negative minimum, which represents the equilibrium separation between a given atom-pair. Over this range of interatomic distances, the interaction is attractive.
3) At short distances, the interaction energy begins to tend towards positive values as repulsive interactions begin to dominate.

The most popular function used to model VdW interactions, which has these characteristics, is the Lennard-Jones (LJ), or 12-6, potential. Other functional forms exist, such as the Buckingham potential, but the LJ form is favored by many since it only has two parameters and is relatively computationally efficient. The latter is related to the functional form, which is defined in Eq.(2.1):

$$E_{VdW} = E_{LJ}(r_{ij}) = 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \quad (2.1)$$

The 12-6 potential contains two empirical parameters, which represent the different interatomic interactions present in a given system. The first of these parameters is ε_{ij}, which represents the potential well-depth and is related to the binding energy between a given atom-pair. The second parameter is σ_{ij}, which is the finite distance at which the interatomic potential is equal to zero. Figure 2.2 illustrates the form of both the Coulomb and LJ potentials as a function of the interatomic distance separating a given atom pair. The reason the LJ potential is more efficient than other potential forms is due to the 12-6 form of LJ potential. When computing exponential functions or roots, the computational expense is around 5 times that of simple multiplications of additions. In the 12-6 form of the LJ potential, a 6th power calculations is required and because of this, the 12th power calculation does not require much additional effort (the square of the 6th power).

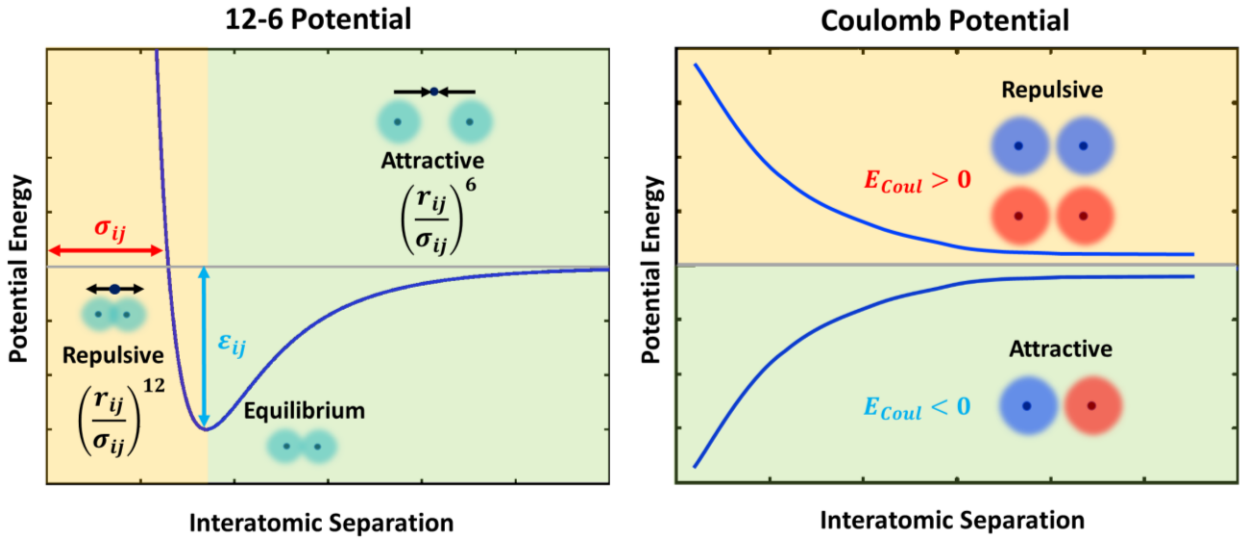

Figure 2.2 *Form of the Potentials Used to Calculate the Interactions Energies during MC/GCMC Simulations.*

Interactions between different atom types, *i* and *j*, are calculated using the Lorentz-Berthelot mixing rules, which are defined as:

$$\sigma_{ij} = \frac{\sigma_i + \sigma_j}{2} \quad (2.2)$$

and,

$$\varepsilon_{ij} = \sqrt{\varepsilon_i \varepsilon_j}. \quad (2.3)$$

In this way, LJ parameters can be calculated for all atom pair combinations in any system, so long as parameters are defined for every unique atom type. Frequently, the parameters in the 12-6 potential are fit to reproduce experimental data and/or high-level QM calculations for a given system. This will be the focus of Chapter 4 of this thesis, and will be discussed in more detail in this chapter.

2.2.5 Grand-Canonical Monte-Carlo (GCMC) Simulations

The principles of statistical mechanics provide a powerful connection between the microscopic properties modeled in computational simulations and the observed macroscopic properties of the bulk system. From statistical mechanics, it is shown that the observed macroscopic properties are simply the limit averages of the microscopic properties examined in simulations. The 'GC' portion of GCMC references the name of the statistical thermodynamic ensemble which is utilized by the simulation. In statistical mechanics, an ensemble is a fundamental concept representing a collection of a large number of replicas of a system of interest, for example a collection of porous material unit cells containing adsorbed gas molecules. In the Grand Canonical ensemble, the chemical potential (μ), temperature (T), and the system volume (V) are kept constant. The chemical potential is a fundamental thermodynamic property of a system, which is defined as the change in potential energy of a system with respect to the number of particles in that system. In a GCMC simulation, the number of particles (N) is allowed to vary and, therefore, the number of guest within a host system may either increase of decrease. As such, since the chemical potential and the temperature are held constant, the results yielded by GCMC simulations are representative of the system at chemical and thermal equilibrium.

The 'MC' portion of GCMC refers to the Monte Carlo method which was appropriately named after a region in Monaco, known best by visitors for its famous Monte Carlo casino and resort. The choice of this name reflects the core principle of the Monte Carlo method, which can simply be thought of as a sophisticated random sampling method. The Monte Carlo method has found success in many field including mathematics, finance, physics, biology and chemistry for systems involving a large number of repetitive stochastic processes.[26–30] In the context of chemical systems, a Monte Carlo simulation starts by randomly selecting a guest molecule in the system, be that a gas molecule or a counter-ion, and perturbs the molecule in some way through different move types including rotation, translation, or a conformational change. Once the newly perturbed configuration is created, the potential energy of the entire system is calculated via empirical potentials simulating the electrostatic and Van der Waals interactions of a system.

Once the potential energy of the new configurations has been calculated, three possible scenarios can arise: the potential energy of the new configuration can be significantly higher, or marginally higher, or lower than that of the prior configuration. If the energy of the new configuration is lower than the previous configuration, it is automatically accepted and added to the ensemble average, which is the average over all accepted configurations during the simulation. If the energy of the new configuration is slightly higher than the configuration prior to the perturbation, it has a chance of being accepted that is determined by the Boltzmann weighting scheme. This enables the sampling to surmount energy barriers in the potential energy surface, which allows for a better sampling of the entire phase space. Lastly, if the energy of the new configuration is higher than that of the previous configuration, then the new configuration will usually be rejected, and the previous configurations is added back into the ensemble average. It is important in this case to re-add the previous configuration potential energy to the ensemble average since favoured configurations

should be sampled more in order to preserve a physically meaningful distribution. In line with this, even unfavourable, high energy configurations have a small, non-zero probability of existing. Hence, high energy configurations can occasionally be accepted and added to the ensemble average, in agreement with the Boltzmann distribution. This process of generating a new configuration by perturbing the system, evaluating the resultant potential energy and adding it to the ensemble average is repeated iteratively until the ensemble average has converged, and the system energy is relatively constant.

When MC simulations are coupled with the grand canonical ensemble, one has two additional move types available to perturb the system and generate a new configuration: particle insertion and deletion. In a similar manner as described for MC simulations, inserted guest molecules that result in a lower potential energy will have a higher probability of being accepted and used as the starting point for subsequent configurations. The fact that the number of molecules is not held constant in the grand canonical ensemble makes it ideal for describing systems in which guest molecules are free to move into and out of the system. In context of this research, this makes the grand canonical ensemble especially suited for describing the adsorption properties of porous material such as MOFs and zeolites.[31] With the grand canonical ensemble, each subsystem, or replica, is put in contact with a heat bath and with other replicas. The chemical potential and temperature of all the systems composing the ensemble are allowed to reach equilibrium with the heat bath then transferred to a closed, thermally isolated container. All of these subsystems are open systems for which matter and heat may be transferred, and are in thermal equilibrium with each other. The exchange of matter between subsystems is governed by the fact that they all have the same chemical potential. For the grand canonical scheme, which is commonly referred to as the

μVT-ensemble, it can be shown that the probability of finding the system in a state with energy E_i is given by:

$$P_{N,i} = \frac{e^{-\beta(E_{N,i}+N\mu)}}{\sum_N \sum_i e^{-\beta(E_{N,i}+N\mu)}} = \frac{e^{-\beta(E_{N,i}+N\mu)}}{\Omega(\mu VT)} \quad (2.4)$$

As noted earlier, the key feature of the grand canonical Monte Carlo method is that the number of particles may change throughout the simulation. Now that the probability distribution for the grand canonical ensemble has been defined in Eq. (2.4), it is now possible to discuss how this ensemble is applied in Monte-Carlo simulations to determine the physical observables of a given system of interest. For any physical observable A that depends on n independent variables (denoted $\{x_1, \dots, x_n\}$), with a general probability distribution $P(x_1, \dots, x_n)$, the expectation value $\langle A \rangle$ is given by:

$$\langle A \rangle = \int P(x_1, \dots, x_n) A(x_1, \dots, x_n) dx_1 \dots dx_n \quad (2.5)$$

where the integral is taken over the domain of the probability distribution. In context of this study, the property of interest is the number of guest molecules, N (or more specifically in terms of MOFs and zeolites, the number of counter-ions or adsorbate gas molecules). As the number of guest-molecules in the grand canonical ensemble are not fixed, this allows for a simple calculation of $\langle N \rangle$. However, a direct calculation of the integral in equation (2.5) is not feasible as it would require the calculation of the large number of terms contained in $\Omega(\mu VT)$. Therefore, the applied method does not work directly with calculations using the grand canonical partition function, but rather describes how to sample points in phase space such that as the number of points collected increases, the sampled points form a probability distribution which approaches the grand canonical probability

distribution. This method ensures that once a large amount of sample points is collected, the value of $\langle N \rangle$, given by:

$$\langle N \rangle_{obs} = \frac{\int N(\boldsymbol{q},\boldsymbol{p}) e^{-\beta(E(q,p)+N\mu)} d\boldsymbol{q} d\boldsymbol{p}}{\Omega(\mu VT)} \quad (2.6)$$

can be approximated by Eq.(2.7) below:

$$\langle N \rangle_{obs} \approx \frac{1}{\eta} \sum_{i=1}^{\eta} N(\boldsymbol{q}_i \boldsymbol{p}_i) = \langle N \rangle_{trial} \quad (2.7)$$

where η represents the total number of collected sample points. In using this approximation, the configurations that are sampled must be sampled in accordance with the grand canonical probability distribution. In other words, configurations which are more probable (with respect to the grand canonical probability distribution) are to be sampled more frequently.

Figure 2.3 shows a typical plot of the variation in the number of guests and the fluctuations in the potential energy as a function of the number of GCMC steps taken. The two stages which are typical of a GCMC simulation are (1) equilibration and (2) production, which are indicated on the plot shown in Figure 2.3. In the equilibration phase, it is more favourable for a guest to be inserted and adsorbed than to be deleted, which is reflected in the sharp increase in the number of guests until around 3.5 million GCMC steps in Figure 2.3. After numerous insertions are accepted, the pores in the material become more saturated with adsorbate molecules and it becomes less favourable to further add guests to the pores. At this point, the rate of accepted insertions becomes equal to the number of accepted deletions, which is reflected in Figure 2.3 as the number of guests begins to fluctuate about an average number of guests. At this point, the rate of accepted insertions becomes equal to the number of accepted deletions, which is reflected in Figure 2.3 as the number

of guests begins to fluctuate about an average number of guests. This indicates that the simulation has reached an equilibrium state.

The fact that the chemical potential of the adsorbed gas and that of the free gas are equal at equilibrium allows for μ_{ads} to be obtained easily from μ_{gas}, which is known for ideal gases. In this way, the expression for ideal gases can be used to determine the acceptance probability as molecules of gas are inserted or deleted from the system. An important result of this is that the equilibrium phase data generated by GCMC simulations can then be further applied towards the calculation of equilibrium adsorption properties of a given system, including gas uptake isotherms and heats of adsorption, to name a few. Because of the relative simplicity of this method, GCMC has been extensively used in the study of gas adsorption in porous materials.[32]

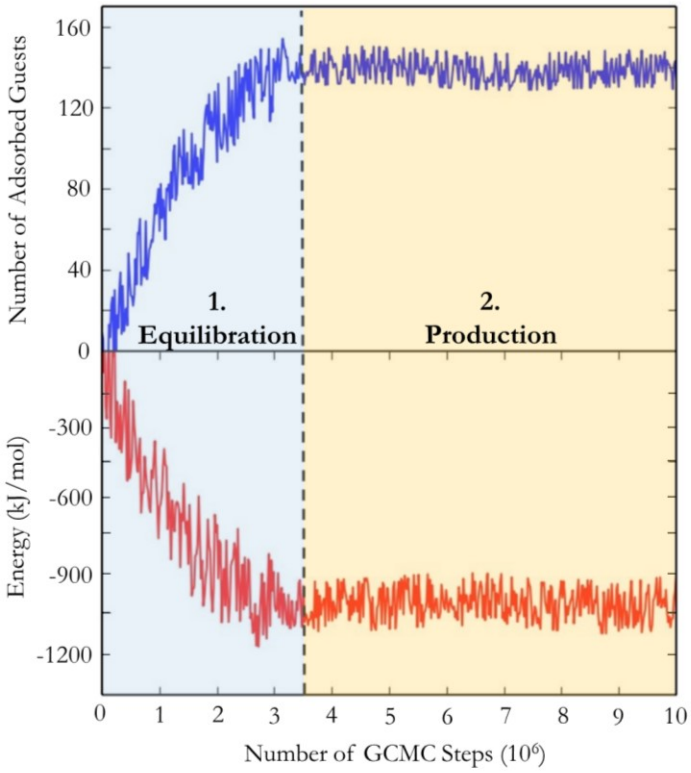

Figure 2.3 *Plots showing the number of adsorbed guests (top) and the total energy (bottom) during a typical GCMC simulation with a MOF. The equilibration and productions phases of the simulations are indicated.*

2.2.6 Special Sampling for Counter Ions

For charged systems, in which there are counter-ions or charged guests, proper sampling of the guest locations is ensured by the addition of a new move type developed by the Woo lab called the 'double-jump' (Figure 2.4). The double jump move was implemented such that two jump moves of two randomly selected counter ions are performed simultaneously before the potential energy of the new configuration is evaluated. The reasoning behind the addition of this move type in charged systems arises from unfavourable interaction between same sign charges, as would be the case with counter-ions or charged guests. If only a single ion is perturbed at a time, the number of energetically unfavorable configuration due to repulsive electrostatic interactions becomes inflated. Hence, without the double jump move energetically favourable configurations would not be sufficiently sampled, thereby resulting in a nonphysical probability distribution.

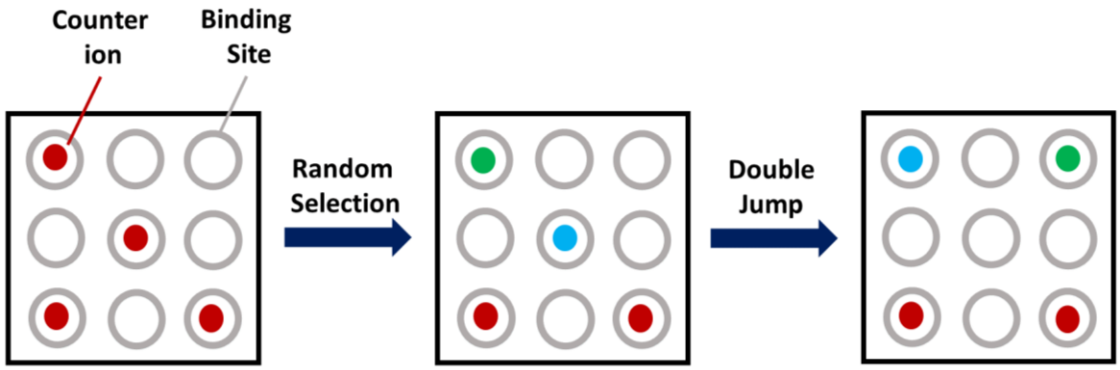

Figure 2.4 *Depiction of the 'Double Jump' moved used in MC/GCMC simulations for systems with a net charge.*

2.3 Electronegativity Equalization Based Charge Methods

In section 2.1 the REPEAT charge derivation method was discussed. This method is a rigorous atomic charge derivation method that is based on DFT calculations of the QM ESP. As such, it is a compute intensive method for evaluating the partial atomic charges in a given system that is best suited for in depth studies that require a high level of accuracy. However, when

performing screening studies of a large number of materials the associated computational expense of the REPEAT method makes it less desirable. In this section, parameterized charge derivation methods based on the electronegativity equalization principle will be discussed. These methods are better suited towards large screening studies, owing principally to their reduced computational expense relative to QM based methods such as REPEAT.

2.3.1 The Charge Equilibration (QEq) Model

The charge equilibration method, first proposed by Rappé and Goddard [33] is the simplest of the electronegativity equalization charge models. In essence, it is a second order expansion of the molecular energy in terms of the partial atomic charges. The energy expression takes on the following form:

$$E_Q^{QEq} = \sum_i^N \left(\frac{1}{2}\,\kappa_i Q_i^2 + \chi_i Q_i\right) + \frac{1}{2}\sum_i^N \sum_{j\neq i}^N J(r_{ij})\, Q_i Q_j \qquad (2.8)$$

Here, Q_i is the atomic charge, κ_i is the atomic hardness, and χ_i is the atomic electronegativity. The second half of the equation represents the Coulombic term, describing the electrostatic interaction between point charges. The atomic hardness and electronegativity are parameters that are specific to each system. They simulate the charge distribution throughout the atomic system by loosely defining a given atom's ability to attract charge. More specifically, the atomic electronegativity and atomic hardness are related to the ionization potential (IP) and the electron affinity (EA) of a given atom:

$$\chi_i = {}^1\!/_2\,(IP + EA), \quad (2.9)$$

$$\kappa_i = IP - EA \qquad (2.10)$$

What makes the QEq method, as well as extensions of the QEq model, so appealing is that they are easily parameterized to specific atomistic systems. In the energy expression given in Eq.(2.8), both the atomic hardness and atomic electronegativity terms may be treated as adjustable model parameters. Differentiation of the energy expression in Eq. (2.8) with respect to each of the N atomic charges yields a set of N linear equations:

$$\frac{\partial E_Q^{QEq}}{Q_i} = \chi_i + \kappa_i Q_i + \sum_{i \neq j}^{N} J(r_{ij}) Q_j = 0 \quad (2.11)$$

When set to zero, the linear set of equations may be solved for a set of charges that will minimize the energy. The minimization itself is constrained by the condition that the second order derivative of the Hessian matrix, defined by the following elements:

$$\frac{\partial^2 E_Q^{QEq}}{\partial {Q_i}^2} = J_{ii} = \kappa_i , \quad (2.13)$$

and

$$\frac{\partial^2 E_Q^{QEq}}{\partial Q_i {Q_j}^2} = J_{ij}, \quad (2.13)$$

must be positive definite.

Expressed in matrix form, the equations may be written as:

$$\mathbf{HQ} = -\boldsymbol{\chi}, \quad (2.14)$$

where $\boldsymbol{\chi} = [\chi_1 \chi_2, \ldots, \chi_N]^T$ and $\mathbf{Q} = [Q_1, Q_2, \ldots, Q_N]^T$ are vectors formed from the atomic electronegativities and atomic charges for a system of N atoms, and the symmetric matrix **H** is

assembled from the Hessian elements defined in Eq. (2.12) and Eq. (2.13). From these definitions, the diagonal of **H** corresponds to the atomic hardness for each atom i, while the off-diagonal elements correspond to Coulombic repulsions between atoms i and j. In practice, the set of linear equations must be augmented by an additional constraint on the net charge of the system. In general, the charges are constrained so as to guarantee molecular charge neutrality:

$$\sum_{i=1}^{N} Q_i = 0. \quad (2.15)$$

However, this can be generalized so as to incorporate charged systems by using a Lagrange multiplier and substituting a neutral charge by a total charge, Q_{tot} :

$$\sum_{i=1}^{N} Q_i = Q_{tot}, \quad (2.16)$$

in this way, both neutral and charged systems can be considered. The corresponding Lagrangian for Eq. (2.8) is then given by:

$$L_Q^{QEq} = E_Q^{QEq} - \lambda\left(\sum_{i}^{N} Q_i - Q_{tot}\right), \quad (2.17)$$

and variational minimization of the function given in Eq. (2.17) with respect to each of the N atomic charges yields N linear equations:

$$\lambda = \kappa_i Q_i + \chi_i + \sum_{j \neq i}^{N} J(r_{ij}) Q_j. \quad (2.18)$$

This can be written in 2x2 block matrix form to give:

$$\begin{pmatrix} \boldsymbol{H} & \mathbf{1} \\ \mathbf{1}^T & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{q} \\ \lambda \end{pmatrix} = \begin{pmatrix} -\boldsymbol{\chi} \\ Q_{tot} \end{pmatrix}, \quad (2.19)$$

where $\boldsymbol{H}$ is an N × N matrix with the same form as in Eq. (2.13), $\mathbf{1}$, $\boldsymbol{q}$, and $\boldsymbol{\chi}$ are column N-vectors. The vector $\boldsymbol{q} = [Q_1, Q_2, \ldots, Q_N]^T$ contains all of the partial atomic charges, and the vector $\boldsymbol{\chi} = [\chi_1\, \chi_2, \ldots, \chi_N]^T$ contains the atomic electronegativities. Using this modification to the original QEq model proposed by Rappé and Goddard, it can now be applied to systems having overall net charge.

2.3.2 The Split-Charge Equilibration (SQE) Model

The split-charge equilibration (SQE) method is an extension of the QEq modeled discussed in the previous section. It was developed by Mueser *et al.* [34] to overcome many of the limitations of the QEq method. In particular, the SQE method introduces the concept of split-charges that adds additional information about the covalent bonds in the system being modeled, and the energy expression is written in terms of these so called split charges, q_{ij}. Atom centered partial atomic charges are then obtained from the split-charges given by Eq. (2.20). Here the split charge represents the charge that flows from atom *j* to atom *i*, and the sum is over all atoms directly bonded to atom *i* (Figure 2.5).

$$Q_i = \sum_j q_{ij} \quad (2.20)$$

As a generalization of the atomic hardness and electronegativity, additional bond hardness, κ_{ij}, and bond electronegativity, χ_{ij}, parameters are introduced in the SQE energy expression given in Eq. (2.21), where the second term is simply the QEq energy expression in terms of the atomic charges as given previously in Eq. (2.8).

$$E_Q^{SQE} = \frac{1}{2}\sum_i^N \sum_j^N \left(\chi_{ij} q_{ij} + \frac{1}{2}\kappa_{ij} q_{ij}^2\right) + \sum_i^N \left(\frac{1}{2}\kappa_i Q_i^2 + \chi_i Q_i\right) + \frac{1}{2}\sum_i^N \sum_{j\neq i}^N J(r_{ij})\, Q_i Q_j \quad (2.21)$$

$$E_Q^{SQE} = \frac{1}{2}\sum_i^N \sum_j^N \left(\chi_{ij} q_{ij} + \frac{1}{2}\kappa_{ij} q_{ij}^2\right) + E_Q^{QEq} \quad (2.22)$$

$$E_Q^{SQE} = E_Q^{AACT} + E_Q^{QEq} \quad (2.23)$$

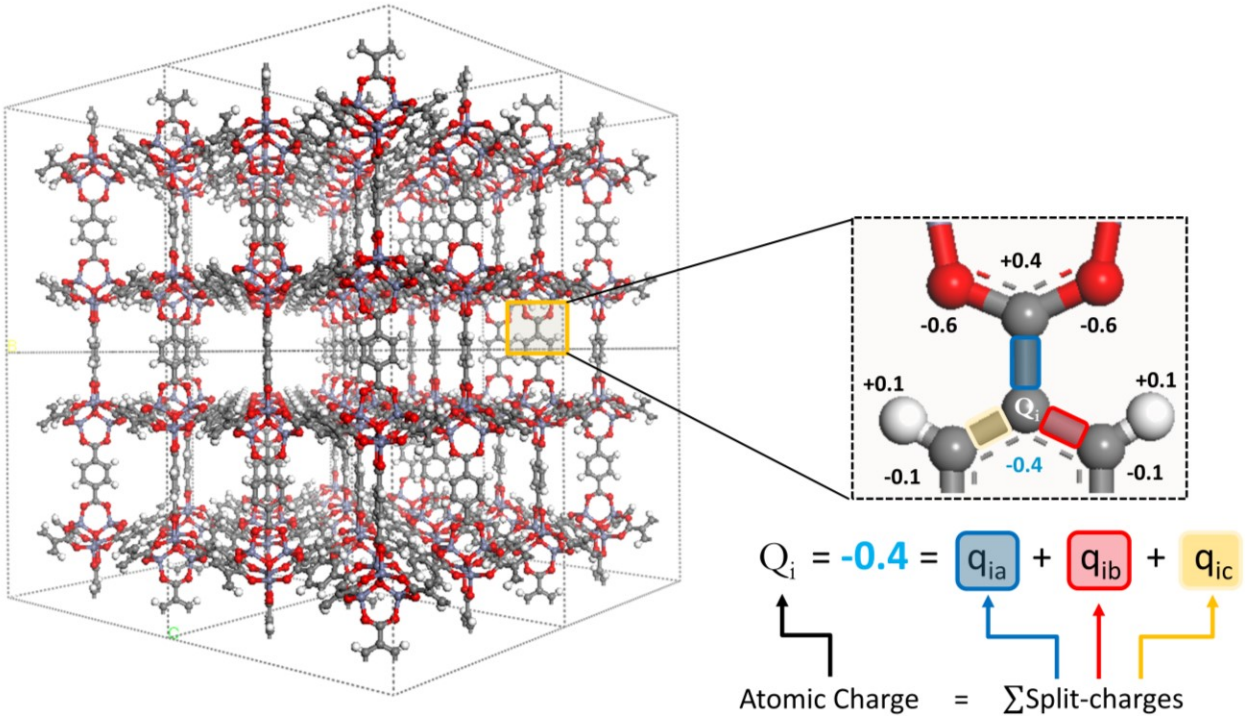

Figure 2.5 *Illustration of how partial atomic charges are determined as the sum of split charges, for IRMOF-1. The concept of split-charges is central to the SQE method, and imparts the model with information regarding connectivity in covalently bonded systems, such as MOFs and zeolites.*

When either the bond or atomic parameters are set to zero, it can be seen that the SQE model encompasses both the QEq and the atom-atom charge transfer (AACT)[35] models, respectively. With the addition of bonding information in the form of split-charges, the SQE method overcomes the main limitation of the QEq model: its failure to accurately predict bond dissociation limits, particularly in covalently bonded systems. Additionally, the greater number of parameters present in

the SQE method (2 additional parameters for each unique bond-type), results in a greater number of degrees of freedom during the charge fitting procedure, resulting in a more flexible charge model. Due to the antisymmetry property of the split-charges, however, the SQE model is limited to charge neutral systems. This property, expressed in Eq. (2.24), is meant to conserve the flow of charge between two atoms, but does so by conserving charge neutrality of the system. Unlike with QEq, no trivial mathematical transformation can be applied to allow for the accommodation of charged systems. The circumvention of this limitation will be the focus of Chapter 3, where it will be discussed in greater detail.

$$q_{ij} = -q_{ji} \quad (2.24)$$

2.4 Parameter Optimization Methods

The force field optimizations performed for aluminosilicate zeolites involved the optimization of empirical parameters for both the electrostatic and Van der Waals components of the interaction energies being modeled. In order to do this, a custom genetic algorithm was used, followed by fine-tuning of the resultant parameters using a gradient based optimizer. The genetic algorithm used to optimize parameters for the charge models was developed by Sean Collins of the Woo group and was not modified in anyway. For the optimization of the Lennard-Jones parameters used to model the dispersion interactions in aluminosilicate zeolites, certain modifications were needed, however, the overall routine was left untouched. For the fine tuning of the parameters, I have written my own gradient descent algorithm. In this section, a general overview of the optimization methods employed in this work will be given in the sequential order in which they were used. A discussion of the more specific details related to the parameter optimizations performed will be outlined in Chapter 4 of this thesis.

2.4.1 Initial Parameter Optimization using a Genetic Algorithm(GA)

Genetic algorithms (GAs) are heuristic solution search algorithms inspired by Darwinian evolution, which have been employed towards efficient optimization of large and complex search spaces.[36] They are built upon the genetic representation of a solution, or *chromosome*, which is evolved through successive *generations* until an optimal solutions is obtained. The *population* of a given generation represents a subset of all possible, or encoded, solutions to the given problem. In the context of the work presented in this thesis, the solution for which the GA is searching is an optimal set of parameters for modeling the electrostatic and VdW interactions within aluminosilicate zeolites. Remaining consistent with its definition, the population represents a subset of all possible parameter sets, that are randomly generated within a fixed a range, and where each of the solutions is referred to as a chromosome. In an analogous manner, a *gene* corresponds to a unique parameter within a particular chromosome, while the *allele* is the value taken on by the parameter. As the algorithm progresses, the population evolves towards a more fit solution, with each successive stage in the process being referred to as a *generation*.

Unlike conventional search algorithms, such as grid searches or gradient descents, a GA converges upon an optimal solution from multiple directions in the search space. This results in the primary advantage of GAs with respect to other optimization techniques, which is that it is a global search algorithm. In other words, the algorithm is not as easily trapped in local minima that exist throughout the search space. The GA overcomes the shared drawback of all traditional optimization techniques by sampling a larger portion of the total search space, in contrast with traditional methods that focus their search in a more localized region.

In the first step of the genetic algorithm, a population of set size, N_{pop}, is randomly generated. In general, it is important that the population size is not too small in order to avoid having the GA quickly converged to a local minimum. Just as in nature, a small population will have little genetic diversity and will quickly stagnate, preventing a more thorough exploration of the search space. Thus, the greater the population size the better the chances of discovering a global minimum. However, a larger population generally corresponds with a greater requirement for computational resources. Therefore, a balance should be found and, for the optimizations carried out for this thesis, a population size of 30 – 35 was found (through trial and error) to be a sufficiently large population to avoid a poor 'mating pool'.

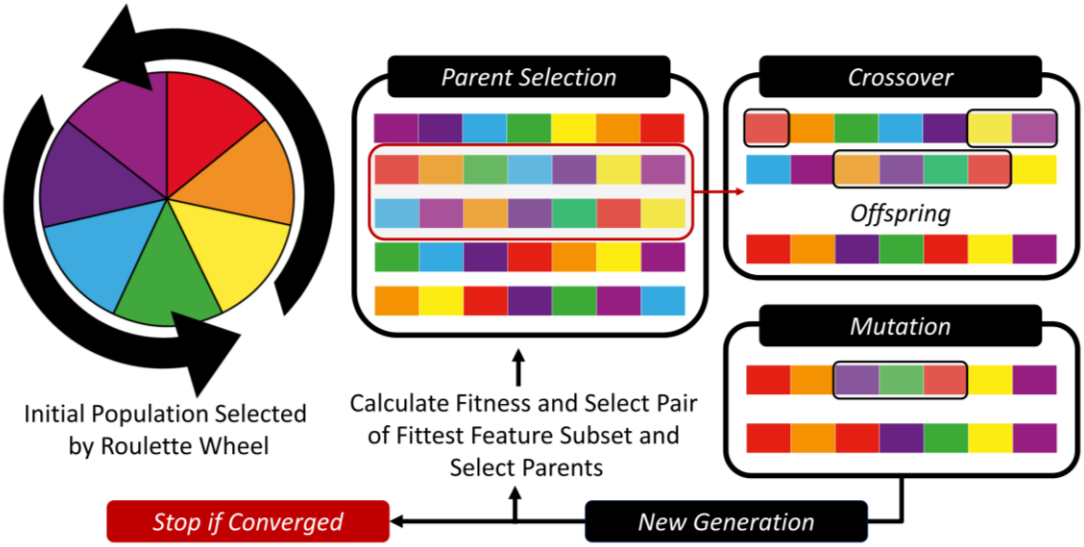

Figure 2.6 *The Steps involved in the Genetic Algorithm*

Once the initial population has been generated, the next step in the GA is to calculate the fitness of each solution and rank each member of the generation according to their fitness score. A fitness function, also known as a cost or lost function, is not a concept exclusive to the GA, but is employed in all optimization techniques. A fitness function provides a relationship between the candidate solution to the problem and the quality of said solution. There is an endless possibility of

functions that can be used to evaluate the fitness of candidate solutions, but one should avoid the use of overly complicated fitness functions. Such use may reduce the computational efficiency of the algorithm with no improvement with regards to the identification and ranking of fit solutions or, chromosomes. The specifics of the fitness functions used for the parameter optimizations carried out in this thesis will be discussed in greater detail in Chapter 4.

The next step of the GA is where the evolutionary principle of 'survival of the fittest' comes into action. After each member of the generation is ranked according to its fitness score, individuals within the generation then mate with one another in a process known as *crossover*. From this process, a new generation of offspring of the same population size is obtained. This is known as a generational population model, which is in contrast to the steady state population model, or incremental GA. In this scheme, a population of one or two offspring's is generated in each generation (or iteration) to replace one or two individuals from the initial population. In general, a generational model is best suited for complex and intractable optimization problems whereas a steady state model is best suited for optimizations that involve a very large number of parameters (when a generational model becomes to computationally costly).[37] It should be noted, that the GA used for the optimizations carried out in this work utilizes the generational population model for crossover.

The pair of solutions that undergo crossover are known as *parents*. The parent selection process is crucial to the convergence rate of the GA as quality parents drive individuals to better and fitter solutions. However, care should be taken to prevent one extremely fit solution from taking over the entire population in a few generations, as this will lead to a loss of diversity known as premature convergence. Preservation of good diversity in the population is imperative for the success of the GA. [38] In this regard, fitness proportionate selection is one of the most popular

methods employed in the parent selection process. In this method, each individual can become a parent with a probability that is proportional to its fitness score, and fitter individuals will have a higher probability of mating and propagating their features to the next generation. Such a selection scheme is meant to mimic natural selection, in which more fit individuals will preserve their genetic characteristics through many generations.

Two implementations of fitness proportionate selection[39] are possible: (1) roulette wheel selection and, (2) stochastic universal sampling (SUS). In a roulette wheel selection, the wheel is divided into N slices, where each slice represents an individual in the population. Each individual is assigned a portion of the wheel, with the size of this slice determined its fitness score: more fit solutions will take a larger portion. Then, a fixed point is chosen on the wheel circumference as shown in Figure 2.7a and the wheel is spun. The slice, or region, in which the fixed point lies after rotation is chosen as the first parent. This process is then repeated in order to selec the second parent for mating. From Figure 2.7a it is clear that a more fit individual has a greater piece of the pie and, therefore, has a great probability of being selected as a parent. The SUS method for parent selection is very similar to roulette wheel selection, however in this case there can be multiple fixed points placed around the wheel as shown in Figure 2.7b. In this scheme, all the parents can be selected in a single spin of the wheel. Furthermore, such a parent selection scheme encourages the highly fit individuals to be selected at least once. It should be noted that fitness proportional selection methods are not valid for cases where the fitness can take on a negative value. In such cases, methods such as tournament selection or rank selection may be employed. However, as these methods are outside the scope of this thesis, they will not be discussed in further detail.

There are different crossover (mating) schemes, including one-point crossover, multi-point crossover and arithmetic recombination (Figure 2.8). In a one-point crossover scheme, a random

crossover point is selected in each chromosome and the tails of the two parents are swapped to generate the new offspring. Multi-point crossover is similar, but in this case multiple points in each chromosome are used and alternating segments are swapped between parents to generate the new offspring. Finally, arithmetic recombination is more common crossover scheme when performing numerical optimizations such as those carried out in this thesis. In this scheme, the parent's genes hold numerical values and the offspring will inherit a value somewhere between that of both the parents. Each gene is then generated for each offspring from both parents, with each parent having the possibility of contributing more or less of their characteristics to either offspring. As with parent selection, the weighting of parents in this crossover scheme is generally related to the fitness of the parents. In this way, favourable characteristics are preserved and propagated through subsequent generations of the GA.

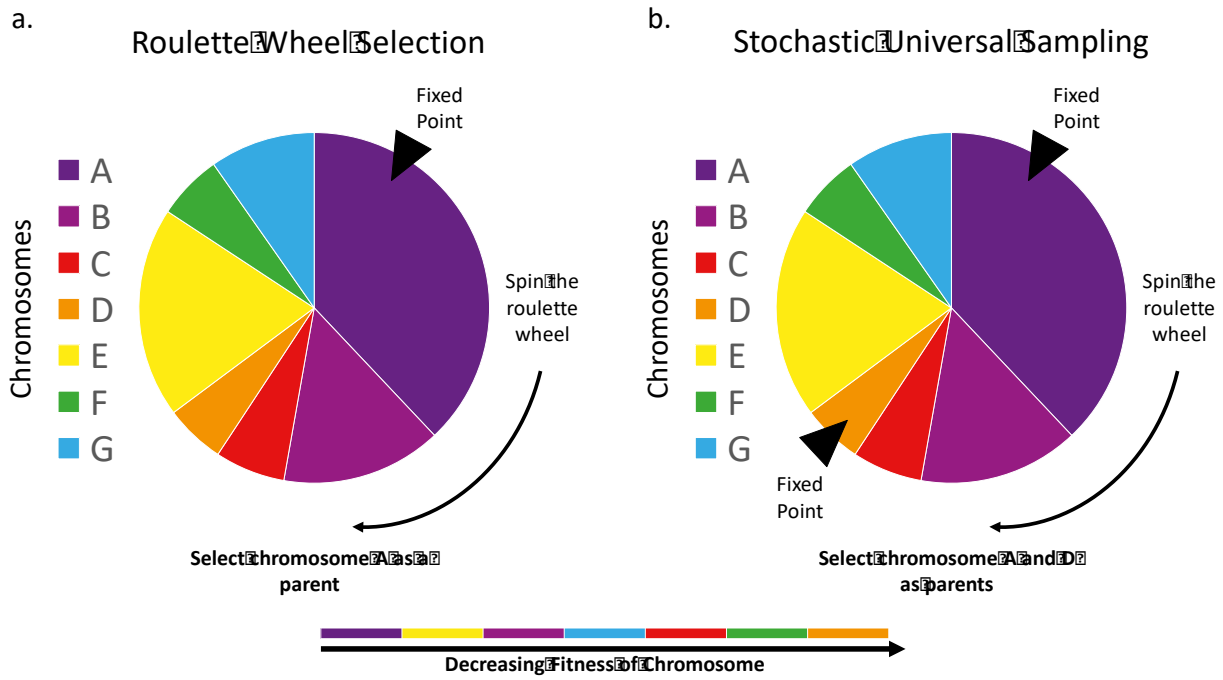

Figure 2.7 *Implementations of Fitness Proportionate Selection: a) Roulette wheel selection b) Stochastic Universal Sampling*

The next step in the genetic algorithm, which marks the end of one generation and the start of the next, is mutation.[40] As in nature, there is the chance genetic mutations to occur, resulting in a slightly altered chromosome and a new solution. Allowing mutations to occur in the GA helps promote diversity within the population, and prevent premature convergence. Some of the possible mutation types that have been implemented in GAs include swap mutations, scramble mutations and random mutations. In a swap mutation, two positions on the chromosomes of the parents are selected at random, and the values are interchanged. A scramble mutation, as the name indicates, involves randomly shuffling a subset of the genes in the entire chromosome. Random mutations are implemented in the GA used in this work, where randomly selected genes in a chromosome are selected, then randomly perturbed. A summary of each of the mutation types discussed here are illustrated in Figure 2.9. In the example given in Figure 2.9 for random mutations, for each gene a random number, *r*, is generated and if it is below the mutation rate, M, then the gene is randomly mutated within a range of $+/- 2r \cdot (1 - \alpha_i)$. After a possible mutations are applied to a given chromosome, the population for the next generation is again evaluated, and the entire process is repeated. The GA will continue until the evolution stagnates and the same chromosome maintains the greatest fitness for a set number of generations. At this point the GA has converged, and the algorithm stops.

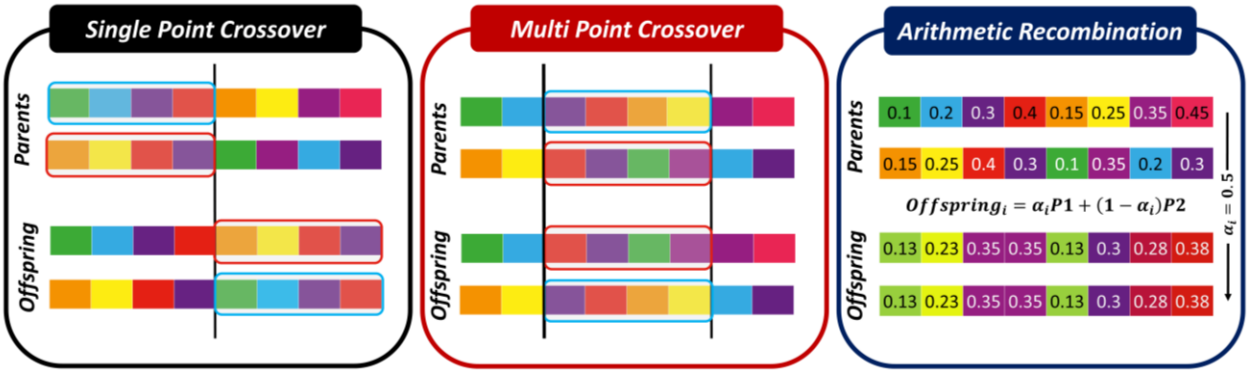

Figure 2.8 *Different possible crossover schemes. Arithmetic recombination is the crossover scheme employed by the GA used for the parameter optimizations carried out in Chapter 4.*

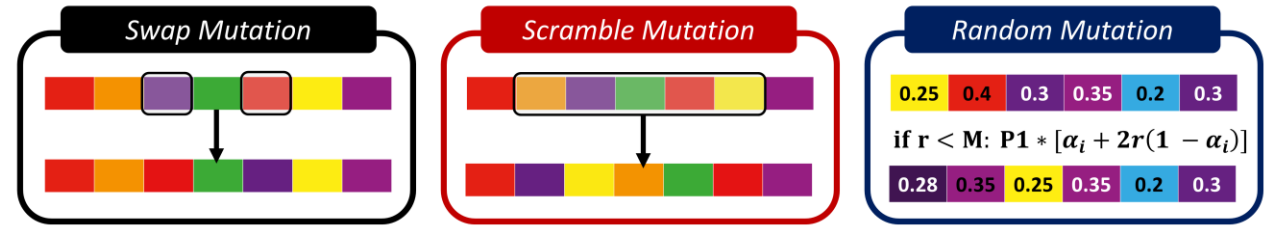

Figure 2.9 *The different types of possible mutations in the GA. Random mutations have been implemented in the GAs used in this work.*

2.4.2 Latin Hypercube Sampling (LHS)

In section 2.3.1, the initial selection of the population was said to be randomly generated. When optimizing parameters for the 12-6 Lennard Jones potential, the increase in computational expense is significant in comparison to when attempting to optimize parameters for QEq and SQE. Thus, in order to increase the efficiency of the GA for the VdW parameter optimization, Latin Hypercube Sampling was implemented in order to achieve a more uniform random sampling of candidate solutions in the initial population. The idea behind LHS is to improve sampling by removing the clustering that can arise when using a simple random number generator. The best way to understand why such clustering can occur in random systems, is to think about the toss of a coin. Since each toss of the coin is independent, it is just as statistically likely to get two tails in a row as it is to flip heads then tails. In a similar way, random number generators generate new sample points without taking into account previously generated points. This method becomes inefficient when clustering of sample points occurs, since it leaves larger areas of the search space unexplored.[42]

With LHS, however, subsequently generated sample points are constrained to a range which is dependent on the previously selected points. In the context of statistical sampling, a square grid containing sample positions is a Latin square if, and only if, there is only one sample point in each row and each column. LHS is the generalization of this concept to an arbitrary number of dimensions, where there is only one sample point in each axis-aligned hyperplane. This is analogous

to a game of chess with N rooks on the board that are of no threat to each other. In other words, LHS aims to spread the sample points more evenly across all possible values by portioning each input distribution into N intervals of equal probability, and selecting one sample from each interval. In this way, a more uniform set of solutions is generated in the first generation. This is turn give the GA finds a better chance of finding a diverse of solutions in the first generation, thereby reducing the total number of iterations needed for the GA to converge.[41] An illustration of both random sampling and LHS is given in Figure 2.10, where the impact of LHS is visually evident in the probability densities shown for either sampling method.

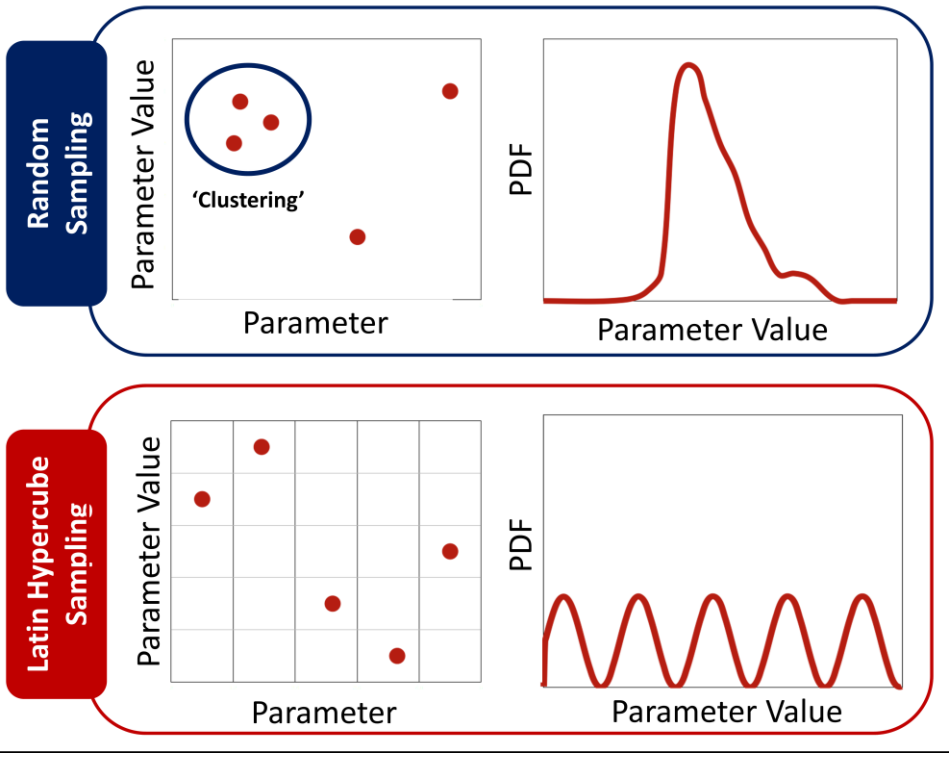

Figure 2.10 *Comparison of Random Sampling and Latin Hypercube Sampling for a 2-dimensional problem with 4 sample points. In the GA, each sample point would correspond to a parameter value (allele) in the corresponding chromosome.*

2.4.3 Parameter Refinement using a Gradient Descent (GD) Optimization

Although a GA is a global search algorithm, it is not assured that the GA will attain precisely the global minimum (Figure 2.11). Therefore, it is general practice to refine the parameters yielded by the GA using a traditional optimization technique such as a gradient descent (GD). As the name of this technique indicates, the working principle behind a GD is to use the gradient of the cost function with respect to the parameter value to shift the parameters in the direction of the greatest descent, or minimum error (for minimization problems). This process is repeated iteratively until a set tolerance in the change of the fitness of the solution is achieved. Thus, the GD serves the purpose of ensuring that the parameters obtained are fully optimized, as shown in Figure 2.11.

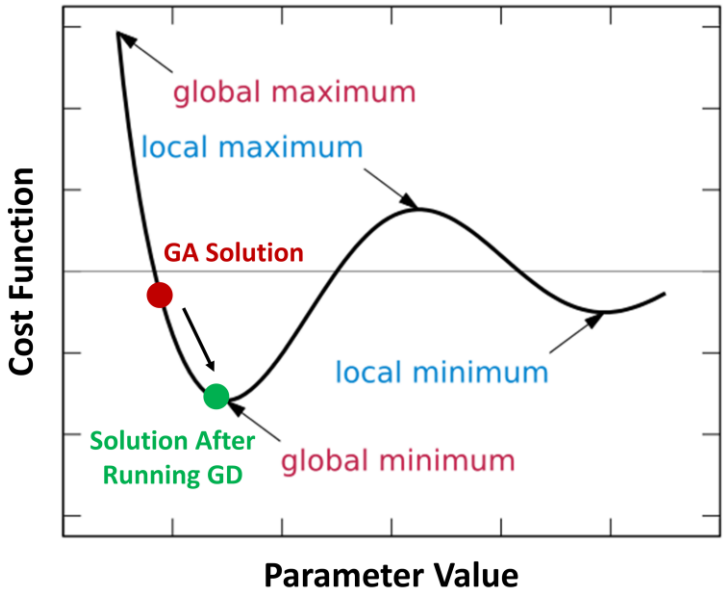

Figure 2.11 *Illustration of GA parameter refinement using a GD.*

To better understand the gradient descent algorithm, consider a scalar function, $f(x)$, with $x \in \mathbb{R}^n$. The gradient of such a function, $\partial f(x)/\partial x$, at a point x points toward a direction where the function is increasing. Conversely, when the gradient is negative at a point x it is generally referred to as the *direction of steepest descent*. The plain gradient descent method for finding the

minimum of $f(x)$ starts from an initial point x_0, then iteratively takes a step along the direction of steepest descent, until convergence. In the algorithm, the size of the step taken can be optionally scaled to ensure that the minimum is not passed over. Hence, keeping the step size constant is not generally desired, and most GD algorithms (including the one used for this thesis) will have some sort of step size adaptation implemented.

Hence, the first drawback of a GD algorithm is the step size, which could be arbitrarily chosen to be any given value.[44] However, selecting too large or too small of a step size could result in decreased efficiency of the algorithm, or even prevent the algorithm from ever reaching a minimum. In fact, the more one thinks about such examples, the more one would conclude that the magnitude of the gradient should not be trusted, and used, as the step size. Rather, one should only use the directional information provided by the gradient. In this sense, robust GD algorithms must permanently rescale the step size empirically depending on the local properties of the function. This requires some apriori knowledge of the range over which the expected solution should be found. In general, application of step size adaptation scheme will allow the algorithm to find an appropriate step size through trial and error within the first couple of iterations. In this regard, there are two simple heuristics that should be applied to any good GD algorithm: (1) when an iteration has been completed and the function value increases, the step size was too large; revert back to last step and decrease the step size, (2) when an iteration has been completed and the function value decreases, perhaps the step could have been larger: increase the step size. Reversion to a previous 'better' solution is a process known as *back propagation*, and is used in all robust optimization techniques.[45] The scheme presented below summarizes the steps taken in a gradient descent algorithm with both step size adaptation and back propagation implemented.

Input: starting point $x \in \mathbb{R}^n$, a scalar function $f(x)$, and its gradient $g(x) = \partial f(x)/\partial x$, initial step size α and tolerance θ.

Output: some x, hopefully minimizing $f(x)$

1. Initialize $f_x = f(x) \in \mathbb{R}^n,\ g = \partial f(x)/\partial x \in \mathbb{R}^n$
2. $y \leftarrow x - \alpha\, g/|g|, f_y \leftarrow f(y)$
3. If $f_y < f_x$ then:
   $y \leftarrow x,\ f_y \leftarrow f_x$
   $g \leftarrow \partial f(x)/\partial x$
   $\alpha \leftarrow l\alpha \quad l \geq 1$ *Increase Step Size*

   else:
   $x \leftarrow x,\ f_x \leftarrow f_x$ *Back Propagation*
   $g \leftarrow \partial f(x)/\partial x$
   $\alpha \leftarrow l\alpha \quad 0 < l < 1$ *Decrease Step Size*

   end if:
   $|y - x| < \theta$ for 10 iterations in sequence

Furthermore, depending on the shape of the function, the GD will become increasingly slow as it approaches the minimum or may not be applicable to the relevant optimization problem. For example, with poorly conditioned convex problems the GD will exhibit a 'zigzag-like' behaviour as the gradients point nearly orthogonally to the shortest direction to a minimum point. Further, for non-differentiable functions, the gradient descent is ill-defined and an alternative method, such as the subgradient projection method, must be used.[46] The function being minimized is generally some sort of cost-function, which measures the error in the predicted solution relative to a target solution, such as experimental data, and so care should also be taken in selecting an error function that properly differentiates between good and bad solutions. It should be noted that in almost all cases, the explicit form of the function being minimized is not known, and the gradients must be calculated numerically using the following expression for calculating the gradient centred between two points:

$$g(x) = \frac{f(x + \alpha) - f(x - \alpha)}{2\alpha} \quad (2.25)$$

The final, and probably most important drawback of GD algorithms is the selection of the starting point.[47] As was mentioned in the outset of this section of the chapter, the GD algorithm is not a global search algorithm, and the solution which it converges upon may not necessarily be a global minimum. This tendency of the GD to converge to a local minimum depends strongly on the point at which the optimization begins. For example, in Figure 2.11 if the starting point was chosen to be the local maximum, the GD would then have an equal chance of converging to the local minimum as to the global minimum. In many cases it is hard to avoid this issue, and the only solution is to run the GD multiple times from different starting points. This is why the GD is used in this work as a means of refining parameters generated by a global search algorithm, namely the GA. Nonetheless, GD algorithms have proven useful in many applications including ontological studies in geological settings, development of neural network parameters as well as other machine learning applications.[48–50]

2.5 References

(1) Koch, W.; Holthausen, M. C. *A Chemist's Guide to Density Functional Theory*, 2nd ed.; Wiley, 2001.

(2) Dronskowski, R. *Computational Chemistry of Solid State Materials*, 1st ed.; Wiley, 2007.

(3) Tirado-Rives, J.; Jorgensen, W. L. Performance of B3LYP Density Functional Methods for a Large Set of Organic Molecules. *J. Chem. Theory Comput.* **2008**, *4* (2), 297–306.

(4) Becke, A. D. Density-functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98* (7), 5648–5652.

(5) Cohen, A. J.; Mori-Sánchez, P.; Yang, W. Challenges for Density Functional Theory. *Chem. Rev.* **2012**, *112* (1), 289–320.

(6) Dronskowski, R. *Computational Chemistry of Solid State Materials*, 1st ed.; Wiley, 2007.

(7) Francl, M. M.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A. Self-Consistent Molecular Orbital Methods. A Polarization-Type Basis Set for Second-Row Elements. *J. Chem. Phys* **1982**, *77*, 3654–3665.

(8) Kresse, G.; Hafner, J. *Ab Initio* Molecular Dynamics for Liquid Metals. *Phys. Rev. B* **1993**, *47* (1), 558–561.

(9) Kresse, G.; Hafner, J. *Ab Initio* Molecular-Dynamics Simulation of the Liquid-Metal–amorphous-

Semiconductor Transition in Germanium. *Phys. Rev. B* **1994**, *49* (20), 14251–14269.

(10) Kresse, G.; Furthmüller, J. Efficient Iterative Schemes for *Ab Initio* Total-Energy Calculations Using a Plane-Wave Basis Set. *Phys. Rev. B* **1996**, *54* (16), 11169–11186.

(11) Kresse, G.; Marsman, M. Vienna Ab-initio Simulation Package: VASP the GUIDE http://cms.mpi.univie.ac.at/vasp/vasp.pdf (accessed May 19, 2016).

(12) Soler, J. M.; Artacho, E.; Gale, J. D.; Garcia, A.; Junquera, J.; Ordejon, P.; Sanchez-Portal, D. The SIESTA Method for Ab Initio Order-N Materials Simulation. **2001**.

(13) Blöchl, P. E. Projector Augmented-Wave Method. *Phys. Rev. B* **1994**, *50* (24), 953–979.

(14) Perdew, J.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865–3868.

(15) Paier, J.; Hirschl, R.; Marsman, M.; Kresse, G. The Perdew-Burke-Ernzerhof Exchange-Correlation Functional Applied to the G2-1 Test Set Using a Plane-Wave Basis Set. *J. Chem. Phys.* **2005**, *122* (23).

(16) Campan, C.; Mussard, B.; Woo, T. K. Electrostatic Potential Derived Atomic Charges for Periodic Systems Using a Modified Error Functional. **2009**, 2866–2878.

(17) Hamad, S.; Balestra, S. R. G.; Bueno-Perez, R.; Calero, S.; Ruiz-Salvador, A. R. Atomic Charges for Modeling Metal–organic Frameworks: Why and How. *J. Solid State Chem.* **2015**, *223*, 144–151.

(18) Collins, S. P.; Woo, T. K. Split-Charge Equilibration Parameters for Generating Rapid Partial Atomic Charges in Metal–Organic Frameworks and Porous Polymer Networks for High-Throughput Screening. *J. Phys. Chem. C* **2017**, *121* (1), 903–910.

(19) Ray, K. G.; Olmsted, D. L.; Burton, J. M. R.; Houndonougbo, Y.; Laird, B. B.; Asta, M. Gas Membrane Selectivity Enabled by Zeolitic Imidazolate Framework Electrostatics. *Chem. Mater.* **2014**, *26* (13), 3976–3985.

(20) Sutrisno, A.; Terskikh, V. V; Shi, Q.; Song, Z.; Dong, J.; Ding, S. Y.; Wang, W.; Provost, B. R.; Daff, T. D.; Woo, T. K.; Huang, Y. Characterization of Zn-Containing Metal-Organic Frameworks by Solid-State 67 Zn NMR Spectroscopy and Computational Modeling. *Chem. Eur. J.* **2012**, *18* (39), 12251–12259.

(21) Morris, W.; He, N.; Ray, K. G.; Klonowski, P.; Furukawa, H.; Daniels, I. N.; Houndonougbo, Y. A.; Asta, M.; Yaghi, O. M.; Laird, B. B. A Combined Experimental-Computational Study on the Effect of Topology on Carbon Dioxide Adsorption in Zeolitic Imidazolate Frameworks. *J. Phys. Chem. C* **2012**, *116* (45), 24084–24090.

(22) Burtch, N. C.; Jasuja, H.; Dubbeldam, D.; Walton, K. S. Molecular-Level Insight into Unusual Low Pressure CO $_2$ Affinity in Pillared Metal–Organic Frameworks. *J. Am. Chem. Soc.* **2013**, *135* (19), 7172–7180.

(23) Vaidhyanathan, R.; Iremonger, S. S.; Shimizu, G. K. H.; Boyd, P. G.; Alavi, S.; Woo, T. K. Direct Observation and Quantification of CO2 Binding Within an Amine-Functionalized Nanoporous Solid. *Science (80-.).* **2010**, *330* (6004), 650–653.

(24) Taylor, J. M.; Vaidhyanathan, R.; Iremonger, S. S.; Shimizu, G. K. H. Enhancing Water Stability of Metal−Organic Frameworks via Phosphonate Monoester Linkers.

(25) Watanabe, T.; Manz, T. a.; Sholl, D. S. Accurate Treatment of Electrostatics during Molecular Adsorption in Nanoporous Crystals without Assigning Point Charges to Framework Atoms. *J. Phys. Chem. C* **2011**, *115* (11), 4824–4836.

(26) Furness, P. Applications of Monte Carlo Simulation in Marketing Analytics. *J. Direct, Data Digit. Mark. Pract.* **2011**, *13* (2), 132–147.

(27) Mode, C. J. *Applications of Monte Carlo Methods in Biology, Medicine and Other Fields of Science*; InTech, 2011.

(28) Binder, K. Applications of Monte Carlo Methods to Statistical Physics. *Reports Prog. Phys.* **1997**, *60* (5), 487–559.

(29) Bart Verberck. *Characterizing Molecular Rotations Using Monte Carlo Simulations.*; INTECH Open Access Publisher, 2011.

(30) Moebs, W. D. C. A Monte Carlo Simulation of Chemical Reactions. *Math. Biosci.* **1974**, *22*, 113–120.

(31) Frenkel, Daan, Smit, B. *Understanding Molecular Simulations*; Academic Press: San Diego, 2002.

(32) Duren, T.; Bae, Y. S.; Snurr, R. Q. Using Molecular Simulation to Characterise Metal-Organic Frameworks for Adsorption Applications. *Chem. Soc. Rev.* **2009**, *38* (5), 1237–1247.

(33) Rappe, A. K.; Goddard, W. A. Charge Equilibration for Molecular-Dynamics Simulations. *J. Phys. Chem.* **1991**, *95* (8), 3358–3363.

(34) Müser, M. H.; Mosey, N.; Mosey, N. J. A Generalization of the Charge Equilibration Method for Nonmetallic Materials. **2006**, *125*.

(35) Chelli, R.; Procacci, P.; Righini, R.; Califano, S. Electrical Response in Chemical Potential Equalization Schemes. *J. Chem. Phys.* **1999**, *111* (18), 8569–8575.

(36) Chambers, L. Practical Handbook of Genetic Algorithms. *Van Nostrand Reinhold* **1995**, *II*.

(37) Vavak, F.; Fogarty, T. C. Comparison of Steady State and Generational Genetic Algorithms for Use in Nonstationary Environments. In *Proceedings of IEEE International Conference on Evolutionary Computation*; IEEE; pp 192–195.

(38) Elbeltagi, E.; Hegazy, T.; Grierson, D. Comparison among Five Evolutionary-Based Optimization Algorithms. *Adv. Eng. Informatics* **2005**, *19* (1), 43–53.

(39) Thierens, D.; Goldberg, D. Convergence Models of Genetic Algorithm Selection Schemes. *Parallel Probl. Solving from Nature—PPSN III* **2005**, 119–129.

(40) Rawlins, G. J. E. *Foundations of Genetic Algorithms. Volume 1*; Elsevier, 1991.

(41) Viana, F. A. C. Things You Wanted to Know about the Latin Hypercube Design and Were Afraid to Ask. In *10th World Congress on Structural and Multidisciplinary Optimization*; GE Global Research: Orlando, Florida, USA, 2013.

(42) Iman, R. L.; Conover, W. J. A Distribution-Free Approach to Inducing Rank Correlation among Input Variables. *Commun. Stat. - Simul. Comput.* **1982**, *11* (3), 311–334.

(43) McKay, M. D.; Beckman, R. J.; Conover, W. J. A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics* **1979**, *21* (2), 239.

(44) Dai, Y. H.; Yang, X. Q. A New Gradient Method with an Optimal Stepsize Property. *Comput. Optim. Appl.* **2006**, *33*, 73–88.

(45) Rojas, R. *The Backpropagation Algorithm*; Springer-Verlag, 1996.

(46) Kiwiel, K. C. Convergence and Efficiency of Subgradient Methods for Quasiconvex Minimization. *Math. Program.* **2001**, *90* (1), 1–25.

(47) Bengio, Y.; Simard, P.; Frasconi, P. Learning Long-Term Dependencies with Gradient Descent Is Difficult. *IEEE Trans. Neural Networks* **1994**, *5* (2), 157–166.

(48) Wu, H.; Zhao, B.; Gao, W. Analysis of Gradient Descent Ontology Iterative Algorithm for Geological Setting. *Geol. Ecol. Landscapes* **2017**, *1* (1), 41–46.

(49) Fernando, C.; Banarse, D.; Blundell, C.; Zwols, Y.; Ha, D.; Rusu, A. A.; Pritzel, A.; Wierstra, D. PathNet: Evolution Channels Gradient Descent in Super Neural Networks. 2017.

(50) Bottow, L. Large-Scale Machine Learning with Stochastic Gradient Descent. *Lechevallier Y., Saporta G. Proc. COMPSTAT'2010* **2010**, *2*, 94–101.

3 Atomic Charges for Periodic Frameworks with a Net Charge

3.1 Summary

The work presented in this chapter was taken from the manuscript of a paper that was published in the *Journal of Chemical Theory and Computation (JCTC)* on May 11, 2017. The content of the manuscript itself has been left unchanged, however there have been formatting changes made to ensure consistency of this chapter with all other chapters of this thesis including changing numbering of tables, figures and equations. My contribution to this work was in the testing and validation of both the REPEAT method the novel SQE_{AB} method for use with charged periodic systems. This involved generating the appropriate structures, including the strategic placement of counter-ions in large pore MOFs POST-1 and IRMOF-16. The REPEAT charges were then calculated and compared for these systems with and without the charge neutralizing ions placed. Additionally, Al atoms had to be placed in all zeolite structures used in this work as all crystallographic data for zeolites only includes Si atoms in the frameworks due to the experimental inability to differentiate Al and Si by X-ray diffraction. To do this, I wrote an algorithm that randomly places Al atoms in the framework according to the Löwenstein rule. To show the effect of counter-ions on the framework atomic charges, ions were placed within a small pore zeolite, CAS, which had 4 Al atoms in the unit cell, placed using this algorithm. Again, the REPEAT charges were calculated for the charged CAS framework with and without the inserted counter-ions. Furthermore, the algorithm was applied towards generating a training set that I then used to optimize parameters for QEq, SQE and SQE_{AB}. The frameworks populating this small training set was a representative subset of frameworks taken from a larger training set that will be presented in Chapter 4. As a final test of the REPEAT ESP charges derived from periodic systems with a net charge, we used the charges to evaluate a physical observable. To do so, I calculated the REPEAT charges on a MOF where the counter-ion positions were known, and used these charges to predict the positions of the counter-ions using an MC simulation. These results where then compared to those obtained when using QTAIM, or Bader charges. It should be noted that I did not contribute to the mathematical derivation and implementation of the SQE_{AB} model. This was done by a post-doctorate in our research group, Mykhaylo Krykunov.

A New Split Charge Equilibration Model and REPEAT Electrostatic Potential Fitted Charges for Periodic Frameworks with a Net Charge

Mykhaylo Krykunov, Christopher Demone, Jason W.-H. Lo, Tom K. Woo*

Centre for Catalysis Research and Innovation

The Department of Chemistry and Biomolecular Science

University of Ottawa, Ottawa, Canada

3.2 Abstract

Periodic frameworks that possess a net charge, such as zeolites, are an important class of materials in wide use. For guest−host interactions to be simulated in these materials, partial atomic charges are often used. In this work, we investigate two methods for the generation of partial atomic charges in periodic systems having a net framework charge. We first examine the validity of generating REPEAT electrostatic potential fitted charges derived from periodic electronic structure calculations, where a constant background charge is added to neutralize the net charge on the framework. The constant background charge obviates the need to add neutralizing counter-ions, which may induce artifacts such as polarization in the infinite periodic system. The second method we explore is the split charge equilibration (SQE) method for the rapid generation of partial atomic charges. The original formulation of the SQE method cannot be applied to systems with a net charge. In this work, we reformulate the SQE method by transforming the split charges into an atomic charge basis that allows for non-neutral systems to be treated. The new SQE model, which we call SQE_{AB} (for atomic basis), was validated with a series of tests using both charged and neutral

metal organic frameworks and zeolites. It was shown that SQE_{AB} gives equivalent results to those of the original SQE model for neutral systems. We then demonstrated that the SQE_{AB} method is able to "capture" the chemical structure of a charged framework better than that of the charge equilibration model by comparing to REPEAT electrostatic potential fitted charges.

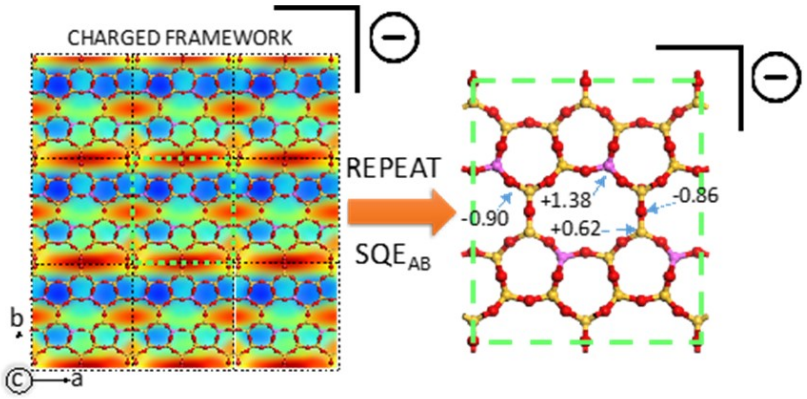

Figure 3.1 *Generation of partial atomic charges in periodic frameworks with a net charge from the QM ESP.*

3.3 Introduction

Nanoporous materials such as zeolites, metal-organic frameworks (MOFs), covalent organic frameworks (COFs) and porous polymer networks (PPNs) have attracted significant scientific attention as enabling materials for catalysis, gas capture and storage, separations, fuel cell membranes, drug delivery to name a few. MOFs are made by combining inorganic and organic linking groups to form crystalline solids and have seen an explosive growth in the literature with over 20,000 different MOFs having been reported and studied in the past decade alone.[1] Although not as structurally diverse as MOFs, zeolites have been developed for a longer period and have found their way into many large-scale industrial applications. For instance, zeolites are used to catalytically refine all of the gasoline that is produced and are also employed as a solid sorbent in large pressure swing adsorption units used to scrub CO_2 in natural gas processing.[2]

Interestingly, many commonly used nanoporous materials possess frameworks that have a net charge, which is balanced by counter ions that intercalate through the pores. For example, virtually all zeolites that are used in practice have a proportion of the Si sites replaced with Al, which bestows the framework with a negative charge. In other words, there are no purely siliceous zeolites with neutral frameworks that are in wide use. The zeolite ZSM-5 which was one of the first zeolites used in the petroleum industry[2] has a Si to Al ratio as low as ~4:1 which gives the unit cell of the material a -27e charge. MOFs with charged frameworks are also commonly reported, where the net formal charges are often localized to the metal centers. For example, Dincă and Long[3] reported a Mn benzene-tris-tetrazolate MOF, **1**, with extra framework counter ions which showed good hydrogen storage capacity. Neutron diffraction studies revealed strong H_2 binding sites on the charged Mn centers of the MOF. In the case of MOFs being developed for proton conducting membranes for fuel cells,[4,5] the acidic sites on the MOF organic groups play a fundamental role in the proton conduction mechanism and there have also been reports of the extra framework ions in charged MOFs aiding in the conduction mechanism.[6]

As in many areas of science and engineering, computer simulation has developed into a powerful tool to examine the physical and chemical processes involving nanoporous materials.[7] To study materials at the molecular scale, classical simulations that utilize empirical force fields to evaluate the potential energy surface are in wide use. Classical molecular dynamics and Monte Carlo simulations have been used extensively to study guest host interactions in nanoporous materials. Recently, Smit and coworkers[8] have used Monte Carlo simulations to screen the methane storage capability of over 650,000 MOFs, zeolites, and PPNs. In most classical simulations, atomic pair potentials are used to model the non-bonded guest-host interactions, which are typically broken up into a van der Waals component and an electrostatic component. The electrostatic component is commonly accounted for

by a fixed partial atomic charge approximation. Since electrostatic interactions are strong and long ranged, the assignment of the charges is then critical to the accuracy of the simulations. Atomic charges that are fit to reproduce the electrostatic potential derived from a first principles quantum mechanical (QM) calculation are most often used for this purpose. The general procedure for calculating these electrostatic potential fitted charges, or ESP charges, is to perform a QM calculation on a suitable conformation of a molecule and fit the partial atomic charges on each of the atoms as to minimize the difference between the QM electrostatic potential and those due to the point charges on a set of grid points surrounding the molecule. Although ESP charges have been used for decades in the atomistic simulation of sd, it was not until relatively recently that a similar formulation was developed to extract ESP charges from periodic QM calculations. The technical problem is that the electrostatic potential of an infinitely periodic system is ill-defined and periodic QM methods can have an arbitrary offset to the electrostatic potential. The first ESP charge method to overcome this problem was the REPEAT method[9] where the ESP off-set value, δ_ϕ, is considered as another variable in the fit. In this treatment the charges are fit to minimize the relative differences in the ESP rather than the absolute values. The REPEAT method[9,10] and its enhancements[11] have become the method of choice for deriving partial atomic charges from periodic QM calculations for use in molecular simulation.

One drawback of ESP charges is that a relatively expensive electronic structure calculation is required. While this isn't typically a problem when studying a few materials, it can be problematic when screening hundreds of thousands or millions of compounds.[8,12] One approach to rapidly generate partial atomic charges is to use a computationally efficient electronegativity equalization method, the most popular of which is the Charge Equilibration (QEq) method of Rappé and Goddard.[13] The QEq method allows for charges to be generated in seconds for large systems, but it is also highly parameterized with atomic hardness and electronegativity parameters required for each

element or atom type. Snurr and Wilmer, for example, have used such an approach to generate charges to perform GCMC simulations to screen more than 130,000 MOFs for their CO_2 uptake properties. In order to generate appropriate charges for molecular simulation, one can parameterize the QEq method to reproduce ESP fitted charges. For example, Kadantsev *et al.*[14] developed a set of QEq parameters fit to reproduce the DFT derived ESP charges of a set of 543 MOFs. Using these parameters to evaluate the QEq charges on a set of 693 different MOFs, the CO_2 gas uptake capacities were found to reproduce those calculated with DFT derived REPEAT charges with a Spearman rank-order correlation coefficient of >0.97.

A generalization of the QEq method called the Split Charge Equilibration(SQE) method, has been developed by Müser and co-workers.[15] In addition to the atomic hardness and electronegativity parameters, the SQE method uses atom-pair specific bond hardness and bond electronegativity parameters. If the bond hardness and electronegativity parameters are set to zero, the SQE method elegantly reduces to the QEq method. The SQE method has gained attention as a potentially more flexible alternative to the QEq method and has been used in a variety of applications.[16,17,18,19,20,21,22] In the context of nanoporous materials, Verstraelen and coworkers[23] developed SQE parameters fit to reproduce DFT derived ESP charges (and Hirschfeld charges) of a large number of silicate cluster compounds. These SQE parameters were then tested on a number of compounds including periodic all-silicon zeolites showing good agreement with the DFT derived charges. One reason it was only applied to all-silicon zeolites is that the SQE method has the limitation that it can only be applied to neutral systems. This is due to the fundamental anti-symmetry property of the split charges which constrains the system to a zero net-charge.

Despite the importance of nanoporous materials with charged frameworks, particularly zeolites, we are not aware of any ESP charges that have been derived from periodic QM calculations

of frameworks with a net charge. Recently, Fischer and Bell[24] attempted to develop generic charges and Lennard-Jones parameters for zeolites where the charges were fit to REPEAT ESP charges. However, the charges were derived from neutral, Si only zeolites. A number of empirical force fields that have been developed for simulating zeolites, including Al containing charged frameworks, however, the partial atomic charges used in these models were not derived from periodic QM calculations of the Al containing zeolites. The partial atomic charges on the Si, Al and O atoms from some of these force fields are given in Table 3.1. It is interesting to note that there is a large variation in the charges used in these zeolite force fields.

Table 3.1 Partial atomic charges used in the simulation of zeolites

		partial atomic charge (e)			
author (year)	zeolite	Si	Al	O_{Si}[a]	O_{Al}[a]
Losinger *et al.* (1991)	generic	+1.93	+1.02	-0.49	-0.49
Hirotani *et al.* (1997)	ZSM-5 Si/Al=94	+0.783	+0.587	-0.40	-0.40
	ZSM-5 Si/Al=46	+0.766	+0.575	-0.40	-0.40
Faux *et al.* (1997)		+3.70	+2.775	-1.869	-1.869
Jaramillo *et al.* (1999)		+2.05	+1.75	-1.025	-1.20
Calero (2009)	generic	+0.786	+0.486	-0.393	-0.414
Fang *et al.* (2013)	ZSM-5	+2.21	+2.08	+1.105	-1.32
Fischer *et al.*[24] (2013)	generic	+1.034	-	-0.5017	-

[a] *O_{Si} refers to an oxygen bonded only to two Si atoms. O_{Al} refers to an oxygen bonded to one Al atom and one Si atom.*

In charged frameworks, the location of the counter ions is often not known experimentally, as in the case of zeolites. Here, molecular simulation could be used to determine the location of the counter ions. The counter ions can also be highly fluxional and therefore must be allowed to move during the simulation. In both cases, the partial atomic charges assigned to the framework are critical to the simulations. To derive ESP charges from periodic QM calculations of frameworks that possess a net charge, one can perform the QM calculation with the counter ions positioned somewhere in the pores in order to neutralize the overall charge. However, the resulting framework charge distribution

will be polarized towards the counter ions. When using these polarized charges in simulations where the counter ions are then free to move, the simulation is likely to bias positioning the counter ions in these pre-determined locations. Deriving ESP charges using a counter ion may also be problematic in smaller pore materials since there may be very few ESP fitting points between the counter ion and the framework atoms, leading to 'unphysical' charges that may reproduce the ESP well, but which may not be transferrable when the counter-ion is removed. This is the well-known buried atom problem in ESP fitted charge methods.[25] Although restraints can be imposed to prevent the 'unphysical' charges from arising, this will not prevent the ions from polarizing the framework. Thus, it would be ideal to perform the ESP fit without the counter ions in the periodic QM calculation. However, without the counter ion, the simulation cell will possess a net charge and the infinitely periodic system will have an infinite net charge and infinite energy.

As first presented, the SQE method can only treat neutral systems. This is a manifestation of the anti-symmetry property of the so-called split charges such that they necessarily sum to zero, restricting the total charge to zero. To overcome this limit, Verstraelen and co-workers[26] took the net charge of the molecule, Q_{tot}, and distributed the excess charge equally over all atoms of the system. A natural generalization of this approach is the SQE+Q^0 model,[17] which involves an unequal distribution of the excess charge among atoms or groups justified by some reasonable physical assumptions. SQE+Q^0 model has been applied to zwitterions.[27] However, these results serve more as "a proof of concept" rather than a practical scheme, because the excess charge distribution has not been derived from a minimization procedure. On the other hand, the excess charge distribution that differs from a constant value can be easily obtained in models that are formulated in an atomic basis such as the QEq or QTPIE model.[28] The latter model is initially derived in the bond representation similar to the SQE model. However, unlike SQE, the QTPIE model involves a transformation to the atomic basis

from the bond representation with the help of a pseudo inverse. The transformation from the split-charge representation (equivalent to the one in SQE) to the atomic representation is also employed in the ACKS2 model.[29] However, this model is formally equivalent to the SQE+Q^0 model, if the same parameters are employed. We note that neither the ACKS2 or the QTPIE methods have been applied to the charged and/or periodic systems.

In this work, our goal is to address a few gaps in the literature we see when deriving partial atomic charges for periodic frameworks with a net charge. First, we show that the REPEAT method can be used to derive atomic charges of periodic frameworks with net charges without the need for adding explicit counter ions in the periodic QM calculation. This is achieved by adding a constant neutralizing background charge[30,31] to the periodic QM calculation. Next, we present a new formulation of the SQE method for charged systems that will allow for rapid partial atomic charges to be generated for Al containing zeolites and other materials with charged frameworks.

3.4 Methodology and Background Theory

3.4.1 REPEAT Charges for Periodic Systems with a Net Charge

The typical formalism to obtain ESP charges for molecular systems involves the minimization of the difference in the electrostatic potential obtained from a QM calculation of the system and those obtained from the partial atomic charges, q_i, on a set of grid points surrounding the molecule. In its simplest form the minimization function is given by:

$$F(\{q_j\}) = \sum_{grid} (\phi_{QM}(\vec{r}_{grid}) - \phi_q(\vec{r}_{grid}))^2 + \lambda(\sum_j q_j - q_{tot}) \qquad (3.1)$$

where ϕ_{QM} and ϕ_q are the electrostatic potential due to the QM wave function (and nuclei) and the point charges, respectively, λ is a Lagrange multiplier used to constrain the atomic charges to sum to

the total system charge, q_{tot}, and the first summation runs over the grid points and the second summation runs over all atoms. For periodic systems, there is an additional complication associated with the fact that the electrostatic potential in an infinitely periodic system is ill defined. As a result, the ESP obtained from a periodic QM calculation can be offset by an arbitrary, but fixed amount. To overcome this problem, one can introduce an offset, δ_ϕ that is adjusted just like the charges to minimize a modified REPEAT error function:

$$F(\{q_j\},\delta_\phi)=\sum_{grid}\left(\phi_{QM}(\vec{r}_{grid})-(\varphi_q(\vec{r}_{grid})+\sigma_\phi)\right)+\lambda\left(\sum_j q_j-q_{tot}\right) \qquad (3.2)$$

Minimization of the error function Eq. (3.2) with respect to δ_ϕ, gives the following expression for δ_ϕ:

$$\delta_\phi=\frac{1}{N_{grid}}\sum_{grid}\left(\phi_{QM}(\vec{r}_{grid})-(\varphi_q(\vec{r}_{grid})\right) \qquad (3.3)$$

which reveals that the ESP fit is performed to minimize the relative differences with respect to the average value of the potential on the grid points. This resolves the problem of the arbitrary offset in the ESP, allowing meaningful ESP charges to be derived from periodic QM calculations. In fact, it can be shown that the REPEAT charges reduce to the conventional 'molecular' ESP charges in the limit of an infinitely large simulation box.[9]

In periodic QM calculations, one simulates an infinite periodic crystal via the use of periodic boundary conditions. If the unit cell has a net charge, then the system will have an infinite charge and the energy of the periodically repeated system will diverge. In order to investigate charged defects or contaminants in solids, Leslie and Gillan[32] first proposed adding a constant background charge to neutralize the cell's net charge. This is equivalent to immersing the system in a jellium background which fills the simulation cell so that the net charge is zero.[30] Most periodic QM packages employ this

technique to treat charged systems. The constant background charge that is added has the effect of offsetting the QM electrostatic potential by a constant amount. Thus, the original REPEAT method will naturally account for this offset, and as a result, meaningful ESP fitted charges should be obtained for charged periodic systems, just as they are with neutral periodic systems.

3.4.2 Split Charge Equilibration Method for Charged Frameworks

The split charge equilibration method, first proposed by Müser and co-workers,[15] is a generalization of the QEq method. In the QEq method the potential energy of a system consisting of N atoms is written as a function of effective atomic charges, Q_i:

$$E_Q^{QEq} = \sum_i^N \left(\frac{1}{2} \kappa_i Q_i^2 + \chi_i Q_i \right) + \frac{1}{2} \sum_i^N \sum_{j \neq i}^N J(r_{ij}) Q_i Q_j \quad (3.4)$$

where κ_i and χ_i are the electronic hardness and electronegativity of atom i, respectively. The electrostatic potential, $J(r_{ij})$, between the atoms i and j depends on the interatomic distance, r_{ij}. For a molecular system, $J(r_{ij})$ is given as a standard Coulomb potential between point charges while in the case of periodic systems, N in Eq. (3.4) corresponds to the number of atoms in the unit cell, and the Coulomb potential is evaluated with the Ewald summation[33] or equivalent. In both the molecular and periodic case, it is worth noting that the potential $J(r_{ij})$ is modified for atoms at close distances to mimic the interaction of bonded atoms. With a given set of atomic κ_i and χ_i parameters, the atomic charges of the system are found from the minimization of the energy expression of Eq. (3.4). In order to introduce the constraint on the net charge, Q_{net}, the energy expression Eq. (3.4) is augmented with the Lagrange multiplier, λ, as given in Eq. (3.5).

$$L^{\mathrm{QEq}}(\mathbf{Q}, Q_{net}) = E^{\mathrm{QEq}}(\mathbf{Q}) - \lambda \left(\sum_i^N Q_i - Q_{net} \right) \quad (3.5)$$

In the SQE method, the atomic charges are defined in terms of the so-called split charges:

$$Q_i = \sum_j^N q_{ij} \qquad (3.6)$$

where the "split charge" q_{ij} represents the charge flow from a covalently bonded neighbour atom j to atom i. In Eq. (3.5), one normally only sums over the number of bonded neighbour atoms, n_i, but for convenience we sum over all atoms N allowing for the fact that q_{ij} is zero for non-neighbour atom pairs. After the substitution of the split charges Eq. (3.6) into Eq. (3.4), the potential energy can be rewritten as given in Eq. (3.7).

$$E^{\mathrm{QEq}}(\mathbf{Q}) = \sum_i^N \sum_j^N \chi_i q_{ij} + \tfrac{1}{2}\sum_i^N \sum_j^N \sum_k^N \kappa_i q_{ij} q_{ik} + \tfrac{1}{2}\sum_i^N \sum_{j\neq i}^N J(r_{ij}) \sum_k^N \sum_l^N q_{ik} q_{jl} \qquad (3.7)$$

So far, this ansatz does not bring any advantages, since both Eq. (3.4) and Eq. (3.7) yield identical atomic charges for single molecules. (The previous statement is true for the connected graphs, while for the non-connected ones, such as the case of two molecules at a large distance, the model given by Eq. 3.7 will yield two neutral species unlike QEq given by Eq. 3.4. We shall consider in this work only connected graphs. The SQE method differentiates itself from the QEq method by the introduction of bond dependent parameters, κ_{ij}^b and χ_{ij}^b, as a generalization of the atomic hardness and electronegativity counterparts, respectively. (For a discussion of the physical interpretation of the parameters see the work of Verstraelen *et al.*[29]) Thus, in addition to element only or atom-type only parameters, κ_i and χ_i, the SQE method employs bond pair specific parameters. Introducing these terms into Eq. 3.7 gives the SQE potential energy as shown in Eq. 3.8.

$$E^{\mathrm{SQE}}(\mathbf{Q}) = \sum_i^N \sum_j^N \chi_i q_{ij} + \tfrac{1}{2}\sum_i^N \sum_j^N \sum_k^N \kappa_i q_{ij} q_{ik} + \sum_i^N \sum_{j>i}^N \left(\tfrac{1}{2}\kappa_{ij}^b q_{ij} + \chi_{ij}^b\right) q_{ij} + \tfrac{1}{2}\sum_i^N \sum_{j\neq i}^N J(r_{ij}) Q_i Q_j \qquad (3.8)$$

One important property of the split charges is that they satisfy the anti-symmetry condition such that:

$$q_{ij} = -q_{ji}\,. \quad (3.9)$$

As a result, the sum of all charges is always constrained to zero in this formalism.[15] Although this property eliminates the need of a Lagrange multiplier to constrain the total system charge, it also restricts this model to neutral systems.

In order to treat charged systems, one straightforward approach would be to add a local contribution to the partial atomic charge[17] as given in Eq. 3.10, where the local charge corrections, q_i , are treated as independent variables.

$$\tilde{\zeta}_{\iota} \quad {}_{\iota} \cdot \sum_{j}^{N} q_{ij}\,. \quad (3.10)$$

Unfortunately, when this ansatz is substituted into the SQE energy expression (3.8), the modified SQE scheme, which we will call $\mathrm{SQE}_{\mathrm{LCC}}$ for Local Charge Correction, gets reduced back into the QEq model. We show this non-trivial result in Appendix A.

An alternative approach to accommodating net charges within the SQE model is to transform the model into the atomic basis, i.e. the basis of partial atomic charges, Q_i . As an example, consider defining the split charges, q_{ij}, in terms of atomic charges as in Eq. 3.11.

$$q_{ij} = \tfrac{1}{2}\left(Q_i - Q_j\right). \quad (3.11)$$

The substitution of Eq. 14 into Eq. 7 yields the SQE energy expression that depends exclusively on local variables. Unfortunately, the formula (3.11) is valid only for diatomic molecules. In general, it is necessary to invert Eq. (3.5), but the number of covalent bonds does not usually coincide with the number of atoms (even in the case of two atoms there is only one split-charge). As a result, the corresponding system of linear equations would be rank deficient in most cases. To overcome this problem, one can apply the Moore-Penrose pseudoinverse.[34] Below we shall employ such a general

inversion for the reformulation of the SQE model that can accommodate charged systems. We shall call this new model, SQE_{AB}, where AB stands for Atomic Basis.

Let us consider first the transformation of Eq. (3.5) in matrix-vector notation for a neutral system as given by Eq. (3.11).

$$\mathbf{Q} = \mathbf{T}\bar{\mathbf{q}} . \quad (3.12)$$

Here, the transformation matrix, $\mathbf{T}$, has the size N_{atoms} times N_{bonds}, and it consists of three types of elements: $+1$ (if there is a bond between atoms with such indices i and j that $j > i$), -1 (if there is a bond, but $j < i$) and 0 (if there is no bond between atoms). Then, the vector of split-charges, $\bar{\mathbf{q}}$, can be obtained by applying the pseudo inverse, $\mathbf{T}^+$, on the vector of atomic charges

$$\bar{\mathbf{q}} = \mathbf{T}^+\mathbf{Q} \quad (3.13)$$

where

$$\mathbf{T}^+ = \left(\mathbf{T}^\dagger\mathbf{T}\right)^{-1}\mathbf{T}^\dagger . \quad (3.14)$$

Here, $\mathbf{T}^\dagger$ is the conjugate transpose of $\mathbf{T}$, and the inverse of the square matrix, $\mathbf{T}^\dagger\mathbf{T}$, is obtained according to

$$\left(\mathbf{T}^\dagger\mathbf{T}\right)^{-1} = \mathbf{U}\begin{pmatrix}\mathbf{\Lambda}^{-1} & 0\\ 0 & 0\end{pmatrix}\mathbf{U}^{-1} . \quad (3.15)$$

where $\mathbf{U}$ is the unitary matrix composed of eigenvectors and $\mathbf{\Lambda}$ is the diagonal matrix composed of the corresponding non-zero eigenvalues of $\mathbf{T}^\dagger\mathbf{T}$. Note that $\mathbf{T}^\dagger\mathbf{T}$ is the Laplacian of a graph G defined by the incidence matrix $\mathbf{T}$. The Laplacian matrix has the following structure: $\mathbf{L}(G)=\mathbf{D}(G)-\mathbf{A}(G)$, where $\mathbf{A}(G)$ is the adjacency matrix, and $\mathbf{D}(G)$ is the diagonal matrix of vertex degrees. The important feature of the Laplacian spectrum is that it has all non-zero eigenvalues except one (if the graph is connected).[35] Thus, we have to discard only one eigenvalue in Eq. (3.15).

Other important results from graph theory include the relation[36] of the second largest eigenvalue of the Laplacian matrix to the edge (or bond) connectivity,[37] and the relation of the function derived from all eigenvalues of the Laplacian to the Wiener index.[38] Both quantities, i.e. the bond connectivity and Wiener index, are extensively used in cheminformatics applications.[39] Thus, although the pseudo inverse brings an additional cost to the SQE_{AB} method in the form of matrix diagonalization, the by-product of this operation, i.e. the Laplacian eigenvalues, can be used for processing of chemical information.

Once we obtained the pseudo inverse $\mathbf{T}^{+}$, the SQE model of Eq. (3.8) can be expressed in the atomic representation as given by Eq. (3.16).

$$\begin{aligned} E_{\mathrm{AB}}^{\mathrm{SQE}}\left(\mathbf{Q}\right) &= \sum_{i}^{N}\chi_i Q_i + \tfrac{1}{2}\sum_{i}^{N}\kappa_i Q_i^2 + \tfrac{1}{2}\sum_{i}^{N}\sum_{j>i}^{N}\sum_{k}^{N}\sum_{l}^{N}\kappa_{ij}^{b}\left(\mathbf{T}^{+}\right)_{ij,l}\left(\mathbf{T}^{+}\right)_{ij,k} Q_l Q_k \\ &+ \sum_{i}^{N}\sum_{j>i}^{N}\sum_{k}^{N}\chi_{ij}^{b}\left(\mathbf{T}^{+}\right)_{ij,k} Q_k + \tfrac{1}{2}\sum_{i}^{N}\sum_{j\neq i}^{N} J\left(r_{ij}\right) Q_i Q_j \end{aligned} \qquad (3.16)$$

Although the SQE_{AB} energy expression now corresponds to the original SQE model formulated in the split-charges basis, it is necessary to apply constraints to the atomic charges during the optimization procedure in order to maintain the charge neutrality in the atomic basis. We will return to the issue of the equivalence of two models later. Here, we stress that the purpose of reformulating the SQE model in the atomic basis is that it allows us to apply similar constraints to the atomic charges to allow for a non-zero net charge of the system. This gives the Lagrange functional expressed in Eq. (3.17) where **1** is the unit vector and the SQE energy is expressed in terms of the gradient, **g**, (Eq. 3.18) and the Hessian, **H**, (Eq. 3.19).

$$L_{\mathrm{AB}}^{\mathrm{SQE}}\left(\mathbf{Q}, Q_{net}\right) = \mathbf{Q}^T\mathbf{g} + \tfrac{1}{2}\mathbf{Q}^T\mathbf{H}\mathbf{Q} + \lambda\left(\mathbf{Q}^T\mathbf{1} - Q_{net}\right), \quad (3.17)$$

$$g_i = \partial E_{\mathrm{AB}}^{\mathrm{SQE}}\left(\mathbf{Q}\right)/\partial Q_i, \quad (3.18)$$

$$H_{ij} = \partial^2 E_{\mathrm{AB}}^{\mathrm{SQE}}\left(\mathbf{Q}\right)/\partial Q_i \partial Q_j \quad (3.19)$$

Note that the derivatives in Eqs. (3.18) and (3.19) are taken at the point $\mathbf{Q}=0$. The solution of the SQE$_{AB}$ Lagrangian (Eq. 3.17) is given by[40] Eq. 3.20.

$$\begin{pmatrix}\mathbf{Q}\\ \lambda\end{pmatrix}=\begin{pmatrix}-\mathbf{H}^{-1}\left(\mathbf{g}+\lambda\mathbf{1}\right)\\ -\dfrac{Q_{net}+\mathbf{1}^{T}\mathbf{H}^{-1}\mathbf{g}}{\mathbf{1}^{T}\mathbf{H}^{-1}\mathbf{1}}\end{pmatrix} \qquad (3.20)$$

We will now express the SQE$_{AB}$ model for a charged system in terms of the neutral system with Q_{net}=0 where Eq. (3.20) becomes Eq. (3.21).

$$\begin{pmatrix}\mathbf{Q}^{0}\\ \lambda^{0}\end{pmatrix}=\begin{pmatrix}-\mathbf{H}^{-1}\left(\mathbf{g}+\lambda^{0}\mathbf{1}\right)\\ -\dfrac{\mathbf{1}^{T}\mathbf{H}^{-1}\mathbf{g}}{\mathbf{1}^{T}\mathbf{H}^{-1}\mathbf{1}}\end{pmatrix}. \qquad (3.21)$$

Here, $\mathbf{H}$ and $\mathbf{g}$ are the same as in Eq. (20), and $\mathbf{Q}^{0}$ is a vector consisting of the partial atomic charges of a neutral system. Now we can write down the solution for the atomic charges that explicitly depends on the net charge by combining Eqs. (3.20) and (3.21):

$$\mathbf{Q}=\mathbf{Q}^{0}+Q_{net}\frac{\mathbf{H}^{-1}\mathbf{1}}{\mathbf{1}^{T}\mathbf{H}^{-1}\mathbf{1}}. \qquad (3.22)$$

Thus, the atomic charges of the SQE$_{AB}$ model for charged systems is composed of the SQE$_{AB}$ charges of the neutral system plus an additional term that explicitly depends on the net charge and the Hessian. We also note that when SQE$_{AB}$ is applied to periodic systems that the Ewald algorithm implicitly introduces a uniform background charge that effectively neutralizes the simulation cell.[41]

It can also be shown (See Appendix B) that for neutral system the working equations of the SQE$_{AB}$ model given by Eq. 3.21 are conditionally equivalent to the corresponding equations in the split-charge representation of the original SQE model. This condition is often not satisfied in practical applications. Therefore, we have numerically compared the SQE$_{AB}$ and SQE models in the Results and Discussion section.

It is notable that the Moore-Penrose pseudo inverse had been previously applied to in the QTPIE model to transform it from a bond representation to an atomic basis.[28] Although the SQE_{AB} model also employs the Moore-Penrose pseudo inverse for the same purpose, the two models are distinct. The QTPIE model was originally formulated in the bond representation[42] but unlike SQE and SQE_{AB}, the QTPIE model doesn't employ a molecular graph based on covalent bonds. Instead, its bond variables and parameters correspond to a fully connected graph where all atoms are connected to each other. Although these models are similar, such a difference in a pairwise basis leads to distinct mathematical features of two models. For example, linear equations of the QTPIE model in the bond representation are ill defined due to a large number of bond variables. On the other hand, SQE doesn't have such a problem in a much smaller split-charges basis. The QTPIE model in the atomic representation affords a simple analytical formula for the pseudo inverse due to the completeness of a graph. Therefore, a numerical transformation is not necessary in that case.

Even in the case of a simple diatomic molecule, where one might expect the SQE_{AB} and QTPIE models to be equivalent, they are not identical. Namely, the QTPIE model as well as its progenitor PE-CC-QVB2,[43] do not contain the quadratic terms with the bond hardness (see the third term in Eq. 3.16).

3.5 Computational Details

All periodic DFT calculations were performed with the Vienna Ab Initio Simulation Package (VASP)[44] with the projector augmented wave (PAW) method[45] and plane wave basis sets. A plane wave cut-off of 500 eV was used and only the gamma point in the Brillouin-zone was sampled. The Perdew-Burke-Ernzerhof (PBE) exchange and correlation functional[46] was utilized for all DFT calculation. For calculations using an experimental crystal structure, the positions of the hydrogen atoms were optimized.

Monte Carlo (MC) simulations were performed using an in-house code based on DL_POLY 2 molecular dynamics package[47] that we have used to successfully model the gas uptake in MOFs.[48] The atomic positions of the MOF framework were fixed in the simulations. Non-bonding interactions were calculated using Lennard-Jones (LJ) potentials and parameters from the universal force field (UFF)[49] for the MOF and partial atomic charges as described. To determine the positions of the counter ions, we employed a double jump MC move to improve sampling wherein two ions are each moved to a random position within the MOF.

Both the REPEAT and SQE calculations were performed using in-house developed code and are freely available upon request.

3.6 Results and Discussion

To show that the REPEAT ESP fitted charges obtained by adding a constant background charge in the periodic QM calculation are valid, we have performed a series of calculations on charged and neutral systems (Figure 3.2) that should give near identical partial atomic charges for the framework atoms. We will first evaluate the ESP charges of a neutral framework MOF (Figure 3.2a) via a standard neutral periodic QM calculation. These will then be compared to the ESP charges evaluated on the same MOF where the unit cell has been 'charged' with ions that are added into the pores of the framework as shown in Figure 3.2b. The periodic DFT calculation of the 'charged' system will be performed with a constant background charge. If there is no polarization of the framework by the ions, the ESP charges obtained should be the same barring any fitting artifacts. We will also compare the ESP charges of a charged framework MOF (Figure 3.2c) and compare it to the ESP charges derived from a calculation where the framework has been neutralized with ions added to the center of the pores (Figure 3.2d). Again, if the ions are inserted into a large pore and there is minimal polarization, the charges from the two calculations should be nearly identical.

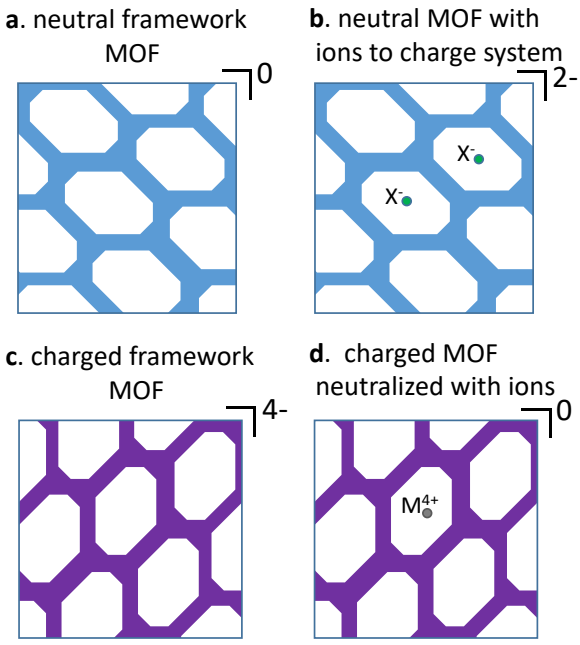

Figure 3.2 *Schematic diagram showing test systems used to evaluate the REPEAT ESP charges. A) A neutral framework MOF. B) The same MOF with two anions added to the center of the pores, giving the unit cell a net charge of -2 e. c) A MOF with a charged framework whose unit cell carries a -4 e charge. d) The same MOF as c) where the framework charge has been neutralized by adding a counter-ion.*

For the first test depicted in Figure 3.2a and 3.2b, we have determined the ESP charges of a neutral large pore framework, IRMOF-16,[50] from a periodic DFT calculation of a 2x1x1 supercell which contains two pores. We then placed F^-, Na^+ and Mg^{2+} ions into the centre of the pores, imparting the unit cell with a net charge of -2, -1,+2 , +4e and derived REPEAT ESP charges for each of these charged systems. The pores of IRMOF-16 are large enough to give a 15 Å separation between the ions and the nearest atoms of the framework as to minimize polarization effects. Table 3.2 gives the mean absolute deviation and maximum deviation between the ESP charges derived from the QM calculation of the neutral framework, and those derived from the same 'charged' framework. The mean absolute deviation (MAD) in the ESP charges is no more than 0.020 e, while the maximum deviation is no more than 0.047 e, showing that overall the charges obtained are very similar in all cases. For the system where 2 Mg^{2+} ions have been added, the differences in charges for all of the atoms in the unit cell are shown graphically in Figure 3.3a. The light blue line in the figure is the ideal

1:1 line. It reveals that the charges are neither systematically higher nor lower than the other charges. This is corroborated by a linear fit which gives a slope of 1.030. Plots for the calculations with other net charges are given in the Supporting Materials.

Table 3.2 Comparison of ESP charges determined from a periodic QM calculation of the neutral MOF IRMOF-16 to those evaluated with charged ions inserted.

Net charge	ions inserted	MAD	Max dev.
-2	2 F^-	0.003	0.007
-1	1 F^-	0.003	0.013
+2	2 Na^+	0.003	0.016
+4	2 Mg^{2+}	0.020	0.047

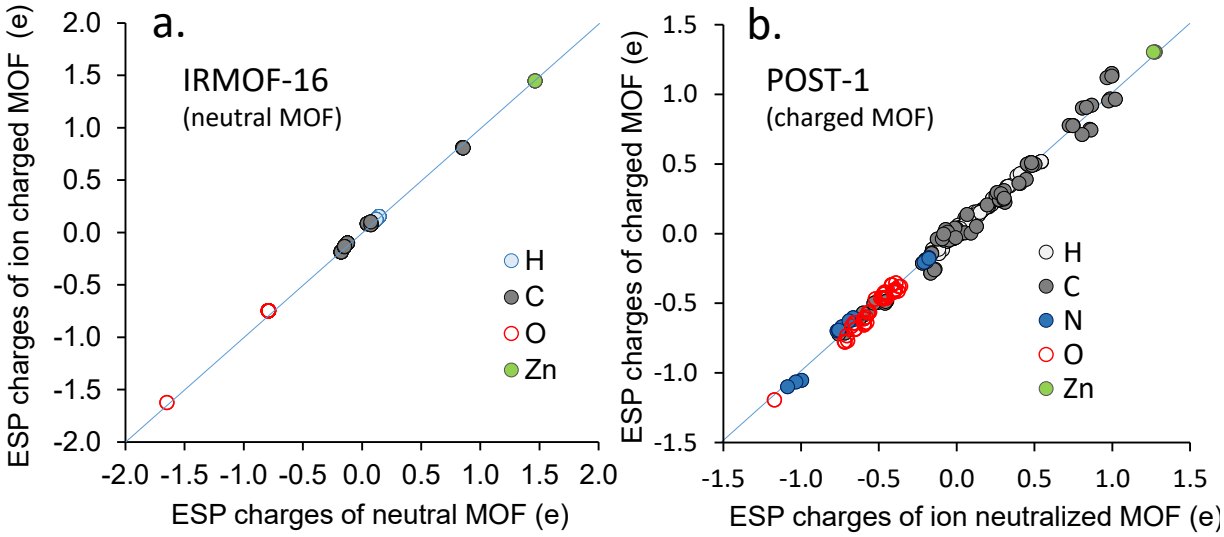

Figure 3.3 *a) ESP charges of IRMOF-16 derived from a periodic QM calculation of the neutral framework compared to those derived from a periodic DFT calculation where 2 Mg^{2+} ions were inserted into the center of the pores to charge the system. b) ESP charges of the charged framework POST-1 derived from periodic QM calculation with a net charge of -4e compared to the charges derived from a calculation with a Ti^{4+} ion inserted to neutralize the framework charge. The light blue line is the line of zero deviation.*

As another numerical test, we have calculated the REPEAT ESP charges of a MOF whose framework possesses a net charge and compared it to the charges derived when counter ions are inserted into the pores to neutralize the system (Figure 3.2c and 3.2d). The MOF POST-1 developed by Seo *et al*,[51] was used for this test since the framework has a net charge (-4 e per unit cell) and since it possesses relatively large, triangular shaped channels with sides of ~13.4 Å in length. To neutralize

the charge in POST-1, a single Ti^{4+} ion was inserted into the center of one of the channels. Figure 3.3b compares the REPEAT ESP charges obtained with and without the counter ion, showing there is excellent agreement. The MAD in the charges was determined to be 0.027 e with a maximum deviation of 0.154 e.

In the two previous examples, large pore MOFs were used as to minimize the polarization of the framework atoms when the ions were added to either charge or neutralize the system. The positioning of the extra-framework ions also ensured that there would be ample ESP fitting points around all atoms of the framework and the ions as to prevent any buried atom artifacts in the ESP fit. To demonstrate that there can be problems with the ESP fitting procedure when neutralizing the framework charge by adding ions, we evaluate the ESP charges in a zeolite of framework type CAS which has small pores of dimensions 2.4 x 4.7 Å.[52] In this case, the pore is small enough that when an ion such as K^+ is added, it nearly fills the pore as shown in Figure 3.4a. In this way, the number of ESP fitting points between the ion and the framework is limited. If the zeolite is synthesized with a 5:1 Si to Al ratio[53] then the unit cell of the CAS has 4 Al atoms that replace Si in the unit cell, giving the framework a net charge of -4 e. We performed QM calculations of the system with a -4 e charge and one where it was neutralized by 4 K^+ ions placed into the center of the pores as shown in Figure 3.4a. REPEAT ESP charges were derived from each and compared. The difference in the derived charges is significant with a MAD of 0.25 e and a maximum deviation of 0.75 e. The difference in all ESP charges is shown graphically in Figure 3.4b. It also worth noting that the ESP fitted charges on the K^+ ions were found to differ by as much as 0.18e from an expected +1. This example demonstrates that it is not always practical to use counter ions to neutralize the charged framework and although this is an extreme example, it suggests that in cases of systems with larger pores there will likely be similar effects albeit to a lesser degree. Thus, we advocate obtaining ESP fitted charges of charged

frameworks by performing periodic QM calculations where the charge is neutralized by a constant background charge rather than actual counter-ions.

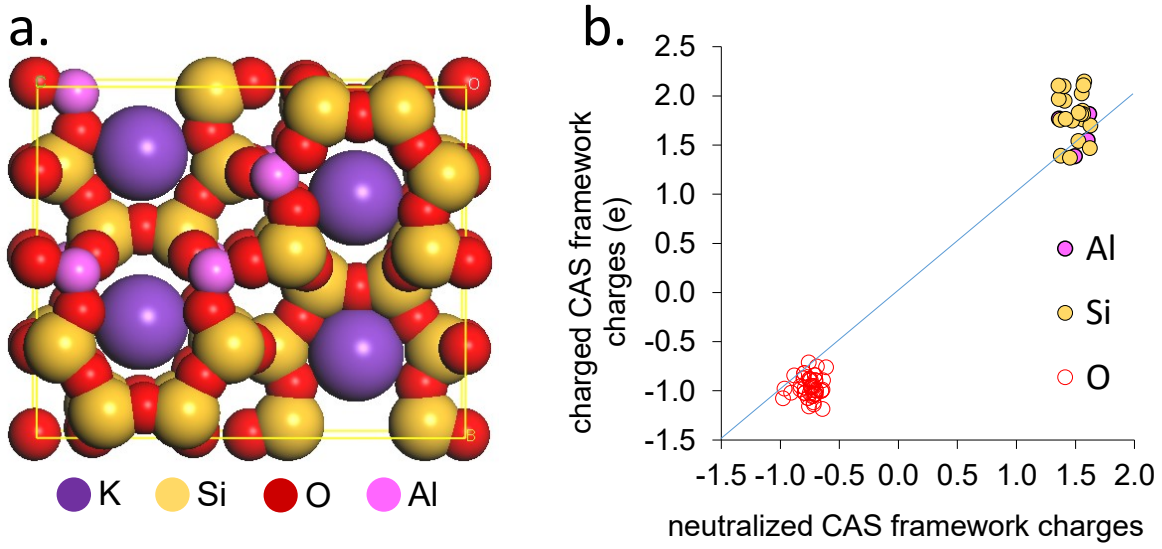

Figure 3.4 *a) A space filling model of the unit cell of the zeolite of framework type CAS with a 5:1 Si to Al ratio, where four K^+ counter ions that neutralize the charge of the framework have been placed in the middle of the pores. b) REPEAT ESP charges of the CAS zeolite shown in a) determined from a QM calculation where the framework as a net charge of -4 to the ESP charges determined from a QM calculation where the framework charge is neutralized with 4 K^+ ions. The light blue line is the line of zero deviation.*

As a final test of the REPEAT ESP charges derived from periodic systems with a net charge, we use the charges to evaluate a physical observable. Dincă *et al.*[3] were able to determine the location of the extra-framework counter ions in a charged framework benzene-tris-tetrazolate MOF, **1**.[54] The single crystal X-ray structure of **1** is shown in Figure 3.5a with the extra-framework Mn^{2+} ions depicted in orange. For the unit cell shown in Figure 3.5a, only 3 Mn^{2+} ions are required to neutralize the charge of the framework, whereas 24 counter-ion locations are identified from the SCXRD experiment (From the view in Figure 3.5a there are 8 sites that are 'on top' of one another and hidden from view). In other words, there are only 3 Mn^{2+} counter-ions on average per unit cell that they are spread across 24 sites within the crystal. Thus, a well sampled Monte Carlo simulation of the Mn^{2+} positions should pick these binding sites given a good representation of the electrostatic potential of the framework.

REPEAT ESP charges were derived from a periodic DFT calculation of the charged MOF, **1**, where the counter ions were removed. Thus, in this periodic DFT calculation the -6 e charge of the framework was neutralized by a constant background charge. Using these charges, we have run a Monte Carlo simulation to determine the most probable positions of the Mn^{2+} extra-framework ions. Figure 3.5b shows an isosurface of the probability distribution of the Mn^{2+} ions determined from the simulation. A comparison of Figure 3.5a with 3.5b reveals that there is excellent agreement in the predicted and experimental extra-framework ion positions. It could be argued that the positions of the ions may not be particularly sensitive to the choice of charges. To investigate this, we have run the same simulation but with atoms-in-molecules, QTAIM, charges. (We note that QTAIM charges are not ideal in this context and higher atomic multipoles should be included in the QTAIM partitioning). An isosurface plot of the Mn^{2+} probability distribution calculated using the QTAIM charges is given in Figure 3.5c, which shows a drastically different counter ion locations compared to those determined with the REPEAT charges and those determined experimentally.

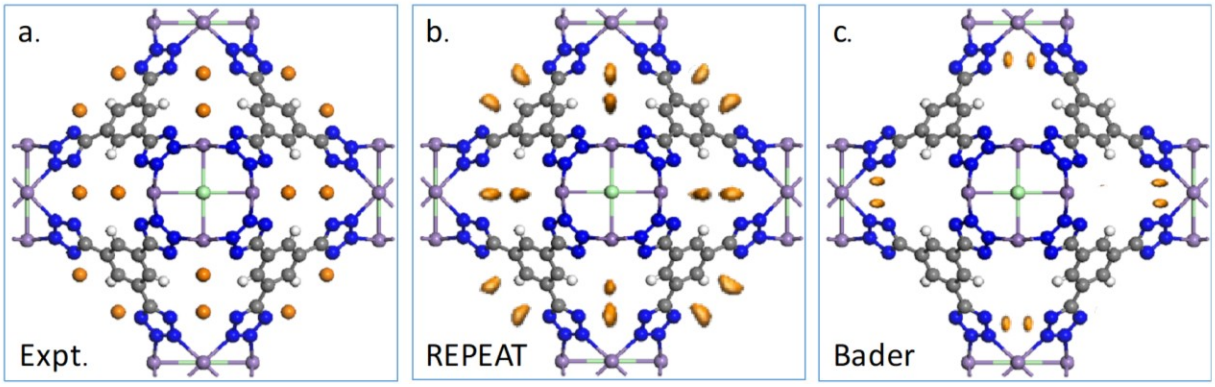

Figure 3.5 *a) Experimental SCXRD structure of the MOF, 1, with the extra-framework Mn^{2+} ions shown in orange. Isosurface (orange) of the extra-framework Mn^{2+} probability distribution determined from a Monte Carlo simulation using b) REPEAT charges and c) QTAIM charges*

We now turn our attention to validating our new split charge model in the atomic basis, SQE_{AB}, which unlike the original SQE method, can be applied to systems with a net charge. It was shown in the Methodology section that the SQE_{AB} model is conditionally equivalent to the original SQE method

for neutral systems. This proof assumes the transformation from split charges to atomic charges is fully reversible, which it is formally not because of the pseudo-inverse. Here we will explore the equivalence numerically by comparing the charges obtained with the two methods for a series of neutral frameworks. For this purpose, we have chosen four neutral zeolite frameworks (RRO, RWR, SFE, UOV) and three neutral MOFs (HKUST-1, IRMOF-16, PCN-69). Although the neutral zeolites are an important class of systems, they only contain one type of bond (O-Si), and may not be considered a rigorous test. As a result, we included a set of MOFs containing between four to six different bond types. For the zeolites we have used the set of SQE parameters calibrated for silicate materials from Verstraelen *et al.*[23], while for the MOFs, we have used our own set of parameters parameterized for neutral MOFs.[55] Table 3.3 summarizes the results and shows that with the exception of one MOF, the charges with the two methods coincide within eight digits ($1x10^{-8}$ e) or more. The exception is that of the MOF PCN-69, which was the largest framework tested with 1008 atoms in the unit cell and 1164 split-charges. In this case the charges differ at most in the sixth digit, but whose mean average deviation is $\sim 10^{-7}$ e. These results suggest that the SQE_{AB} and original SQE models are indeed equivalent for neutral systems.

Table 3.3 Comparison of the partial atomic charges generated with the original SQE model and the SQE_{AB} model for neutral zeolites and neutral MOFs

Framework	Max. deviation (e)	MAD (e)	No. of atoms in unit cell
RRO	1.0e-8	1.1e-9	54
RWR	1.0e-8	1.0e-10	96
SFE	2.0e-9	6.9e-10	42
UOV	1.0e-8	2.1e-10	528
HKUST-1	1.0e-9	1.3e-10	672
IRMOF-16	1.0e-8	5.3e-10	113
PCN-69	4.7e-6	4.2e-7	1008

Because the original SQE model cannot be applied to systems with a net charge, we next compare the SQE_{AB} model to the QEq and the SQE_{LCC} models on charged frameworks. The SQE_{LCC} model is where a local charge contribution is added to the split charges that was introduced in Equations 9-11 in the Methodology section. There it was shown that the SQE_{LCC} model is actually equivalent to the QEq model and therefore does not offer the advantages of the SQE method over the QEq method for charged systems. To validate this numerically we have examined 6 charged zeolites, with varying Si to Al ratios. We have used the QEq-HI parameters optimized for silicate materials based on Hirshfeld-I charges for the atomic parameters, χ_i and κ_i and the SQE-HI parameters for the bond parameters, χ_{ij} and κ_{ij}.[23] The SQE-HI parameters could not be used for both atomic and bond parameters since the atomic electronegativities are not given in the SQE-HI parameter set (It can be shown that the atomic electronegativities can be "absorbed" into a new combined bond electronegativity parameter in the original SQE method.[26]

Table 3.4 Difference in charges between the QEq model and the SQE_{LCC} and SQE_{AB} models for a set of charged zeolites.

base zeolite framework	net charge of unit cell (e)	avg. charge per atom (e)	MAD for Si atoms (e)	
			SQE_{LCC}	SQE_{AB}
ABW	-4	-0.167	2.2E-9	0.046 (0.14)[a]
DDR	-40	-0.111	1.3E-9	0.065 (0.18)
FER	-12	-0.111	2.1E-9	0.068 (0.16)
CDO	-9	-0.083	2.0E-9	0.078 (0.18)
MTW	-7	-0.083	2.2E-9	0.078 (0.19)
LAU	-4	-0.055	2.5E-9	0.089 (0.18)

[a] the parenthetic values correspond to maximum difference.

Table 3.4 compares the charges obtained for the QEq and SQE_{LCC} models using the same parameters for a set of 6 charged zeolites. The third column reports the total net charge divided by the number of atoms in the unit cell giving an indication of how many atoms the total charge must be distributed amongst. For simplicity we only compare the charges on the Si atoms. Table 3.4 reveals

that the QEq and SQE_{LCC} models yield the same partial atomic charges within a given accuracy, since the threshold parameter in the conjugate gradient solver was 10^{-8}. This contrasts the results with the SQE_{AB} model which are also shown in Table 3.4. The MAD in the charges on the Si atoms varies between 0.046 to 0.089 e, while the maximum difference given in parenthesis varies from 0.14 to 0.19 e. The primary result of importance here is that the SQE_{AB} model produces charges that are different than QEq.

Although the new SQE_{AB} model generates different partial atomic charges than the QEq method, one question of interest is how is it an improvement over the QEq model for charged frameworks aside from having more parameters. First we will explore the advantages of the SQE_{AB} model over the QEq model by analyzing the Hessians of two models in the atomic basis. The matrix elements of the Hessian of the QEq model are given by Eq. (3.23),

$$H_{ij}^{\mathrm{QEq}} = \partial^2 E^{\mathrm{QEq}}(\mathbf{Q}) / \partial Q_i \partial Q_j = \begin{cases} \kappa_i & \text{if } i = j \\ J(r_{ij}) & \text{if } i \neq j \end{cases} \qquad (3.23)$$

while a simplified expression for the SQE_{AB} Hessian can be obtained by using Eq. (3.15) for each q_{ij}, is given in Eq. (3.24).

$$H_{ij}^{\mathrm{SQE}} = \partial^2 E^{\mathrm{SQE}}(\mathbf{Q}) / \partial Q_i \partial Q_j = \begin{cases} \kappa_i & \text{if } i = j \\ \kappa_{ij} + J(r_{ij}) & \text{if } i \neq j \end{cases} \qquad (3.24)$$

The SQE_{AB} Hessian, H_{ij}^{SQE}, contains hardness parameters both in the diagonal, i.e. κ_i, and off-diagonal matrix elements, i.e. κ_{ij}. On the other hand, the QEq Hessian, contains adjustable parameters, κ_i, in the diagonal matrix elements only. We note that the original SQE model Hessian has been analyzed before in the split-charge basis.[56] In the split-charge basis the QEq Hessian has atomic hardness parameters in both the diagonal and off-diagonal matrix elements, just as the original

SQE Hessian does. Therefore, it is not immediately transparent how the SQE model should be qualitatively better than QEq. On the other hand, when we compare the SQE_{AB} Hessian given by Eq. (3.24) to the one given by Eq. (3.23), we can immediately see that the incorporation of the adjustable parameters into the off-diagonal matrix elements makes this model more flexible than QEq and the SQE_{AB} model should be able to 'capture' the chemical structure of a given framework better than the QEq method.

To further explore how the SQE_{AB} model can capture the 'chemistry' of a system better than the QEq model for charged frameworks, we have examined the oxygen atom charges generated by the two methods in the LAU framework imbued with a net charge of -4 e. We will divide the oxygen atoms into two categories - those that are covalently bonded to an Al atom, which we have labelled 'Al neighbouring' and those that are not, which we label 'non-neighbouring'. Given in Figure 3.6 are the charges on all the oxygen atoms generated with the SQE_{AB} and QEq methods where the Al-neighbouring atoms are given in red, and the non-neighbouring atoms are shown in green. What is immediately noticeable is that the two oxygen atom types are clearly grouped with the Al-neighbouring atoms being more negatively charged than the non-neighbouring O atoms for the SQE_{AB} model. On the other hand, for the QEq model, the charges on the Al-neighbouring and non-neighbouring atoms are interspersed amongst each other. Moreover, the atom with the least negative charge is an Al-neighbouring atom. Intuitively, one would expect the Al-neighbouring atoms to be more polarized and therefore more negatively charged. To corroborate this notion, we have calculated the REPEAT charges (using a constant background charge) on the same framework. As with the SQE_{AB} model the charges are clearly grouped into neighbouring and non-neighbouring with the Al-neighbouring atoms being more polarized. These results suggest that the SQE_{AB} model is able to capture the chemistry of a system better than the QEq model. We note that the absolute values of the REPEAT charges are

significantly larger than either the SQE_{AB} or QEq charges. However, since the parameters used were somewhat *ad hoc* in nature and not optimized to reproduce ESP charges, this is not necessarily a good indicator of the quality of the models.

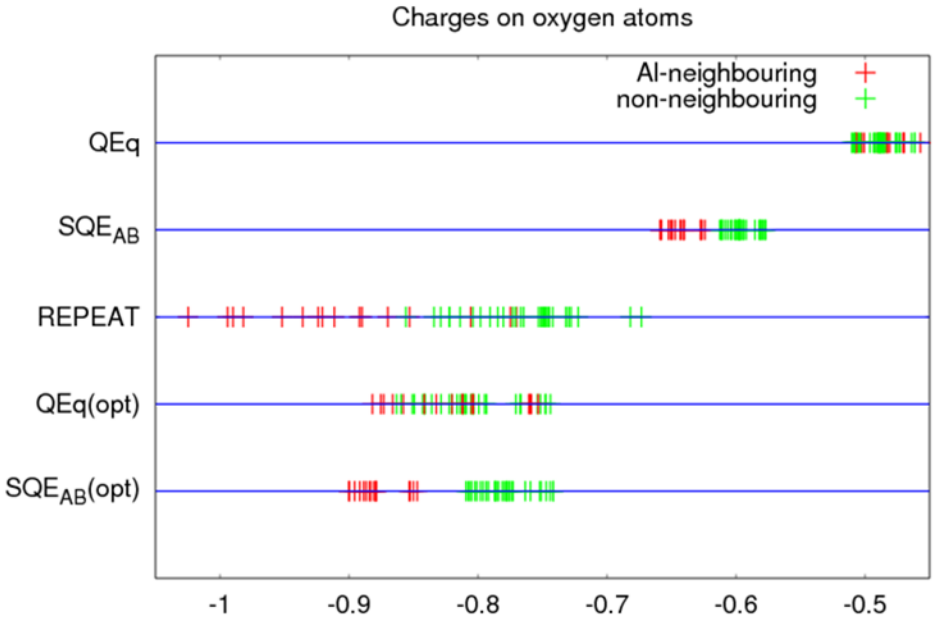

Figure 3.6 The partial atomic charges on oxygen atoms in a charged LAU zeolite framework calculated with various methods. Here, "(opt)" refers to results using a set of parameters optimized to reproduce the REPEAT charges in charged and neutral zeolites.

In the previous comparison, one may argue that the parameter set for the QEq model was not optimal and therefore it did not show the proper polarization of the Al-neighbouring O atoms. Thus, we have performed a parameter optimization of both the QEq and the SQE_{AB} model using a set of 12 zeolites with base topologies ABW, FAU, LTL, EDI, LTJ, MFI, AFG, LAU, RHO, MRE and STO with Si/Al ratios of 1, 3, 5 and ∞ (neutral) (see details in Table 3-S1 of Supporting Information). The parameters were fit to reproduce the REPEAT charges using a custom genetic algorithm, the details of which and the optimized parameters are given in the supporting material. Over the entire training set, the MAD in the charges for the optimized SQE_{AB} model was 16% better (0.101 vs 0.121 e) than with the optimized QEq model. It is worth noting that for zeolites, the QEq model has 6 parameters while the SQE method has only one more free parameter or 7. Figure 3.6 shows partial atomic charges

with the optimized parameters for the LAU framework. We can see that the charges for both the QEq and SQE_{AB} models are both spread over a larger range and shifted to more negative charges that are more in agreement with the REPEAT charges. However, the parameter optimization has not changed the qualitative differences between two models. Namely, that the QEq charges are still interspersed with each other, while the SQE_{AB} charges are noticeably separated in two categories: Al-neighbours and non-neighbours.

3.7 Conclusion

In the first part of this work, the REPEAT method for generating electrostatic potential fitted charges was validated when a constant background charge is used in the periodic electronic structure calculation to neutralize the net charge of the framework. Using simple tests involving the insertion of ions to add or neutralize charge in IRMOF-16 and POST-1, respectively, we demonstrated that the error functional introduced in the REPEAT method naturally accounts for the constant offset in the potential introduced upon adding a constant background charge. This allows for generation of meaningful ESP fitted charges for charged systems without the insertion of counter ions to neutralize the simulation cell, which was shown to be problematic in some cases. Finally, we evaluate the accuracy of REPEAT method for generating meaningful charges for the simulation of charged frameworks, by examining the counter ion locations in a Mn based MOF whose unit cell has a net charge of -6 e. The counter ion positions determined from simulation using the REPEAT charges are in excellent agreement with those determined experimentally. We therefore advocate obtaining REPEAT charges of charged frameworks by performing periodic QM calculations where the charge is neutralized by a constant background charge rather than using counter-ions.

Next we introduced a split-charge equilibration model that we have developed, which we call SQE_{AB}. Whereas the original SQE method can only be applied to neutral systems due to the anti-

symmetry property of the split charges, the SQE_{AB} model can be applied to both charged and neutral systems. Following a similar approach to that used in the QTPIE method,[28] the SQE_{AB} model was constructed by employing a pseudo inverse to express the split-charges in terms of atomic charges, wherein the net charge can be fixed to non-zero values with Lagrange constraints. Compared to the original SQE method, SQE_{AB} model requires an extra matrix diagonalization. However, even for large frameworks with thousands of atoms, the charges can still be evaluated in seconds with modest computing resources.

The SQE_{AB} model was implemented and tested on a variety of neutral and charged zeolites and MOF frameworks. For neutral frameworks, it was shown to reproduce the charges of the original SQE method. An analysis of the Hessian of the SQE_{AB} energy expression compared to that of the QEq method suggests that, the SQE_{AB} model should capture the chemistry of a charged framework better than that of the QEq model. This was explicitly examined for a charged zeolite framework, where it was found that the SQE_{AB} charges on oxygen atoms covalently bonded to an Al atom were more polarized than those that were not – a result that was in agreement with the QM derived REPEAT charges. On the other hand, with the QEq model the charges on the oxygen atoms neighbouring and not-neighbouring aluminum were intermixed. These results were found when using either an *ad hoc* set of parameters or parameters fit to reproduce the REPEAT charges on a set of neutral and charged zeolites.

As charged aluminosilicate zeolites are currently one of the most highly utilized porous materials in industry, and other materials such as MOFs that can have charged frameworks are garnering significant interest, having accurate and fast methods for generating charges for these systems is highly desirable. The validation of the REPEAT method for generating ESP fitted charges of charged periodic frameworks provides a basis for the parameterization the SQE_{AB} method for these

systems. Parameterization of the SQE_{AB} method for charged and neutral zeolites using a large and diverse training set is currently underway in our lab.

3.8 Appendix

Graphical comparison of REPEAT charges determined from periodic QM calculations of the ESP for neutral and charged MOF IRMOF-16, and graphical comparison of QEq and REPEAT charges as well as SQE_{AB} and REPEAT charges for a set of charged zeolites using given optimized parameters. A brief description of the optimization technique is also given.

3.8.1 Appendix A: Equivalency of Two Methods

In order to treat charged systems with the SQE method, one straightforward approach would be to add a local contribution to the partial atomic charge as given in Eq. (3.10), where the local charge corrections, q_i, are treated as independent variables. Unfortunately, the modified SQE scheme, which we have called SQE_{LCC} for Local Charge Correction, reduces into the QEq model for charged systems. In order to demonstrate this non-trivial result, let us derive the working equations with this model. The substitution of Eq. (3.10) in Eq. (3.8) yields the following energy expression:

$$E_{LCC}^{SQE}\left(\tilde{\zeta}_j \quad \sum_i^N \kappa_i \tilde{z}_i \cdot 2\sum_i^N \cdot_i \tilde{z}_i \cdot \sum_i^N \sum_{j>i}^N \left(2 \cdot_{ij} q_{ij} \cdot \kappa_{ij} f_{ij} \cdot 2\sum_i^N \sum_{j\neq i}^N \tilde{} \left(\cdot_{ij}\right) \tilde{z}_i \tilde{z}_j \right. \right. \tag{3.S1}$$

as well as the Lagrange functional

$$L_{LCC}^{SQE}\left(\tilde{\zeta}, z_{net}\right) \quad -_{LCC}\left(\tilde{z}_j\right) \cdot \cdot \left(\left(\sum_i^N \tilde{z}_i \quad z_{net}\right)\right). \tag{3.S2}$$

The subsequent constrained variational minimization of Eq. (3.S2) with respect to $\tilde{\zeta}_i$'s yields

$$\lambda = \kappa_i \tilde{\zeta}_i \cdot \kappa_i \cdot \sum_{j\neq i}^N \tilde{} \left(\cdot_{ij}\right) \tilde{z}_j \cdot \tag{3.S3}$$

Now let us consider the working equations of the QEq model obtained from the variational minimization of Eq. (3.5) with respect to Q_i's:

$$\lambda = \kappa_i Q_i + \chi_i + \sum_{j \neq i}^{N} J\left(r_{ij}\right) Q_j \,. \tag{3.S4}$$

As we can see, Eq. (3.S4) of the QEq model coincides with Eq. (3.S3), in which Q_i variables are replaced with $\tilde{\zeta}_i$'s, thereby showing the two models are equivalent.

3.8.2 Appendix B: Equivalency of SQE_{AB} to SQE for Neutral Systems

We will show that for a neutral system the working equations of the SQE_{AB} model given by Eq. (3.21) are conditionally equivalent to the corresponding equations in the split-charge representation of the original SQE model. We start by rewriting Eq. (3.21) in the following form given in Eq. (3.S5).

$$-\mathbf{H}\mathbf{Q}^0 = \mathbf{g} - \frac{\mathbf{1}^T \mathbf{H}^{-1} \mathbf{g}}{\mathbf{1}^T \mathbf{H}^{-1} \mathbf{1}} \mathbf{1} \,. \tag{3.S5}$$

Further, we shall construct N_{bonds} new equations from Eq. (3.S5) by subtracting the lines with the index *j* from the lines with the index *i* giving:

$$-\sum_{k}^{N} H_{ik} Q_k^0 + \sum_{k}^{N} H_{jk} Q_k^0 = g_i - g_j \,. \tag{3.S6}$$

Here, we use Q_k^0 to indicate that we deal only with the charges of a neutral system. As a result, we have eliminated the contribution from the Lagrange multiplier. Now by making use of the relation between the gradients in the atomic and split-charge representation[15]

$$\frac{\partial}{\partial q_{ij}} = \frac{\partial}{\partial Q_i^0} - \frac{\partial}{\partial Q_j^0} \,, \tag{3.S7}$$

we transform the r.h.s. of Eq. (3.S6) into the split-charge representation

$$-\sum_{k}^{N}\frac{\partial^2 E_{\mathrm{AB}}^{\mathrm{SQE}}}{\partial Q_i^0 \partial Q_k^0}Q_k^0+\sum_{k}^{N}\frac{\partial^2 E_{\mathrm{AB}}^{\mathrm{SQE}}}{\partial Q_j^0 \partial Q_k^0}Q_k^0=\frac{\partial E_{\mathrm{AB}}^{\mathrm{SQE}}}{\partial q_{ij}}. \tag{3.S8}$$

In order to transform the l.h.s. of Eq. (3.S8) into the split-charge representation, we have to remember that the split-charges are constrained by the antisymmetry relations of Eq. (3.9). Therefore, Eq. (3.S8) is equivalent to Eq. (3.S9).

$$-\sum_{k}^{N}\sum_{l>k}^{N}\frac{\partial^2 E_{\mathrm{AB}}^{\mathrm{SQE}}}{\partial Q_i^0 \partial Q_k^0}q_{kl}+\sum_{k}^{N}\sum_{l>k}^{N}\frac{\partial^2 E_{\mathrm{AB}}^{\mathrm{SQE}}}{\partial Q_i^0 \partial Q_k^0}q_{lk}+\sum_{k}^{N}\sum_{l>k}^{N}\frac{\partial^2 E_{\mathrm{AB}}^{\mathrm{SQE}}}{\partial Q_j^0 \partial Q_k^0}q_{kl}-\sum_{k}^{N}\sum_{l>k}^{N}\frac{\partial^2 E_{\mathrm{AB}}^{\mathrm{SQE}}}{\partial Q_j^0 \partial Q_k^0}q_{lk}=\frac{\partial E_{\mathrm{AB}}^{\mathrm{SQE}}}{\partial q_{ij}} \tag{3.S9}$$

Now by exchanging the indices *k* and *l* in the second and fourth terms of the l.h.s. of Eq. (3.S9) and making use of the relation between the gradients twice, i.e.

$$\frac{\partial^2}{\partial q_{ij}\partial q_{kl}}=\frac{\partial^2}{\partial Q_i^0 \partial Q_k^0}-\frac{\partial^2}{\partial Q_i^0 \partial Q_l^0}+\frac{\partial^2}{\partial Q_j^0 \partial Q_k^0}-\frac{\partial^2}{\partial Q_j^0 \partial Q_l^0}, \tag{3.S10}$$

we eventually obtain

$$-\sum_{k}^{N}\sum_{l>k}^{N}\frac{\partial^2 E_{\mathrm{AB}}^{\mathrm{SQE}}}{\partial q_{ij}\partial q_{kl}}q_{kl}=\frac{\partial E_{\mathrm{AB}}^{\mathrm{SQE}}}{\partial q_{ij}}. \tag{3.S11}$$

For the invertible matrix **T** (square and with the full rank) given by Eq. (3.12), $E_{\mathrm{AB}}^{\mathrm{SQE}}$ equals E^{SQE} since relation (3.13) would be simply a unitary transformation between two basis sets. In this case Eq. (3.S11) coincides with the working equations of the original SQE model obtained from the variational minimization of Eq. (3.8) with respect to the split-charges.

3.8.3 Appendix C: Supporting Information

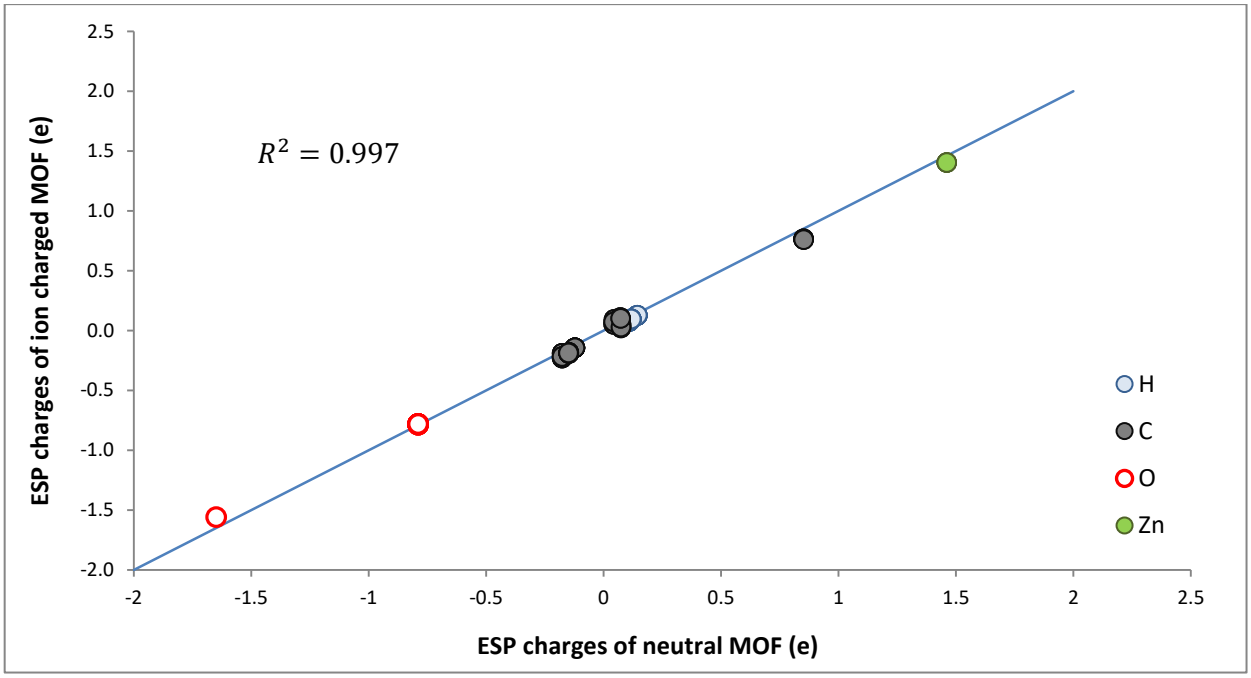

Figure 3-S1: *Comparison of ESP charges determined from periodic QM calculations of the neutral MOF IRMOF-16 to those evaluated with 2 F⁻ ions inserted into the pores.*

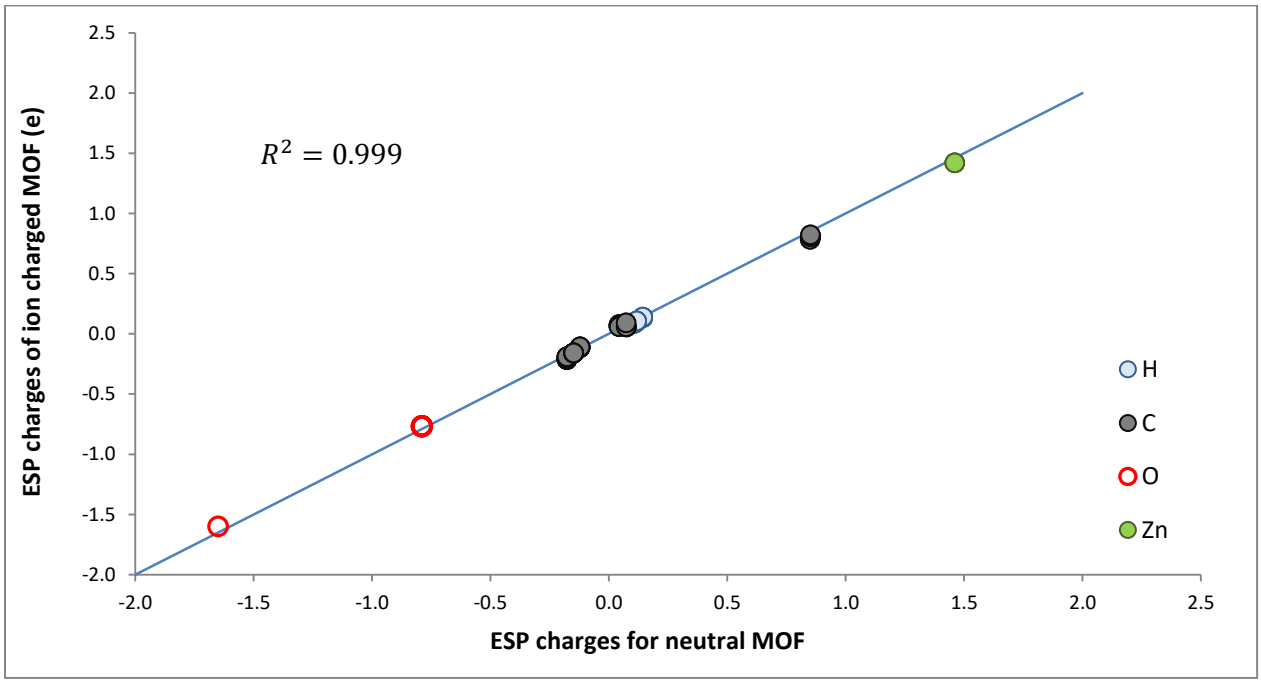

Figure 3-S2: *Comparison of ESP charges determined from periodic QM calculation of the neutral MOF IRMOF-16 to those evaluated with a F⁻ ions inserted into the pores*

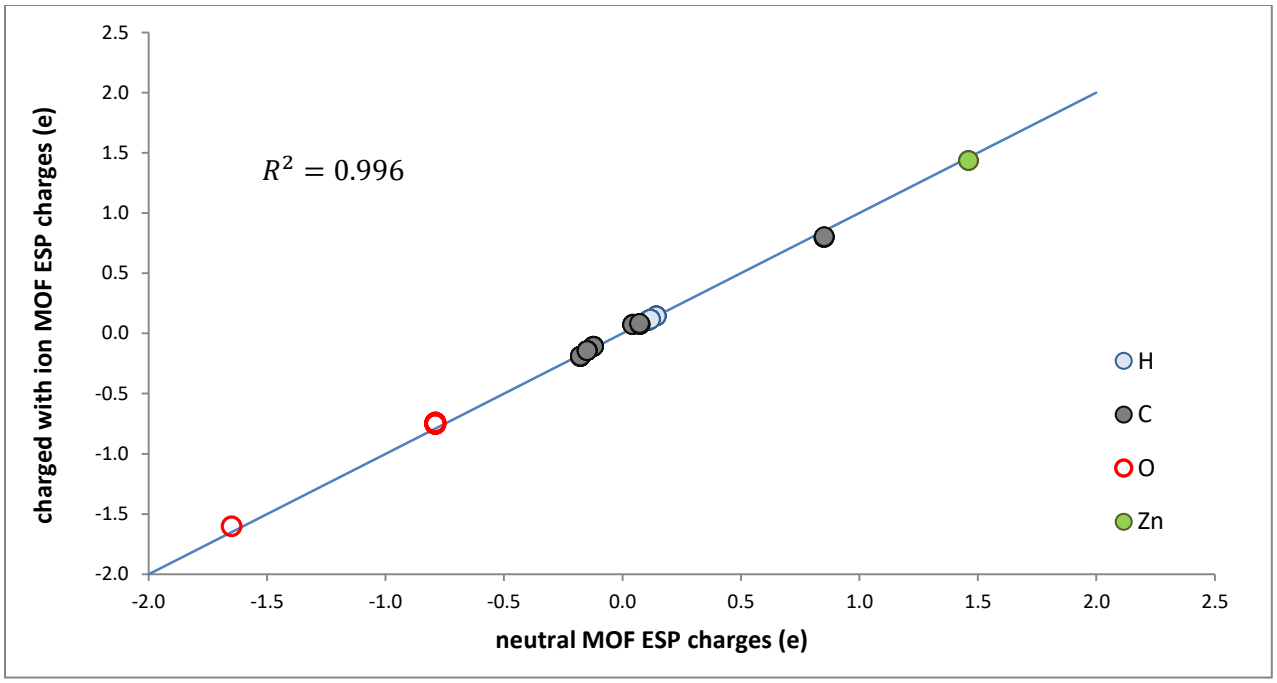

Figure 3-S3: *Comparison of ESP charges determined from periodic QM calculation of the neutral MOF IRMOF-16 to those evaluated with 2 Na^+ ions inserted into the pores*

Table 3-S1: Difference in charges between the QEq model and REPEAT method as well as SQE_{AB} model and REPEAT method for a set of charged zeolites based on the optimized EEM parameters given in Table 3-S2

Base zeolite framework	Net charge of unit cell (e)	Avg. charge per atom (e)	MAD (e)	
			QEq	SQE_{AB}
ABW	-4	-0.167	0.206	0.151
FAU	-70	-0.121	0.149	0.123
LTL	-12	-0.111	0.112	0.107
EDI	-10	-0.083	0.132	0.114
LTJ	-4	-0.083	0.139	0.113
MFI	-24	-0.083	0.153	0.129
AFG	-8	-0.055	0.114	0.105
LAU	-4	-0.055	0.091	0.073
RHO	-8	-0.055	0.094	0.082
MFI	0	0.0	0.076	0.067
MRE	0	0.0	0.099	0.083
STO	0	0.0	0.082	0.070

Table 3-S2: Optimized EEM parameters in eV for the QEq and SQE_{AB} models based on a set of charged zeolite frameworks listed in Table 3-S1.

Atom or bond	QEq	SQE_{AB}
κ(O)	9.9037	3.1632
κ(Al)	15.365	15.699
κ(Si)	16.163	15.238
χ(O)	6.4282	0.24229
χ(Al)	-15.582	-0.14777
χ(Si)	-11.542	0.96987
κ(O-Al)		10.1644
κ(O-Si)		10.5856
χ(O-Al)		-12.9003

3.8.4 Appendix D: Brief overview of the Genetic Algorithm used in this work

A custom genetic algorithm (GA) was used to fit all of the parameters simultaneously for each method separately (QEq and SQE_{AB}). The GA is initialized with the creation of multiple sets of randomly generated parameters, collectively known as a generation, that were then evaluated for how closely they reproduced the QM ESP. The new generation was formed by using a roulette wheel selection algorithm, which chooses two individuals from the generation to act as parents to new individuals by a mating algorithm. The mating algorithm selects a random value for each parameter which is between the values of each corresponding parameter for both the parents'. Subsequent mutations were allowed that would alter a given parameter by ± 30% of the parameter value. The GA was considered converged when the top performer in subsequent generations remained the same for ten generations.

3.9 References

(1) Furukawa, H.; Cordova, K. E.; O'Keeffe, M.; Yaghi, O. M. The Chemistry and Applications of Metal-Organic Frameworks. *Science.*. **2013**, *341*, 974.

(2) Degnan, T. F. Recent Progress in the Development of Zeolitic Catalysts for the Petroleum Refining and Petrochemical Manufacturing Industries. *Elsevier B.V*. **2007**, *170,* 54-65.

(3) Dincă, M.; Dailly, A.; Liu, Y.; Brown, C. M.; Neumann, D. A; Long, J. R. Hydrogen Storage in a Microporous Metal-Organic Framework with Exposed Mn2+ Coordination Sites. *J. Am. Chem. Soc.* **2006**, *128*, 16876-16883.

(4) Hurd, J. A.; Vaidhyanathan, R.; Thangadurai, V.; Ratcliffe, C. I.; Moudrakovski, I. L.; Shimizu, G. K. H. Anhydrous Proton Conduction at 150 °C in a Crystalline Metal-Organic Framework. *Nat. Chem.* **2009**, *1*, 705-710.

(5) Ye, Y.; Zhang, L.; Peng, Q.; Wang, G.-E.; Shen, Y.; Li, Z.; Wang, L.; Ma, X.; Chen, Q.-H.; Zhang, Z.; Xiang, S. High Anhydrous Proton Conductivity of Imidazole-Loaded Mesoporous Polyimides over a Wide Range from Subzero to Moderate Temperature. *J. Am. Chem. Soc.* **2014**, *3*, 0-5.

(6) Xiao, C.; Wang, Y.; Chen, L.; Yin, X.; Shu, J.; Sheng, D.; Chai, Z.; Albrecht-Schmitt, T. E.; Wang, S. Bosting Proton Conductivity in Highly Robust 3D Inorganic Cationic Extended Frameworks Through Ion Exchange with Dihydrogen Phosphate Anions.*Chemistry* **2015**, *21*, 17591-17955.

(7) Duren, T.; Bae, Y. S.; Snurr, R. Q. Using Molecular Simulations to Characterise Metal-Organic Frameworks for Adsorption Applications *Chem. Soc. Rev.* **2009**, *38*, 1237-1247.

(8) Simon, C. M.; Kim, J.; Gomez-Gualdron, D. A.; Camp, J. S.; Chung, Y. G.; Martin, R. L.; Mercado, R.; Deem, M. W.; Gunter, D.; Haranczyk, M.; Sholl, D. S.; Snurr, R. Q.; Smit, B. The Materials Genome in Action: Identifying the Performance Limits for Methane Stoarage. *Energy Environ. Sci.* **2015**, *8*, 1190-1199.

(9) Campañá, C.; Mussard, B.; Woo, T. K. Electrostatic Potential Derived Atomic Charges for Periodic Systems Using a Modified Error Functional. *J. Chem. Theory Comput.* **2009**, *5*, 2866-2878.

(10) Chen, D. L.; Stern, A. C.; Space, B.; Johnson, J. K. Atomic Charges Derived from Electrostatic Potentials for Molecular and Periodic Systems*J. Phys. Chem. A* **2010**, *114*, 10225-10233.

(11) Gabrieli, A.; Sant, M.; Demontis, P.; Suffritti, G. B. Partial Charges in Periodic Systems: Improving Electrostatic Potential (ESP) Fitting via Total Dipole Fluctuations and Multiframe Approaches. *J. Chem. Theory Comput.* **2015**, *11*, 3829-3843.

(12) Fernandez, M.; Woo, T. K.; Wilmer, C. E.; Snurr, R. Q. Large-scale Quantitative Structure-Property Relationship (QSPR) Analysis of Methane Storage in Metal-Organic Frameworks. *J. Phys. Chem. C* **2013**, *117*, 7681-7689.

(13) Rappe, A. K.; Goddard, W. A. Charge Equilibration for Molecular Dynamics Simulations. *J. Phys. Chem.* **1991**, *95*, 3358-3363.

(14) Kadantsev, E. S.; Boyd, P. G.; Daff, T. D.; Woo, T. K. Fast and Accurate Electrostatic in Metal Organic Frameworks with a Robust Charge Equilibration Parametrization for High-Throughput Virtual Screening of Gas Adsorption. *J. Phys. Chem. Lett.* **2013**, *4*, 3056-3061.

(15) Nistor, R. A.; Polihronov, J. G.; Müser, M. H.; Mosey, N. J. A Generalization of the Charge Equilibration Method for Nonmetallic Materials *J. Chem. Phys.* **2006**, *125*, 94108-94118

(16) Verstraelen, T.; Van Speybroeck, V.; Waroquier, M. The Electronegativity Equalization Method and the Split Charge Equilibration Applied to Organic Systems: Parameterization, Validation, and Comparison. *J. Chem. Phys.* **2009**, *131*, 44127-44146.

(17) Verstraelen, T.; Pauwels, E.; De Proft, F.; Van Speybroeck, V.; Geerlings, P.; Waroquier, M. Assessment of Atomic Charge Models for Gas-Phase Computations of Polypeptides. *J. Chem. Theory Comput.* **2012**, *8*, 661-676.

(18) Smirnov, K. S. Assesment of Split-Charge Equilibration Model for Development of Polarizable Force Fields. *Model. Simul. Mater. Sci. Eng.* **2015**, *23*, 74006-74023.

(19) Dapp, W. B.; Müser, M. H. Redoc Reactions with Empirical Potentials: Atomistic Battery Discharge Simulations. *J. Chem. Phys.* **2013**, *139*, 64106-64129.

(20) Dapp, W. B.; Müser, M. H. Towards Time-Dependent, Non-Equilibrium Charge-Transfer Force Fields. *Eur. Phys. J. B* **2013**, *86*, 337-348.

(21) Müser, M. H. The Chemical Hardness of Molecules and the Band Gap of Solids Within Charge Equilibration Formalisms. *Eur. Phys. J. B* **2012**, *85*, 135-149.

(22) Nistor, R. A.; Müser, M. H. Dielectric Properties of Solids in the Regular and Split-Charge Equilibration Formalisms. *Phys. Rev. B* **2009**, *79*, 104303-104311.

(23) Verstraelen, T.; Sukhomlinov, S. V.; Van Speybroeck, V.; Waroquier, M.; Smirnov, K. Computation of Charge Distribution and Electrostatic Potential in Silicates with the Use of Chemical Potential Equalization Models. *J. Phys. Chem. C* **2012**, *116*, 490-504.

(24) Fischer, M.; Bell, R. G. Modelling CO_2 Adsorption in Zeolites Using DFT-Derived Charges: Comparing System-Specific and Generic Models. *J. Phys. Chem. C.* **2013**, *117*, 24446-24454.

(25) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: the RESP Model *J. Phys. Chem.* **1993**, *97*, 10269-10280.

(26) Verstraelen, T.; Bultinck, P.; Van Speybroeck, V.; Ayers, P. W.; Van Neck, D.; Waroquier, M. The Significance of Parameters in Charge Equilibration Models. *J. Chem. Theory Comput.* **2011**, *7*, 1750-1764.

(27) Verstraelen, T.; Pauwels, E.; De Proft, F.; Van Speybroeck, V.; Geerlings, P.; Waroquier, M. Assessment of Atomic Charge Models for Gas-Phase Computations of Polypeptides. *J.*

Chem. Theory Comput. **2012**, *8*, 661-676.

(28) Chen, J., Hundertmark, D. & Martínez, T. A. A Unified Theoretical Framework for Fluctuating-Charge Models in Atom-Space and in Bond-Space *J.Chem.Phys* **2008**, *129*, 214113-214124.

(29) Verstraelen, T., Ayers, P. W., Van Speybroeck, V. & Waroquier, M. ACKS2: Atom-Condensed Kohn-Sham DFT Approximated to Second Order. *J.Chem.Phys* **2013**, *138*, 74108-74128.

(30) Makov, G.; Payne, M. Periodic Boundary Condition in Ab Initio Calculations. *Phys. Rev. B.* **1995**, *51*, 4014-4022.

(31) Bruneval, F.; Crocombette, J. P.; Gonze, X.; Dorado, B.; Torrent, M.; Jollet, F. Consistent Treatment of Charged Systems Within Periodic Boundary Conditions: The Projector Augmented-Wave and Pseudopotential Methods Revisited. *Phys. Rev. B - Condens. Matter Mater. Phys.* **2014**, *89*, 1-13.

(32) Leslie, M.; Gillan, N. J. The Energy and Elastic Tensor of Defects in Ionic Crystals Calculated by the Supercell Method. *J. Phys. C Solid State Phys.* **2000**, *18*, 973-982.

(33) Ewald, P. Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Ann. Phys.* **1921**, *369*, 253-287.

(34) Penrose, R. A Ganeralized Inverse for Matrices. *Proc. Cambridge Phil. Soc.* **1955**, *51*, 406-413.

(35) Merris, R. Laplacian Matrices of Graphs: A Survey. *Linear Algebra Appl.* **1994**, *197–198*, 143-176.

(36) Fiedler, M. Algebraic Connectivity of Graphs. *Czechoslov. Math. J.* **1973**, *23*, 298-305.

(37) Estrada, E. Graph Theoretical Invariant of Randić Revisited. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 1022-1025.

(38) Merris, R. Doubly Stochastic Graph Matrices II. *Linear Multilinear Algebr.* **1998**, *45*, 275-285.

(39) Todeschini, R.; Consonni, V. Mannhold, R., Kubinyi, H., Folkers, G., In Methods and Principles in Medicinal Chemistry; *Eds.; John Wiley & Sons: Weinheim*, **2009**, 41, 1150-1165.

(40) Chen, J.; Martinez, T. J. *J. Chem. Phys.* **2009**, *131*, 44114.

(41) Hub, J. S.; de Groot, B. L.; Grubmüller, H.; Groenhof, G. *J. Chem. Theory Comput.* **2014**, *10* (1), 381.

(42) Chen, J.; Martínez, T. J. QTPIE: Charge Transfer with Polarization Current Equalization. A Fluctuating-Charge Model with Correct Asymptotics. *Chem. Phys. Lett.* **2007**, *438*, 315-320.

(43) Morales, J.; Martínez, T. J. A New Approach ot Reactive Potentials with Fluctuating Charges: Quandratic Valence-Bond Model. *J. Phys. Chem. A* **2004**, *108*, 3076-3084.

(44) Kresse, G.; Furthmüller, J. Efficient Iterative Schemes for Ab Initio Total-Energy

Calculations Using a Plane-Wave Basis Set. *Phys. Rev. B* **1996**, *54*, 11169-11186.

(45) Blöchl, P. E. Projector Augmented-Wave Method. *Phys. Rev. B* **1994**, *50* (24), 953-979.

(46) Perdew, J.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple *Phys. Rev. Lett.* **1996**, *77* (18), 3865-3868.

(47) Forrester, T. R.; Smith, W. The DL-POLY 2 User Manual. *CCLRC Daresbury Laboratory: Daresbury, UK* **1995**, 1-217.

(48) Vaidhyanathan, R.; Iremonger, S. S.; Shimizu, G. K. H.; Boyd, P.; Alavi, S.; Woo, T. K. Direct observation and quantification of CO_2 binding within an amine-functionalized nanoporous solid; *Science,*.**2010**, *330*, 650-653.

(49) Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. UFF, a Rule-Based Full Periodic Table Force Field for Molecular Mechanics and Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **1992**, *114*, 10024-10035.

(50) Yaghi, O. M.; O'Keeffe, M.; Ockwig, N. W.; Chae, H. K.; Eddaoudi, M.; Kim, J. Reticular Synthesis and the Design of New Materials. *Nature* **2003**, *423*, 705-714.

(51) Seo, J.; Whang, D.; Lee, H.; Jun, S.; Oh, J.; Jeon, Y.; Kim, K. A Homochiral Metal-Organic Porous Material for Enantioselective Separation and Catalysis. *Nature* **2000**, *404*, 982-986.

(52) Baerlocher, C.; McCusker, L. B.; Olson, D. H. Atlas of Zeolite Framework Types. *Elsevier Science: Amsterdam,* **2007**, *6,* 1-404.

(53) Gatta, G. D.; Rotiroti, N.; Zanazzi, P. F.; Rieder, M.; Drabek, M.; Weiss, Z.; Klaska, R. Synthesis and Crystal Structure of the Feldspathoid $CsAlSiO_4$ Open-Framework Silicate and Potential Nuclear Waste Disposal Phase. *Am. Mineral.* **2008**, *93*, 988-995.

(54) Dinca, M.; Long, J. R. Hydrogen Stoarage in Microporous Metal-Organic Frameworks with Exposed Metal Sites. *Angew. Chemie-International Ed.* **2008**, *47*, 6766-6779.

(55) Collins, S. P.; Woo, T. K. Split-Charge Equilibration Parameters for Generating Rapid Partial Atomic Charges in Metal-Organic Frameworks and Pourous Polymer Networks for High-Throughput Screening. *J. Phys. Chem. C* **2017**, *121*, 903-910.

(56) Lee Warren, G.; Davis, J. E.; Patel, S. Origin and Control of Superlinear Polarizability Scalin in Chemical Potential Equalization Methods. *J. Chem. Phys.* **2008**, *128*, 144110-144124.

4 Force Field Optimization for Gaseous Separations with Aluminosilicate Zeolites

4.1 Summary

The focus of this chapter is the optimization and validation of a force field for modeling gas adsorption in aluminosilicate zeolites. Using REPEAT charges along with the SQE_{AB} method introduced in Chapter 3, significant improvement over current strategies for modeling electrostatic interactions in aluminosilicates is observed. The atom- and bond-specific SQE_{AB} parameters were optimized to reproduce the gauge-corrected quantum mechanical (QM) electrostatic potential (ESP), using a training set comprising 63 unique framework types. This was done using a genetic algorithm written by a Ph.D. student of the Woo group, Sean Collins, for optimizing SQE parameters to reproduce the QM ESP. Parameters for the QEq model were also obtained by the same method, but the SQE_{AB} parameters will be the focus of the discussion in this chapter. The parameters were then validated on a set of 48 unique framework types that were not part of the training set. In this work, the dispersive interactions as modeled by Lennard-Jones potential were also parameterized to reproduce experimental uptake data for CO_2, CH_4, N_2. The parameters were fit simultaneously to 39 experimental isotherms (13 for each gas) taken from independent sources. Through the development of a diverse training set that accounts for the structural diversity found amongst aluminosilicates, significant improvements over the force fields currently used to model these gaseous processes are achieved.

4.2 Introduction

Zeolites are three-dimensional microporous crystalline solids with well-defined structures, and are composed primarily of Si, O and Al. The Si and Al atoms are tetrahedrally coordinated with each other through shared oxygen atoms, to form what is sometimes referred to as a '4-connected

net.' In addition to these siliceous and aluminosilicate zeolites, there is also the possibility for aluminophosphate materials in which PO_4 tetrahedra are incorporated into the frameworks, as well as the possible incorporation of trace amounts of transition metals such as Mn or Co. Many of the most commonly used zeolites in industry possess frameworks that have a net charge, which is balanced by counter-ions that intercalate through the pores. The charge on the framework results from the substitution of some of the Si atoms by Al atoms. Because Al is in it +3 oxidation state, the charge of the surrounding oxygen atoms is no longer balanced when it replaces Si in the framework. Thus, for each Al atom incorporated into the framework there is a net negative charge that is added to the framework.

In practice, virtually all zeolites that are utilized in industry have a proportion of the Si sites replaced with Al and, in other words, there are no purely siliceous zeolites with neutral frameworks that are in wide use. Whether it is for the oil industry, where zeolites are used for catalytic cracking, or for the natural gas industry, where zeolites are used to purify methane gas, zeolites are an indispensable part of our world. For example, aluminosilicate zeolites NaX and NaY (FAU topology with 77 – 96 Al per u.c and less than 77 Al per u.c, respectively) are most commonly used for natural gas purification, where methane gas is separated from CO_2. Assuming each Al atom adds a -1e charge to the unit cell, this would allow for these zeolites to have a net framework charge ranging from -77e to -96e. The addition of aluminum and corresponding counter-ions effects the adsorption properties of zeolites in two important ways: (1) it creates stronger binding sites as a result of the additional interactions with the charge balancing cations, and (2) it decreases the saturation uptake of because of the reduction in free-volume, again due to the counter-ions. This modification to the adsorption properties is reflected as an increase in the uptake at low pressures and a decrease in the uptake at high pressures. In this regard, there have been many experiments conducted on

aluminosilicate zeolites that have demonstrated a significant enhancement in the adsorption properties relative to purely siliceous zeolites.[1–4]

In recent decades, computational modeling has proven to be a powerful tool for identifying structure-property relationships, allowing for the rational design of porous materials for different gaseous separations.[5–8] When performing simulations of chemical systems there are a variety of methods at one's disposal. First principles-methods, based on quantum mechanics (QM), are generally considered high-level methods that give a complete description of the chemical system in question. However, high-accuracy QM calculations come at the expense of efficiency, since the electronic structure is treated explicitly. Due to the elevated computational cost of QM methods, typically only systems containing hundreds of atoms can be evaluated. For much larger systems, for example biological systems such as proteins, accurate representation of the systems chemistry requires the treatment hundreds to thousands of atoms. In a similar manner, screening studies involve the simulation of thousands to millions of materials, each with unit cells containing hundreds of atoms. In both cases, the use of QM methods would be intractable, and classical atomistic simulations are routinely used for such large problems. These methods have greatly improved efficiency relative to QM methods because they by-pass the Schrödinger equation for representing the potential energy surface of the system. Instead, the potential energy surface is evaluated through parameterized atom-pair potentials, or so-called force fields, which greatly reduces the number and complexity of the calculations that must be performed. Although force field based methods are faster than QM methods, the potentials are highly parameterized with unique parameters for describing the different interactions present, in a given system. A great deal of work is put into continually developing more and more accurate potentials for describing a variety of interacting atomic systems.

When studying the non-bonded guest-host interaction in zeolites, classical molecular dynamics (MD) and Monte Carlo (MC) simulations have been used extensively. A recent example comes from Smit *et al.*[9], who employed MC simulations to screen the methane storage capability of the International Zeolite Associate (IZA) structure data base of realized siliceous zeolites, in addition to a hypothetical zeolite database of over 100,000 siliceous frameworks. Studies such of these use parameterized force fields that allow for a large number of calculations to be completed in a practical amount of time. To model the non-bonded interactions when simulating gas adsorption in porous materials, the interaction energy is broken into two components: (1) the electrostatic interactions, and (2) the van der Waals (VdW) interactions. The electrostatic interactions are commonly accounted for by the fixed charge approximation, wherein each atom is assigned a charge that, collectively, represents the electrostatic potential (ESP) of the system. Due to the strength and range of electrostatic interactions, the assignment of the atomic charges is critical for obtaining accurate results when modeling guest-host behaviors.

There is a variety of methods used to assign partial atomic charges, among which so-called ESP-fitted methods have long been used in molecular simulation. Such methods involve fitting partial atomic charges to reproduce the electrostatic potential derived from first principles QM calculations. Although, ESP-fitted charges have been used for decades in the simulation of molecules, it was not until relatively recently that a method was introduced for extracting ESP-fitted charges directly from periodic QM calculations. When dealing with periodic systems the ESP is ill-defined and periodic QM methods can have different arbitrary offsets in the potential. The method introduced by Campagñá *et al.*[10], known as REPEAT, circumvents this issue by treating the relative differences in the ESP rather than the absolute values. Since its introduction, REPEAT has had much success in modeling electrostatic interactions in porous materials. [11–16]

Table 4.1 *Different generic charges used for modelling zeolites.*

Author (year)	Zeolite	Partial Atomic Charge (e)				
		Si	Al	O_{Si}	O_{Al}	Na^+
Losinger *et al.* (1991)	Generic	+1.93	+1.02	-0.49	-0.49	+1.00
Hirotani *et al.* (1997)	ZSM-5 Si/Al = 94	+0.783	+0.587	-0.40	-0.40	+100
	ZSM-5 Si/Al = 46	+0.766	+0.575	-0.40	-0.40	+1.00
Faux *et al.* (1997)		+3.70	+2.775	-1.869	-1.869	+1.00
Jaramillo *et al.* (1999)		+2.05	+1.75	-1.025	-1.20	+1.00
Calero *et al.* (2009)	Generic	+0.786	+0.486	-0.393	-0.414	+0.3834
Fang *et al.* (2013)	LTA-4A Si/Al = 1	+2.21	+2.08	-1.105	-1.32	+0.99
Fischer *et al.* (2013)	Generic	+1.034	--	-0.502	--	--
Sun *et al.* (2017)	Clusters	QEq Charges (UFF Parameters)				

Generally, the electrostatic interactions in zeolites are modeled using generic charges sets, where every atom of the same type within the framework is assigned the same charge. Examples of such charges, taken from different zeolite force fields, have been listed in Table 4.1.[17–23] What is most interesting to note is the variation in the magnitude of the charges across the different force fields, reflecting the various methodologies applied in selecting the charges. Although this simple model has the advantage of straight-forward and fast, a recently published study by Fischer *et al.*[22] identifies an important drawback associated with using generic charge sets. In that study, the authors optimized and evaluated their own generic charges against system specific QM derived REPEAT charges for a series of siliceous zeolites. To compare the two charge methods, the authors analyzed the simulated CO_2 uptakes in a series of 24 siliceous zeolites obtained applying either method. The authors found that in half the cases the difference between the REPEAT uptakes and those obtained using generic charges on the framework atoms were non-negligible, with root-mean-squared percentage errors (RMSPEs) exceeding 10%. Further comparisons revealed that, in the

majority of cases evaluated, generic charges generally failed to reproduce all features of the ESP as well as the system specific REPEAT charges.

It is also important to note that all of the zeolite force fields presented in Table 4.1 have modeled the electrostatics using charge neutral unit cells. To do so in charged periodic systems such as aluminosilicates, counter-ions must be placed within the unit cell to balance the framework charge, prior to simulation. The difficulty with this is that the counter-ion positions within charged zeolite frameworks are often not known experimentally. Consequently, the number of framework topologies that are sampled when optimizing such charges for zeolites is limited. Furthermore, placing counter-ions in the pores in order to neutralize the overall charge can result in the framework charge distribution being polarized towards the counter-ions. Because counter-ion mobility is critical for reproducing experimental uptake in pourous materials during simulations, such polarization in the framework charges can result in a bias in the counter-ions movement towards the pre-determined locations. To overcome this, Krykunov *et al.*[24] (Chapter 3 of this thesis) have recently validated the REPEAT model for charged periodic systems, including MOFs and zeolites. Using the REPEAT method, it was shown that one can derive atomic charges directly from periodic QM calculations of the ESP for charged systems, without the need to place counter-ions within the unit cell. However, a serious drawback of the REPEAT method is that it requires expensive electronic structures calculations for each framework being considered. This becomes a significant issue when performing screening studies involving thousands or millions of materials.[9,25] In such cases, other more computationally efficient approaches are required. Such an alternative is offered by electronegativity equalization methods, which allow for system specific charges to be generated in seconds rather than hours. These methods include the charge equilibration (QEq) model and the more general split-charge equilibration (SQE) model. Of the two, QEq is the simplest

and most popular formulation, introduced by Rappé and Goddard,[26]. The SQE method developed by Müser and co-workers.[27] is simply a generalization of QEq, which uses the same atomic parameters but introduces additional bond hardness and bond electronegativity parameters. When the bond parameters are set to zero, the SQE method elegantly reduces to the QEq model.

In the context of zeolites, Smirnov *et al.*[28] studied methane adsorption on the surface of MFI- and MEL-type zeolites using QEq charges, which were fit to reproduce Mulliken charges for siliceous zeolite frameworks. Applying the QEq charges to MD simulations, they found that the topology of zeolites plays an important role in determining the degree of polarization of the guest molecules by the host framework. More recently, Vestraelen *et al.*[29] developed both QEq and SQE parameters fit to reproduce DFT derived ESP charges (and Hirschfeld Type I charges) for a large number of silicate cluster compounds. The parameters were then validated on a number of other silicate materials, including periodic zeolite systems that showed good agreement with the DFT derived charges. However, again, in both cases these charges have been developed for charge neutral systems. In the case of the SQE charge parameters developed by T. Verstraelen *et al.*, this is due to that fact that, in its original formulation, the SQE method is limited to charge neutral systems. While a trivial mathematical transformation allows one to apply QEq to charged periodic systems, due to the antisymmetry of the split-charges, no such transformation can be applied to SQE. To overcome this limitation, Krykunov *et al.* introduced a new SQE model, called SQE_{AB} (for atomic basis) that allowed for systems to have a net charge. This new model allows for calculation of atomic charges in charged periodic frameworks by transforming the split-charges from a bond-representation to the atomic basis using a Moore-Penrose pseudoinverse. To validate the method, it was tested on a variety of neutral and charged zeolites and MOFs. When applied to neutral frameworks, SQE_{AB} was shown to generate the same charges as the original SQE model. For charged systems, on the other

hand, SQE_{AB} was shown to accurately capture the chemical structure of aluminosilicate zeolites, by clearly distinguishing between two types of O atoms: those covalently bonded to Al and those bound exclusively to Si. Supporting this result, similar polarization of the O atoms in the same LAU-type framework was observed when comparing with DFT-derived REPEAT charges. Hence, for modeling the electrostatic interactions in charged periodic systems, REPEAT and SQE_{AB} have both been shown to generate chemically intuitive charges.

For modeling the dispersive interactions in aluminosilicate systems, the Lennard-Jones (LJ) potential is typically employed. It consists of two atom-specific parameters, ε and σ, which describe the potential well-depth and the interatomic distance at which the potential is zero, respectively. In general, there are two methods that have been used to optimize these parameters for unique guest-host systems: (1) they are fit to reproduce experimental data or (2) they are fit to results from high-level QM calculations. Empirical fittings procedures have been used extensively to optimize parameters describing dispersive interactions in aluminosilicate systems. In fact, many of the force fields listed in Table 4.1 have been developed using empirical fitting methods. Most of the recent empirical force field optimizations for aluminosilicates have been carried out by Calero *et al.*, who have developed multiple force fields to model various guest-host interactions in aluminosilicate systems.[20,30–32] In particular, these force fields have been optimized to reproduce experimental adsorption data for various gases, including CO_2, CH_4, and N_2, as well as for linear chain alkanes. Empirically derived force fields such as these have successfully been used to model dispersive interactions in both siliceous and aluminosilicate zeolites.[33,34] However, this method is greatly limited by the quantity and quality of the uptake data available. Alternatively, potential parameters are fit to QM calculations of the potential energy surface of the weakly-bound systems. The zeolite force field published by Fang *et al.*[21] provides a recent example of such a first-principles, DFT-derived force

field for predicting CO_2 uptake in aluminosilicates. A major factor effecting the quality of first-principles derived force fields is the quality of the sampled distribution of configurations of the guest-host system. In the optimization carried out by Fang *et al.* proper sampling of CO_2 configurations within the LTA-4A (Si/Al = 1) framework was achieved by combination of random and GCMC sampling. By including configurations taken from GCMC simulations, sufficient sampling of favorable configurations is ensured. Although the force field has been shown to give good agreement with experiment for CO_2 adsorption in LTA-4A, NaY and NaX zeolites at low Si/Al ratios, only a single framework was included in the training set. However, to accurately capture the variety of binding environments present in aluminosilicates, it is important that the training set reflect the diversity in the underlying geometry of these materials. Furthermore, both parameter optimizations discussed required apriori knowledge of either the Al-sitings or the Na^+ ion positions, or both. This, in turn, limits the available data to a small group of frameworks for which the positions have been resolved or the Al distribution does not have an effect on the adsorption properties of the material.

In the context of modeling guest-host interactions in aluminosilicates, this work aims to provide a force field that addresses the use of generic charge sets for modeling electrostatic interactions by offering a more accurate, but similarly efficient alterative. Combining the results obtained by Krykynov *et al.* with the conclusions made by Fischer *et al.*, SQE_{AB} is proposed as the alternative method for modeling the electrostatic interactions in aluminosilicates. Although, electronegativity equalization methods have been applied to silicate systems, to the best of our knowledge, no parameters have been optimized for charged aluminosilicate zeolites. Furthermore, the parameters that are available have been optimized to reproduce the atomic charges from various charge generation methods. A more accurate representation of the electrostatic interactions can be

obtained by fitting parameters to reproduce the QM ESPs of each unique system. Therefore, to ensure a robust charge model, SQE_{AB} parameters will be optimized to reproduce the gauge-corrected QM ESP[10] for a representative training set of both siliceous and aluminosilicate zeolites at different Si/Al ratios and different underlying geometries. In this context, SQE_{AB} provides an alternative method for assigning atomic charges in aluminosilicates, which offers a balance between the speed of generic charges and the accuracy of system specific REPEAT charges. To provide a force field for modeling post-combustion CO_2 capture and natural-gas purification in aluminosilicates, we model the dispersive interactions using a LJ potential. The framework and sodium LJ parameters have been fit simultaneously to 36 experimental isotherms for CO_2, CH_4, and N_2. The uptake data encompasses 3 different framework types (FAU, LTA, MFI) at various Si/Al ratios and temperatures, to ensure transferability. Our force field, which we term the SQE_{AB}-AMP (Split-Charge Equilibration in the Atomic Basis Aluminosilicate MicroPorous) force field, was found to be the most accurate empirical method for determining the ESPs in aluminosilicates. The combination of accuracy and speed offers a significant advantage over the use of generic charges, which generally fail to capture all features of the ESP accurately. In addition to fast and accurate electrostatics, the force field has been optimized for modeling the guests involved in the major gaseous separations performed using aluminosilicates. Comparing to the force fields available for modeling these processes, the SQE_{AB}-AMP model is shown to give more accurate predictions of the gas uptake in aluminosilicates, using a single unified set of parameters. Since charged aluminosilicate zeolites are currently one of the most highly utilized porous materials in industry, having such an accurate and fast force field would be of great value to the zeolite modeling community.

4.3 Methodology and Background Theory

4.3.1 Parameter Optimization Procedure

4.3.1.1 Optimization of SQE_{AB} Parameters for Aluminosilicates

When fitting a mathematical model that gives the best description of the behavior of a given system, one must take care to prepare a robust training set that captures the complexity and diversity of said system. This requires understanding of both the system being model and the mathematical model itself. In the context of parameter optimization for electronegativity equalization models, such as SQE_{AB}, it is important to understand the relationship between the model and the system to which it is being applied. The energy expression for SQE_{AB}, given in Eq. (4.1), demonstrates a dependence upon the local structural environment of a given atom, since the charge and, therefore, the resulting ESP is dependent on the relative positions of the atoms. Additionally, the atom-pair parameters that account for covalent bonding further introduce a dependence on the underlying structure of the system. Therefore, in 3-dimensional porous systems such as zeolites, it is important to take the underlying geometry of the frameworks into consideration. In particular, it is important that the training set is comprised of a diverse set of frameworks that best simulates the possible siliceous and aluminosilicate framework topologies. This will ensure that the model trained will be transferable across different framework types.

$$E_{\mathrm{AB}}^{\mathrm{SQE}}(\mathbf{Q}) = \sum_{i}^{N} \chi_i Q_i + \tfrac{1}{2}\sum_{i}^{N} \kappa_i Q_i^2 + \tfrac{1}{2}\sum_{i}^{N}\sum_{j>i}^{N}\sum_{k}^{N}\sum_{l}^{N} \kappa_{ij}^{b} \left(\mathbf{T}^{+}\right)_{ij,l} \left(\mathbf{T}^{+}\right)_{ij,k} Q_l Q_k + \sum_{i}^{N}\sum_{j>i}^{N}\sum_{k}^{N} \chi_{ij}^{b} \left(\mathbf{T}^{+}\right)_{ij,k} Q_k + \tfrac{1}{2}\sum_{i}^{N}\sum_{j\neq i}^{N} J\left(r_{ij}\right) Q_i Q_j \tag{4.1}$$

In this regard, the frameworks that were selected to populate our training and validation sets were taken from the International Zeolite Association (IZA-SC) database of zeolite frameworks.[35] The database contains all distinct framework topologies for zeolites, which have been experimentally

realized or exists in nature. The first step in the selection process was to identify all frameworks that contain Si and Al. Although Si and Al cannot be distinguished experimentally, the composition of the frameworks can still be determined via combination of XRD and NMR experiments.[36,37] Thus, a chemical formula is listed along with other topological information for each framework. From this chemical formula, it is possible determine the Si/Al ratio of the framework. In some cases, the formula is given such that there is a range of possible Si/Al ratios for a single framework type. For example, the MEI framework has range of Si/Al ratios between ~5 – 15:

$$\textit{Chemical Formula: } |\mathrm{Na}_n^+(H_2O)_{28}|[Al_nSi_{34-n}O_{68}] - MEI,\ n = 2.1 - 5.7$$

There is a total of 232 unique framework types listed in the IZA-SC database. However, not all of these framework types contain Si and Al exclusively but, rather, contain other trace elements such as Be and Zn, or P for aluminophosphate (ALPO) zeolites. After excluding these other groups of zeolites, there was a total of 111 siliceous and aluminosilicate framework types identified in the IZA-SC database. We then divided these frameworks into training and validations sets so as to ensure an accurate representation of the relative proportion geometric features found in all aluminosilicate materials identified in the IZA-SC database. From the 111 frameworks that were identified, 63 unique frameworks were selected to populate the training set while 48 unique frameworks were selected for the validation set.

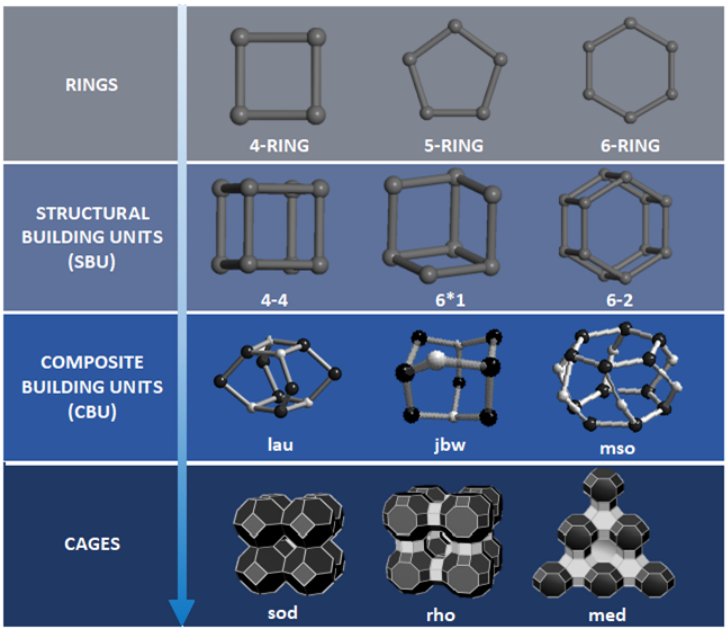

Figure 4.1 *Examples of the different hypothetical structural units listed in the IZA database in order of increasing complexity, from top to bottom.*

To enable comparison between different framework types, zeolites have been broken down into simpler structural motifs. Examples of these structural motifs are given in Figure 4.1, with all variations listed in the IZA-SC database.[35] The simplest of these motifs that was taken into consideration are the rings that are formed upon combination of TO_4 (T = Si or Al) units in the zeolite framework. Generally, when representing these structures, the bridging O atoms are left out for clarity and the vertices represent the T-atoms. Accordingly, the size of the rings is measured by the number of T-atoms that are present in the ring. Structural building units (SBUs), the next structural unit taken into consideration, can contain a maximum of 16 tetrahedrally coordinated T-atoms, and are derived assuming that an entire framework is composed of a single type of SBU. As with all of the sub-structural units used to classify zeolite topologies, there must always be an integral number of SBUs in a unit cell, otherwise the result would be an incomplete structure. Combining together SBUs, composite building units (CBUs) are constructed. The principle difference between SBUs and CBUs is that the latter can contain a much larger number of tetrahedrally coordinated

metal centers and, as a result, form much larger and more complex structures. Combining CBUs, the distinct rings and cages that give rise to the physically important and chemically useful properties of zeolites are formed.

In addition to accounting for the underlying structural features of zeolites, the composition of the frameworks included in the training and validation sets was carefully tailored so as to simulate Al atoms in a diverse set of structural environments. This is especially important to sample during the parameter optimization due to experimental difficulties in resolving Al-sitings in aluminosilicates. Furthermore, due to this lack in experimental data, the crystal structures taken from the IZA-SC database are composed entirely of Si. To add Al to the structures, an algorithm was written to randomly replace Si atoms by Al atoms according to Löwenstein's rule [38], which states that there can be no Al-O-Al linkages. To determine the appropriate ratios to sample in the fitting procedure, an analysis was performed to see which ratios are most frequent amongst the frameworks included in the training and validation sets.

Figure 4.2 summarizes graphically the results of both the structural, and the compositional analysis of the aluminosilicate frameworks included in the training and validation sets. The y-axis corresponds to the proportion of unique frameworks possessing a given property in each set of frameworks considered, while the x-axis corresponds to the property being evaluated (e.g. Si/Al ratio). The results of this analysis clearly show that the frameworks chosen to populate the training and validation sets reproduce the structural features of all siliceous and aluminosilicate frameworks found in the IZA database very well. Figure 2a shows the compositional analysis performed to

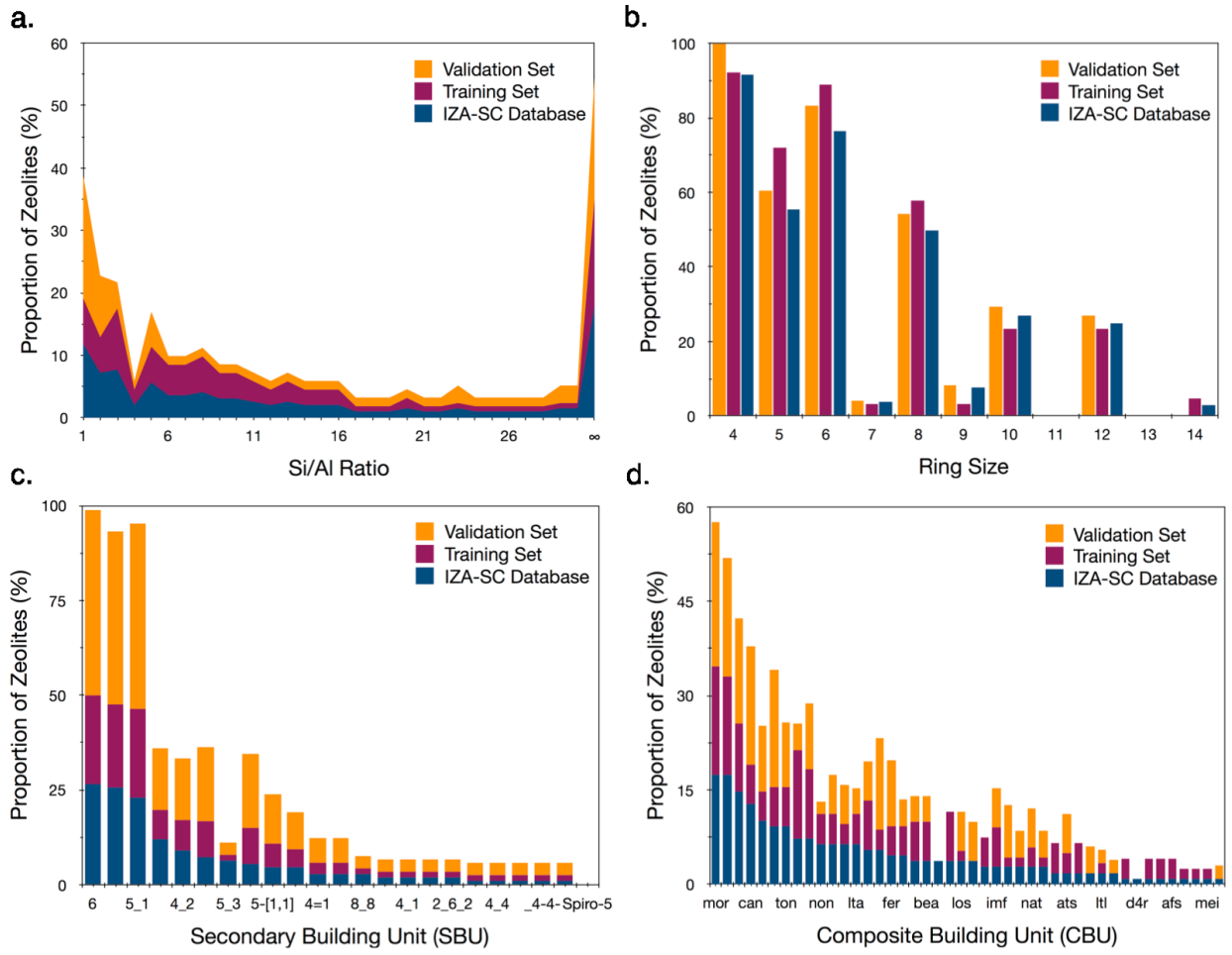

Figure 4.2 (a) *Si/Al Ratio Analysis. The plot shows that Si/Al ratios of 1.0, 3.0 and ∞ are the most populated for the frameworks considered.* ***(b)*** *Frequency of Occurrence of Different Size Rings. Comparison between the number of rings of different size observed in the training and validation sets with respect to all zeolite framework taken from the IZA-SC database.* ***(c)*** *Frequency of Occurrence of Different SBU. Comparison of SBUs found in training and validation sets with respect to the SBUs found in all frameworks containing Al taken from the IZA-SC database.* ***(d)*** *Frequency of Occurrence Different CBU. Comparison of CBUs found in training and validation sets with respect to the CBUs found in all frameworks containing Al.*

identify what Si/Al ratios should be assigned to the frameworks. From this analysis, it is clear that Si/Al ratios of 1.0, 3.0 and ∞ have the greatest frequencies of occurrence for the frameworks considered. Accordingly, the unique frameworks initially selected were then modified to have one of these three ratios. In addition to being the most frequent, using lower Si/Al ratios will allow for a greater diversity of possible Al-sitings to be sampled within a range of topologies. In some cases, unique framework types are assigned various Si/Al ratios, increasing the size of the training and

validation sets. Ignoring the identity (Si or Al) of the tetrahedral sites, the training set was composed of 63 unique framework types, while the validation set was comprised of 48 unique frameworks. After assigning Si/Al ratios, the training set increased in size to 127 frameworks, and the validation set increased to 89 frameworks. A full list of the frameworks in both the training and validation sets are listed along with the corresponding Si/Al ratios that have been assigned in Appendix A (Table 4-S1 – 4-S2).

To fit the parameters, an in-house genetic algorithm (GA) developed for fitting SQE parameters to reproduce the gauge-corrected QM ESP was used.[39] The ESPs resulting from SQE_{AB} point charges are generated using a Wolf summation,[40] which was found to be as accurate as the Ewald summation,[41] while also being significantly more efficient. The QM ESP was obtained for each structure via periodic DFT calculations on a set of real-space grid points. In both cases, the ESP grid points are uniformly spaced by 0.2 Å along all cell dimensions. Additionally, only grid points which fall between 1 and 2 VdW radii of the framework atoms were considered valid and used during the fitting procedure. The SQE_{AB} parameters are then simultaneously optimized to minimize all frameworks' mean absolute difference between the ESP resulting from SQE_{AB} charges, and the QM calculation. In the first step of the custom GA used to optimize the SQE_{AB} parameters, randomly generated parameter sets are constructed, which collectively are known as a generation. Each set of parameters is then evaluated for how well they reproduce the QM ESP in each framework considered, and a new generation is created using a roulette wheel selection algorithm that selects two solutions, or *chromosomes*, which will act as the parents of the new solution, or offspring. This process is biased toward more fit parameter sets that better reproduce the QM ESP, such that the fittest features are propagated to future generations. The mating algorithm then generates a new set of parameters by randomly selecting a value for each parameter that is between

the values of the parameter for both parents. Next, the offspring are allowed to undergo mutations that can alter the parameter by ±30%, with a probability of 25%. The GA was considered fully converged after the best solution remained the same for ten generations. After five independent runs of the GA, the best set of parameters was further refined using a gradient descent method, to yield a set of optimized parameters we refer to as the SQE_{AB}-AMP. A more in-depth discussion on the general features of the GA and the gradient descent algorithm have been included in Chapter 2, and specific details related to the GA used for this optimization are given in Appendix B of this chapter.

4.3.1.2 Optimization of Van der Waals Potential Parameters

The aim of this force field is to model the guest host interactions involved in post-combustion CO_2 separations and methane gas purification using aluminosilicate zeolites. To model the dispersive interactions, a Lennard-Jones potential is used, with the framework atom and sodium cation parameters optimized empirically. To ensure that the parameters are compatible with popular guest force fields, the guest parameters have not been altered during the optimization procedure. Rather, force fields for CO_2, CH_4, N_2 that have been used extensively in modeling gas uptake in aluminosilicates have been adopted, and have been summarized in Table 4.2. It should be noted that the LJ parameters for the sodium cation are generally included as a free variable during the parameter optimization, such as was done in the in this work. For the AMP force field, the charge on Na^+ was chosen to be +1.0 e so that the extra-framework sodium densities match the framework aluminum densities, where each Al atom contributes -1.0 e to the net framework charge. In general, the force field parameters for guest molecules are fit to reproduce experimental phase-equilibria data, typically the vapor-liquid coexistence curve.

Table 4.2 *Summary of the guest force fields used in recent zeolite force fields, and in this work.*

	Calero *et al.*	AMP	Fang *et al.*	References
Na^+	+0.3834 e charge LJ parameters are free variables during optimization	+1.00 e charge LJ parameters are free variables during optimization	+0.99 e charge S. Grimme[42]	Calero *et al.*[20] Fang *et al.*[21]
CO_2	J.G. Harris	J.G. Harris	J.G. Harris	Harris *et al.*[43]
CH_4	TraPPE	5-site		TraPPE[44] 5-Site[45]
N_2	TraPPE	TraPPE		TraPPE[46]

Across the 3 force fields presented in Table 4.2, the CO_2 guests are modeled using the same parameters, published by Harris *et al.*[43] The charges on the C and O atoms were selected to reproduce the experimental quadrupole moment of CO_2, while the LJ parameters were fit to reproduce the liquid-vapor coexistence curve of CO_2 . To model methane, Calero *et al.*[47] use a single-site TraPPE force field, while for the AMP force field we have adopted a 5-site model. As the names indicate, the TraPPE force field employs a united-atom (UA) description of methane, where the entire molecule is modeled as a single spherical particle. Conversely, in the 5-site model, each atom in methane is treated explicitly. Although both models have been shown to accurately describe fluid-phase behavior, the 5-site model has been shown to better describe dense packing of methane in small pores.[48] Due to packing effects in confined spaces, when methane is modeled as a spherical particle as in the 1-site model, methane molecules more easily pack together. Consequently, methane uptake at higher pressures, when molecules are more densely packed within the pores, is generally over-estimated when using a single-site model. Additionally, the TraPPE model represents CH_4 as a single sphere and, therefore, the charge on this sphere must be zero to maintain charge neutrality. On the other hand, the 5-site model assigns charges to each atom in the molecule, allowing for the atoms in methane to interact electrostatically. This, in turn, will provide a more accurate

representation of the behavior of methane within the confined space of a zeolite pore, where neighboring molecules are more likely to induce temporary dipoles. Though the AMP force field does not explicitly account for polarization, allowing for such electrostatic interactions between guest provides a more realistic description of the intramolecular interactions occurring within the pores of aluminosilicate materials. This is particularly desirable when modeling natural gas purification, which is carried out at pressures of up to 20 bar, where methane molecules are more densely packed inside the pores of the zeolite. Ergo, we have adopted the 5-site methane model in the AMP force field, since it more realistically represents the phase behavior of pure methane in small pores and at higher pressures.

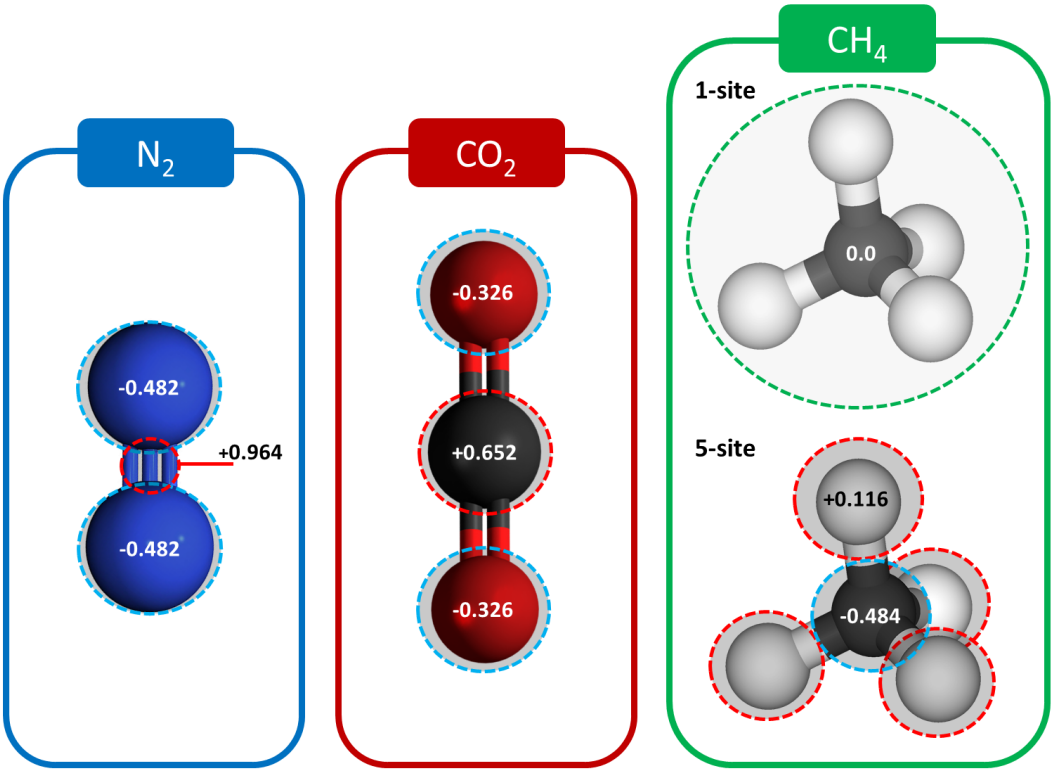

Figure 4.3 *Summary of the different guest models investigated in this work. The COM charge on N_2 (shown with red) does not contribute to the dispersive interactions of the molecule.*

Finally, to model N_2 we have adopted the same TraPPE model used by Calero *et al.*[47], which again has been fit to reproduce phase-equilibrium data, and consists of three sites. Each of the

nitrogen atoms are modeled as VdW interaction sites, which are separated by an experimental bond length of 1.10 Å. The point charges assigned to the nitrogen atoms ($q_N = -0.482\ e$), reproduce the gas phase quadrupole moment of N_2 ($Q = -4.65 \times 10^{-40}$ C m^2).[49] To maintain charge neutrality of the molecule, a third charge is assigned to the center-of-mass (COM). This additional site, however, does not act as a VdW interaction site, and only interacts electrostatically ($q_{COM} = +0.964\ e$). A summary of each of the guest models is given in Figure 4.3, where each dashed circle represents a distinct interaction center. Further, an extensive list of the guest force field parameters considered in this work is given in Appendix A (Table 4-S6).

Since the aim of the AMP force field is to model post-combustion CO_2 capture and natural gas purification in aluminosilicates, the training data was selected within temperature and pressure ranges relevant to these processes. For post-combustion CO_2 capture adsorption typically occurs at 298 K and 1.0 bar.[50–52] while desorption occurs at 8.0– 20 bar and 298 K during natural gas purification.[53–55] These temperatures and pressures are reflected in the training data, which spans temperatures between 248– 373 K and pressures between 0.1 – 10 bar. To incorporate structural and chemical diversity into the training set, the data also spans 3 unique framework types: FAU, MFI, and LTA at various Si/Al ratios. REPEAT charges have been assigned to the framework atoms for all frameworks used during the optimization procedure. An extensive list, including references, for all of the gas uptake data used in both parameter training and validation are included in Appendix A (Table 4-S7 – 4-S8).

As is common practice when simulating adsorption in cation exchanged zeolites,[20,30,56] the zeolite frameworks are assumed to be rigid and the cations are allowed to move freely about the simulation cell. To provide more flexibility in our fitting procedure, we optimize the cross-species parameters representing the distinct intramolecular interactions present by simultaneously fitting them to 36

experimental isotherms. This data has been taken from independent sources, with 12 isotherms for each guest being modeled (CO_2, CH_4, N_2). In total we have identified 20 cross-species interactions, after excluding $Na^+\cdots Na^+$, $Na^+\cdots Si$, and $Na^+\cdots Al$ interactions since these species are prevented from approaching each other due to repulsive Coulombic interactions.

In the work presented, we adopt an approach similar to that employed by Fang *et al.*[21,57] This scheme relies on the assumption that the guest-host interactions in zeolites are represented by pairwise VdW and Coulombic terms, which can be expressed as,

$$E_{Tot} = E_{VdW} + E_{Coul} = s_{12}\frac{C_{12}^{ij}}{R_{ij}^{12}} - s_6\frac{C_6^{ij}}{R_{ij}^6} + \frac{q_i q_j}{R_{ij}}. \quad (4.1)$$

Here, R_{ij} is the interatomic distance separating atom *i* from atom *j*, C_6^{ij} and C_{12}^{ij} are semi-empirical attractive and repulsive coefficients, q_i and q_j correspond to the atomic charges on atom *i* and *j*, and s_6 and s_{12} are global scaling factors. The C_6^{ij} and C_{12}^{ij} parameters represent the attractive and repulsive coefficients for a given atom-pair interaction, respectively. The repulsive coefficients, C_{12}^{ij}, are defined by,

$$\frac{C_{12}^{ij}}{C_6^{ij}} = \frac{\left(R_0^i + R_0^j\right)^6}{2}, \quad (4.2)$$

where R_0^i and R_0^j are the VdW radii of atoms *i* and *j*. This relation is based on the similarity in the VdW terms included in Eq. (4.1) and those in the LJ potential with the form

$$E_{LJ}(R_{ij}) = 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{R_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{R_{ij}}\right)^6\right] = \frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6}, \quad (4.3)$$

where ε_{ij} corresponds to the potential well depth, and σ_{ij} is the interatomic distance at which the potential is zero. In this work we have adopted the atomic C_6 and R_0 parameters for the relevant elements from Grimme's work in our fitting procedure.[42] From the atomic parameters, the cross-species attractive coefficients, C_6^{ij}, are determined as the geometric mean of the corresponding atomic terms for atoms *i* and *j*. During the fitting procedure, the Coulombic interactions are modeled using REPEAT charges that are calculated prior to the parameter optimization for the frameworks considered.

Comparing Eq.(4.1) with Eq(4.3) it is clear that $A_{ij} = s_{12}C_{12}^{ij}$ and $B_{ij} = s_6 C_6^{ij}$. Further comparison allows one to derive expressions for ε_{ij} and σ_{ij} in terms of the scaling factors and the attractive and repulsive coefficients, where

$$\sigma_{ij} = \sqrt[6]{s_{12}C_{12}^{ij}/s_6 C_6^{ij}}, \quad (4.4)$$

and,

$$\varepsilon_{ij} = \frac{s_6 C_6^{ij}}{4\sigma_{ij}}. \quad (4.5)$$

Using these expressions, the corresponding cross-species parameters for the LJ potential can be calculated, and the scaling factors are used as the free variables during the fitting procedure. In the parameter optimization carried out by Fang *et al.*, the scaling factors were determined to allow the closest correspondence between the first-principles interaction energy and the classical force field energy. Two sets of scaling factors were used, with one set used to model guest-host interactions present in siliceous zeolites(Si···X, O···X where X is a guest atom), and a second set to model the additional pairwise interactions present in aluminosilicates (Al···X, Na^+···X). Although using

different scaling factors, s_6 and s_{12}, for each cross-species interaction would give a more flexible model, Fang *et al.* found that this strategy lead to a strong coupling between the magnitude of the scaling factors, and resulted in unphysical potential parameters (negative values). Adopting a similar scheme in our parameter optimization, we have a total of 4 scaling factors (2 pairs of attractive and repulsive factors) for each of the guests modeled, giving a total of 12 scaling factors overall. In this work, the scaling factors are varied simultaneously, and fit empirically to give the closest correspondence between the simulated and experimental uptakes for each of the guests considered.

The advantage of employing this approach is that it only requires a single constraint to be enforced during the parameter optimization to ensure physically relevant parameters: the scaling factors must have non-negative values. Removing the ambiguity in selecting constraints on the parameters during the optimization procedure was the principal motivation for following a similar procedure in this work. Though fewer constraints should generally result in a more flexible parameterization and, ultimately, a more flexible model, the fewer free variables used in the fitting procedure results in the opposite. If one considers the interactions between a framework oxygen with CO_2, there are two unique cross-species interactions that have to be accounted for:

1) the interaction of the zeolite oxygen with the carbon of CO_2 ($O_z \cdots C_x$),
2) the interaction of the zeolite oxygen with a CO_2 oxygen ($O_z \cdots O_x$).

Because only a single set of scaling factors, s_6 and s_{12}, are used to describe both interactions, both interactions must either be scaled up or down simultaneously. Furthermore, since the Si interactions are described using the same scaling factors, the $Si \cdots C_x$ and $Si \cdots O_x$ interaction potentials are scaled by the same amount. The same situation arises with the Al and Na^+ interactions, which are also described together by a single pair of their own scaling factors, s_6' and s_{12}'. Notwithstanding, this

approach has been adopted as we expect the use of physically motivated charges will result in a more accurate model than those presently available, which employ generic charge sets.

Optimization of the scaling factors was carried out using a custom genetic algorithm (GA), with a population size of 30. More specific details for the GA are outlined in Appendix C. In total, the GA yields 21 unique cross-species LJ parameters, which have been fit simultaneously to 39 experimental isotherms taken from independent sources, with 13 isotherms for each guest being modeled (CO_2, CH_4, N_2). A lack of quality adsorption data for different guests and a variety of framework types is compounded by the inability to reliably locate Al atoms in aluminosilicate frameworks. The latter means that only certain framework types for which either the Al locations have been identified, or for which the uptake properties are independent of the Al distribution can be used in the optimization procedure. Thus, to ensure the most robust model parameters, all uptake data that was obtained from the literature was included in the training set. To evaluate the fitness of the parameters, the percent error (PE) between the simulated and experimental uptake is calculated over 6 pressure points for each isotherm. The overall fitness is then taken as the average of the calculated PEs for each simulated isotherm, with the deviations grouped by isotherm and by guest. The PE is calculated according to the following expression:

$$PE_{ij} = \frac{1}{N_{points}} \sum \left[\left(\frac{\left| U_{sim,l} - U_{exp,l} \right|}{U_{exp,l}} \right) \cdot 100 \right], \quad (4.6)$$

where $U_{sim,l}$ and $U_{exp,l}$ are the simulated and measured uptakes at pressure *l*, respectively, and N_{points} is the total number of points evaluated (N_{points}= 6 in all cases) in isotherm *i* for guest *j*. It should be noted that, although the parameters are fit simultaneously for all guests considered, convergence is evaluated with respect to each guest independently. In this way, when the GA converges on a set of cross-species parameters involving a given guest, they are no longer varied and

the optimized parameters are retained. At this point in the optimization procedure the $Na^{+}\cdots O_z$ parameters are also held constant, ensuring that the framework oxygen and extra-framework sodium cross-species parameters are equal regardless of what gas is being adsorbed. This is highly desirable as this interaction should be entirely independent of the guest's identity.

4.3.2 Computational Details

All periodic DFT calculations were performed using the Vienna Ab Initio Simulation Package (VASP)[58] with the project augmented wave (PAW) method[59] and plane wave basis set. A plane wave cutoff of 500 eV was used, and only the gamma point in the Brillouin zone was sampled. The Perdrew-Burke-Ernzehof (PBE) exchange correlation functional[60] was utilized for all DFT calculations. From these calculations, the QM ESP was calculated for each periodic system utilized in this work. For a full list of the zeolite frameworks along with the corresponding Si/Al ratios that were used, refer to Appendix A (Table 4-S1 – 4-S2). The adsorption properties were determined using an in-house GCMC[61] based on the DL_POLY 2 molecular dynamics package,[62] which has been applied to studying gas adsorption in MOFs.[63–65] All guest-framework repulsive steric and attractive dispersion interactions were calculated using LJ potentials. For the validation of SQE_{AB} parameters for aluminosilicates, the CO_2 adsorption isotherms were calculated with LJ parameters for the framework atoms assigned from the UFF,[66] and CO_2 parameters taken from Harris *et al.*[43] In all cases, 30,000 cycles were used for both equilibration and production phases of the GCMC, where each cycle corresponds to a number of steps equivalent to the number of guests in the system. For the GCMC calculations carried out during the LJ parameter optimization, REPEAT charges were assigned to all frameworks. Framework and sodium LJ parameters were simultaneously fit to reproduce experimental gas uptake data for CO_2, CH_4, and N_2, while the guest parameters were held constant. For efficiency, 3 million equilibration and 4.5 million production steps are used for the

GCMC calculations performed during parameter optimization, since this was determined as the minimum number of steps to achieve reliable sampling (see Figure 4-S1 in Appendix A). All GCMC calculations performed for LJ parameter validation, including gas uptake and heats of adsorption (HoA) calculations, were performed using 30, 000 cycles for equilibration and production.

4.4 Results and Discussion

4.4.1 SQE_{AB} Parameter Optimization

In this work, we present SQE_{AB} parameters that have been optimized for 4 atom types and 3 bond types that encompass all siliceous and aluminosilicate zeolites. Our model discriminates between two types of oxygen atoms: those covalently bonded to other Si atoms exclusively, O, and those which are bound to an Al atom, O_{Al}. Figure 4.4 illustrates the atom-typing scheme that was adopted for this parameter optimization, which reflects how the O atoms which are bound to Al in aluminosilicates tend to be more polarized then those which are bound exclusively to Si. By accounting for this difference, the model will more accurately predict the charge distribution in both siliceous and aluminosilicate zeolites. The atomic radii that are used in this work are equal to 1.5 times the covalent bond radii for each element, taken from OpenBabel.[67] The radii are important parameters in the SQE_{AB} model, which define when the conventional Coulomb potential is damped at close range.

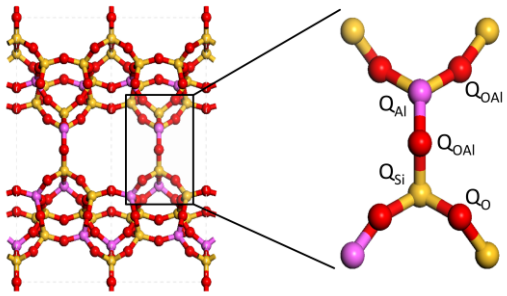

Figure 4.4 *Atom-typing scheme applied in this work. Discriminating between two types of oxygens helps accurately capture the polarization of the charge on the O atoms adjacent to Al atoms in aluminosilicates.*

Table 4.3 *Optimized SQE_{AB}-AMP Parameters*

atom	χ_i (eV)	κ_i (eV)	bond	χ_{ij} (eV)	κ_{ij} (eV)
O	46.426	46.788	O_{Al}-Al	-0.52862	0.02426
O_{Al}	49.825	37.694	O_{Al}-Si	-3.4197	0.045871
Si	0.11856	36.861	O-Si	-1.7844	0.0061678
Al	0.00018457	24.963			

The optimized SQE_{AB}-AMP (Aluminum MicroPorous) parameters are given in Table 4.3. A total of 16 unique parameters are listed, although the atomic electronegativities can be accounted for using combined bond electronegativity parameters,[68] reducing the total number of unique parameters to 10. To validate the parameters, we perform a series of test to evaluate how well the generated charges reproduce the QM ESP for a collection of zeolite framework types not included in the training set. In the first test, we compare directly between different methods for modeling the electrostatics in aluminosilicates to see how well the SQE_{AB}-parameters perform relative to current methods. Of the models considered, REPEAT most accurately reproduces the QM ESP, and is therefore considered as the target solution. We further compare the results to the case where all framework charges are set to zero, giving an indication of the upper limit on the quality of the parameters. Given in Figure 4.5a are the MADs between the QM ESP and the ESP generated by a given charge model (i.e. REPEAT,[10,24] Verstraelen *et al.*,[29] Fang *et.,*[21] Calero *et al.*[20]), for each framework in the training set. The materials are arranged from left to right in ascending order of the corresponding MAD. Figure 4.5b shows the same for the validation set, which contains 48 unique framework types not contained in the training set. Figure 4.5 shows that the REPEAT charges best reproduce the QM ESP for both the training and validation frameworks, with the SQE_{AB}-AMP model lying between REPEAT and the other methods considered. The generic charge models

proposed by either Fang *et al.* and Calero *et al.* outperform the SQE model developed by Verstraelen *et al.* to reproduce ESP-derived charges for silicates, over all structures. Yet, of all the empirical methods studied, the SQE_{AB}-AMP model provides the best descriptions of the QM ESP over all structures in both the training and validation sets.

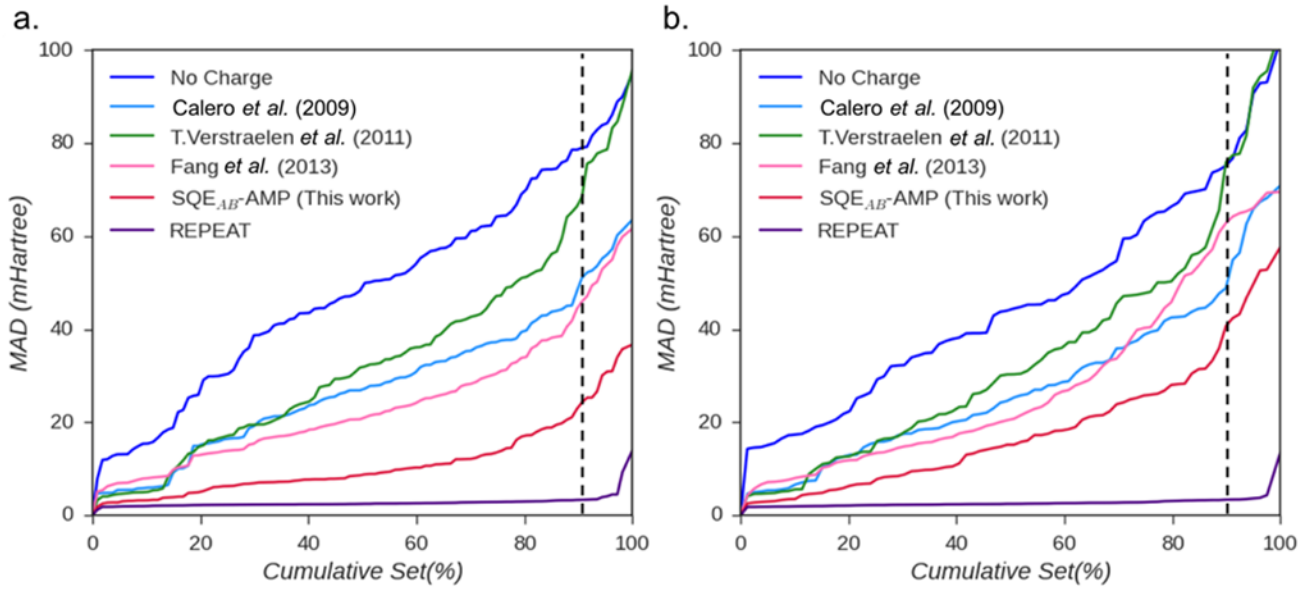

Figure 4.5 *MAD between the QM ESP and the ESP calculated using different charge models for each framework in the* ***(a)*** *training set and* ***(b)*** *validation set sorted in ascending order of MADs. The dashed line corresponds to 90% of the frameworks in the corresponding set.*

Of particular interest for this work is the performance of the SQE_{AB}-AMP model relative to the generic charge models typically employed for aluminosilicates, where system specific SQE_{AB}-AMP charges outperform Fang *et al.* by 22.16% and Calero *et al.* by 35.98% when one averages over the training and validation sets. The results of the comparison between charge methods are summarized in Table 4.4. In addition to listing the mean deviations from the QM ESP for each method, the table also gives the MADs from the QM ESP which encompass 90% of the training and validation set. This number gives an indication of the transferability of the different models to different framework types. For example, if we consider the top 90% of the frameworks in the training set based on how well the atomic charges in these structures reproduce the QM ESP, the

maximum MAD is 20.6 mHartree for SQE_{AB}-AMP, whereas for the generic charge sets published by Fang *et al.* and Calero *et al*, it is 40.6 and 45.0 mHartree, respectively. Of the empirical methods assessed in Figure 4.5, it is clear that the SQE_{AB}-AMP model provides the best description of the QM ESP over all siliceous and aluminosilicate zeolite frameworks considered. To further ensure that the model parameters are not overfit to any features present in the training set and, hence, that the model will be transferable across framework types, we compare the mean of the MADs from the QM ESPs across the training and validation sets. The REPEAT method provides the ideal case, where the mean remains constant between the two sets of frameworks. For the SQE_{AB}-AMP model, the results in Table 4.4 show that there is relatively small change of 1.1 mHartree in the mean deviations across the training and validation sets. On the other hand, for the generic charge models the shift in the mean of the MADs from the QM ESP from the training set frameworks to the validation set framework is on the order of 0.2 – 3.4 mHartree. Interestingly, it should be noted that although the mean in the deviations is greater for the Calero charges relative to SQE_{AB}-AMP charges, this generic charge set is seen to provide a consistent description of the ESP across all frameworks considered.

Table 4.4 *Summary of comparison of ESPs for various charge generation methods to the QM ESP in mHartree*

	Training Set		**Validation Set**	
Charge model	mean	max at 90%	mean	max at 90%
No charge	48.1	76.3	45.1	74.3
Verstraelen *et al.*	33.5	61.7	34.0	64.0
Calero *et al.*	27.5	45.0	27.3	47.7
Fang *et al.*	23.6	40.6	27.5	60.8
SQE_{AB}-AMP	11.0	20.6	12.1	35.9
REPEAT	2.7	3.3	2.7	3.2

To validate the inclusion of additional parameters that distinguish between two types of oxygen atoms, we compare the partial atomic charges calculated with SQE_{AB} and QEq, with and

without atom-typing, to the framework REPEAT charges for the MFI-type zeolite (Si/Al = 3). The parameters used in all cases were optimized following the same procedure that was used to obtain the SQE_{AB}-AMP parameters. A full list of all parameters that have been used for this work can be found in the Appendix A (Table 4-S3 – 4-S5). The results of this comparison are summarized in Figure 4.6, where the additional atom-typing clearly provides a better description of the chemical structure of charged aluminosilicate systems. What is immediately noticeable is that the two oxygen atom types are clearly grouped with the Al-neighboring atoms being more negatively charged than the non-neighboring O atoms, for both QEq and SQE_{AB} when atom-typing is implemented. For SQE_{AB} in particular, the additional parameters result in a framework oxygen charge distribution that matches that calculated with the REPEAT method exceptionally well. Beyond providing further rationalization for implementing an atom-typing scheme in our charge model, this result further validates the quality of the SQE_{AB}-AMP method for generating meaningful atomic charges.

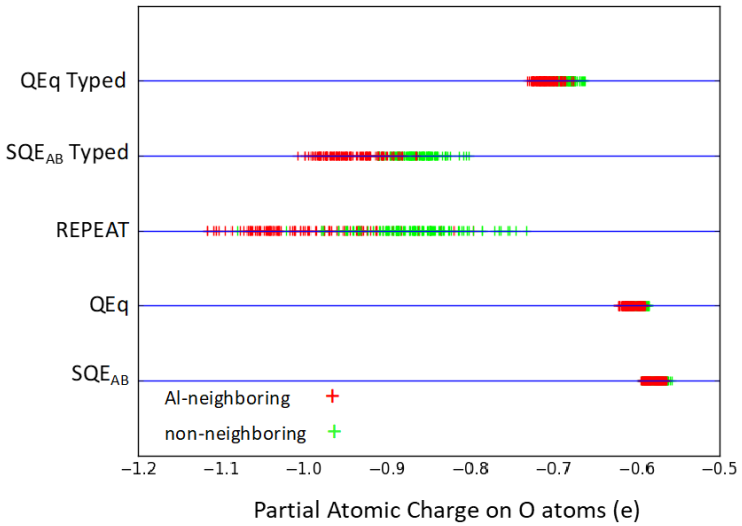

Figure 4.6 *The partial atomic charges on oxygen atoms in a charged MFI (Si/Al = 3) zeolite framework calculated with various methods.*

As a final assessment of the performance of the SQE_{AB}-AMP model, we have evaluated the CO_2 uptake in a series of 29 siliceous and aluminosilicate frameworks, using either SQE_{AB}-AMP or

REPEAT to model the electrostatic interactions. This test is similar to the one carried out by Fischer and Bell,[22] in which they compared the CO_2 uptakes predicted when using either generic charges or REPEAT charges to model the electrostatic interactions. A Lennard-Jones potential is used to model the dispersion interactions, with all framework parameters taken from the UFF force field.[69] For the guest potentials, CO_2 is modeled using the force field published by Harris *et al.*,[43] while the sodium ions are modeled using the force field published by Calero *et al.*[20] The charge on the sodium ions, however, has been adjusted to +1.0 e to match the framework charges imparted upon by the Al atoms. By comparing the simulated uptakes obtained with either charge model, we are able to indirectly evaluate how well the SQE_{AB} charges represent the ESP of a given system. In particular, since CO_2 has a non-negligible quadrupole moment of -13.4×10^{-40} C m^2,[49] we expect that the electrostatic interactions to play an important role in determining the CO_2 adsorption properties of siliceous systems, but notably so for charged aluminosilicate frameworks. Since REPEAT charges have already been shown to reproduce the QM ESP very well, if the SQE_{AB} charges accurately describe the potential in siliceous and aluminosilicate systems, the simulated uptakes should be very similar between the two charge methods for each unique framework considered. To ensure the transferability of these results, a representative cross-section of the frameworks included in the validation set was used for this analysis.

Given in Figure 4.7a is the root-mean-squared percentage error (RMSPE) between the CO_2 isotherms calculated using a REPEAT charges and the corresponding isotherms calculated using SQE_{AB}-AMP charges. In the study carried out by Fischer and Bell, the CO_2 uptakes calculated using generic charges to model the electrostatic interactions in siliceous zeolites were found to differ by an RMSPE of more than 10% in half the cases considered, with a maximum RMSPE of ~23%. Conversely, the RMSPE does not exceed 5% for any of the 29 siliceous and aluminosilicate

frameworks assessed in this work. On average, the SQE_{AB}-AMP charges have been found reproduce the simulated CO_2 adsorption isotherms obtained using REPEAT charges within an RMSPE of 1.47%, with a maximum RMSPE of 4.38%.

Given in Figure 4.7b are the SQE_{AB}-AMP uptakes plotted against the uptakes obtained using REPEAT charges, where the black line indicates an ideal 1:1 correspondence. From this direct comparison of the results obtained using either charge method, an unequivocal relationship between the simulated CO_2 uptakes is observed. Closer inspection reveals a small systematic over-estimation at low pressures, and a small systematic under-estimation at higher pressures relative to REPEAT. However, these deviations are negligible in all the cases considered (RMSPE < 5%). To corroborate this, the Pearson correlation coefficient and the Spearman rank-order correlation coefficient have been calculated, and are listed in Table 4.5. Combined, the Pearson correlation of 0.989 and the Spearman correlation of 0.992 clearly demonstrate that the SQE_{AB}-AMP model provides an approximation of the ESPs in charged and neutral zeolites which are in excellent agreement with the REPEAT method. Conversely, Fischer *et al.* found that their generic charges failed to reproduce the QM ESP as well as REPEAT charges, resulting in errors exceeding 10% in most cases. Here, the difference simulated uptakes obtained using SQE_{AB}-AMP charges and REPEAT charges are all within 5% suggesting that, for the same set of LJ parameters, the two charge methods should yield similar predictions of the adsorption properties of a given zeolite. In the following section, we will turn our attention to the development of a force field for modeling gas adsorption in zeolites, for which both SQE_{AB}-AMP and the REPEAT charges can be used in place of the generic charges sets that are typically used for modeling the electrostatic interactions.

Table 4.5 *RMSPEs, Standard Deviation, and Pearson and Spearman Rank Order Coefficients for CO_2 Uptake in siliceous and aluminosilicate zeolites at 298 K and between 0.1 - 1.0 bar.*

Pearson (R^2)	Spearman (R^2)	$RMSPE_{max}$ (%)	$RMSPE_{min}$ (%)	$\overline{RMSPE}$ (%)	Standard Deviation (%)
0.989	0.992	4.38	0.387	1.47	0.907

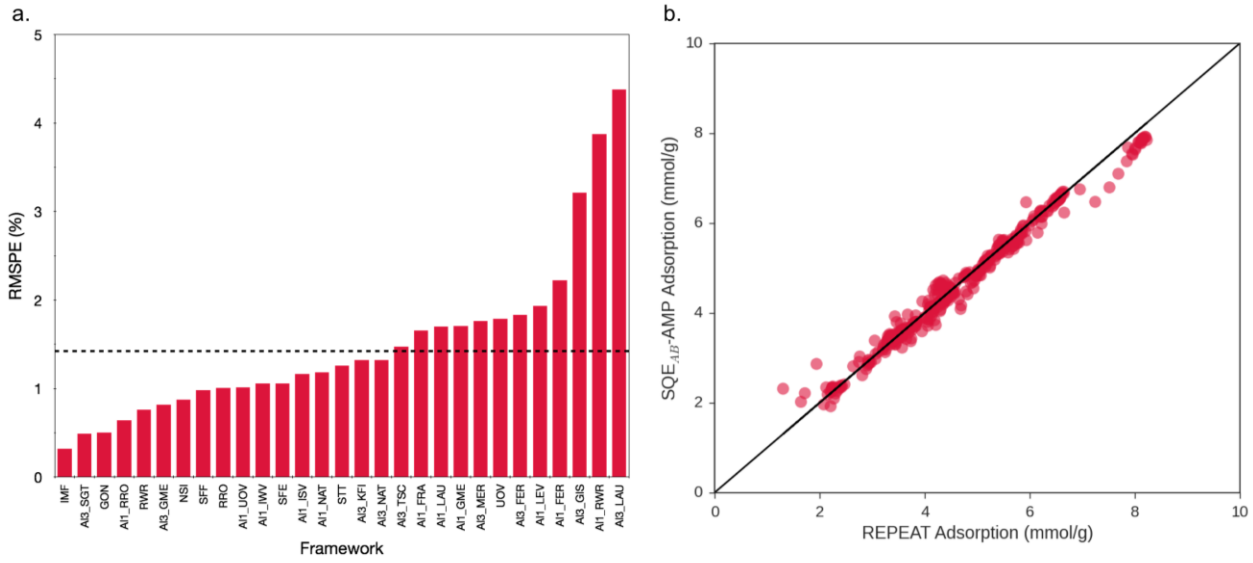

Figure 4.7 (a) *Root-mean-square percentage error (RMSPE) between the CO_2 uptake calculated using a REPEAT/UFF force field and the CO_2 uptakes obtained with a SQE$_{AB}$-AMP/UFF force field. The dashed indicates the average over all 29 frameworks considered (~1.5%).* ***(b)*** *SQE$_{AB}$-AMP/UFF CO_2 uptakes plotted against the uptakes predicted using REPEAT charges. All isotherms were calculated at 298 K, and at 10 pressure points between 0.1 – 1.0 bar. The black line shows the ideal 1:1 correspondence.*

4.4.2 Lennard-Jones Parameter Optimization

In the second half of this work, we focus on the validation of LJ potential parameters for modeling the VdW interactions of CO_2. N_2, and CH_4, with aluminosilicate zeolite frameworks. The optimized cross-species terms describing these interactions are summarized in Table 4.6, and divided by guest for clarity. As the results of section 4.3.1 suggest, using SQE$_{AB}$-AMP charges in place of the REPEAT charges used on the frameworks atoms during the optimization procedure should not result in a significant decrease in the accuracy of the adsorption. Therefore, in order to demonstrate the validity of the model parameters using either charge model, we evaluate their performance in

predicting three different physical observables directly related to accuracy of the simulated guest-host interactions. We then compare our results with other force fields found in the literature.

First, the simulated uptakes are compared to the experimental uptakes for each of the guests modeled, and further evaluated against the Calero *et al.* and Fang *et al.* force fields. The Calero force fields were chosen since there are parameters available for all the guests modeled in this work. Furthermore, the Calero force fields that are used for this comparison are amongst the most highly sited for the types on interactions that the AMP force field aims to model. The AMP force field is additionally compared with the Fang *et al.* force field, allowing the AMP model to be evaluated against first-principles derived LJ parameters for CO_2 adsorption in zeolites. Moreover, the approach taken in optimizing the AMP parameters was adapted from the work of Fang *et al.* The results of these comparisons are presented in Figure 4.8 a – 4.10 a. To further assess the performance of the AMP field, the heats of adsorption have been evaluated with respect to experiment, and compared with the Calero *et al.* Fang *et al.* force fields. The results of this analysis are presented in Figure 4.8 b – 4.10 b. It should be noted that results for Fang *et al.* are only presented for CO_2 as the force field was only optimized to describe CO_2 adsorption zeolites.

Table 4.6 *Optimized Lennard-Jones Parameters by Guest*

CH_4			CO_2			N_2		
Cross species	ε/k_B (K)	σ (Å)	Cross species	ε/k_B (K)	σ (Å)	Cross species	ε/k_B (K)	σ (Å)
$Na^+ \cdots O_Z$	21.35	2.888	$Na^+ \cdots O_Z$	21.35	2.888	$Na^+ \cdots O_Z$	21.35	2.888
$Na^+ \cdots C_M$	79.68	2.792	$Na^+ \cdots C$	39.61	2.963	$Na^+ \cdots N$	42.89	2843
$Na^+ \cdots H_M$	70.92	2.306	$Na^+ \cdots O$	81.45	2.927	$Al \cdots N$	20.23	3.398
$Al \cdots C_M$	38.45	3.324	$Al \cdots C$	19.11	3.527	$Si \cdots N$	31.66	3.699
$Al \cdots H_M$	28.03	2.839	$Al \cdots O$	15.03	3.401	$O_Z \cdots N$	18.765	3.256
$Si \cdots C_M$	35.30	3.694	$Si \cdots C$	38.28	3.729			
$Si \cdots H_M$	5.51	4.078	$Si \cdots O$	29.89	3.600			
$O_Z \cdots C_M$	20.67	3.257	$O_Z \cdots C$	22.41	3.289			
$O_Z \cdots H_M$	16.78	2.732	$O_Z \cdots O$	17.99	3.160			

C_M and H_M correspond to a methane carbon atom and a methane hydrogen atom, respectively.
O_Z is used to distinguish framework oxygen atoms from guest oxygen atoms.

Considering Figure 8a, the predicted N_2 uptakes and heats of adsorptions are presented, respectively. All of the models are shown to be in excellent agreement with experiment over the conditions and framework types considered. In general, the AMP force field is seen to under predict the N_2 uptakes at higher loadings. In contrast, the Calero *et al.* force field is shown to consistently overestimate the gas uptakes across all of the conditions and framework types considered. Looking at the results of the numerical evaluation of the AMP force field given in Table 4.7, the average relative percent error, $\overline{PE_r}$, which considers the average differences between predicted and experimental uptakes, further confirms these observations. The average PE, on the other hand, considers the absolute differences between the simulated and experimental uptakes, and is also listed in Table 4.7 along with the associated standard deviation in the predicted uptakes. The formula for the relative percent error, PE_r^{ij}, is defined as:

$$PE_r^{ij} = \frac{1}{N_{points}} \sum \left[\left(\frac{U_{sim,l} - U_{exp,l}}{U_{exp,l}} \right) \cdot 100 \right], \quad (4.7)$$

where $U_{sim,l}$ and $U_{exp,l}$ are the simulated and measured uptakes at pressure *l*, respectively, and N_{points} is the total number of points evaluated (N_{points}= 6 in all cases) in isotherm *i* for guest *j*. The percent error is defined in Eq. (4.6) in the methods section, with the only difference being that PE considers the absolute differences between the simulated and experimental uptakes, while PE_r^{ij} considers the relative differences. The latter provides insight on the direction of the error (i.e. over-prediction or under-prediction). The average relative percent error is then defined as the average of the PE_r^{ij} over the number of isotherms, N_{iso} :

$$\overline{PE_r} = \frac{1}{N_{iso}} \sum PE_r^{ij}. \quad (4.8)$$

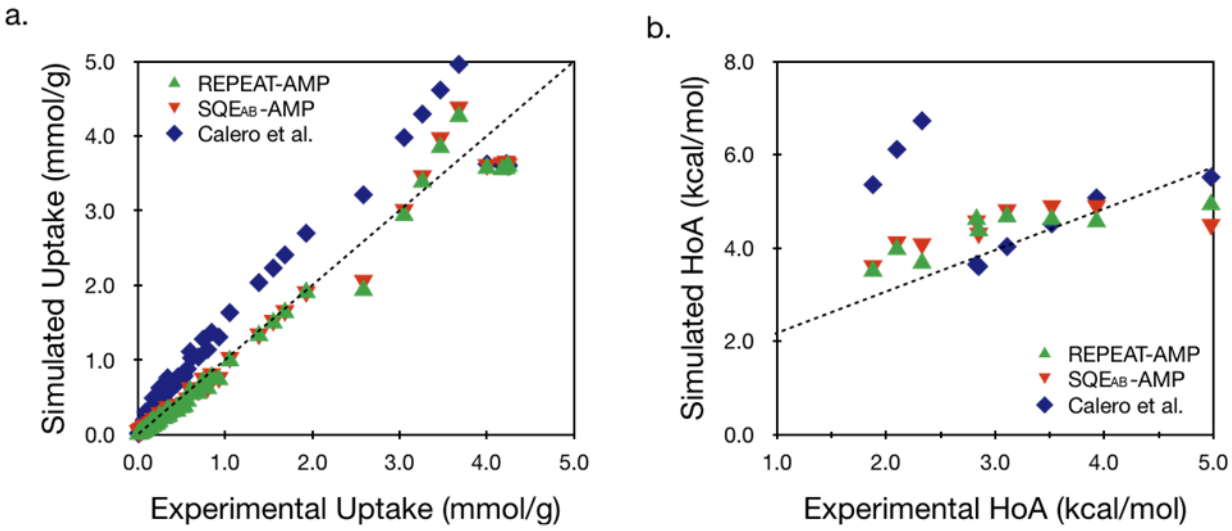

Figure 4.8: *Comparison of adsorption properties calculated with the AMP force field using either DFT derived REPEAT charges or optimized SQE_{AB} charges, and the Calero et al. force field published for N_2 and other small molecule gases between 0.1 – 1.0 bar and 77K – 373 K for (a) N_2 uptake and (b) heats of adsorption. The dotted black line represents the ideal correspondence between the simulated and experimental results.*

Table 4.7: *Average PEs, Average relative PEs, Pearson and Spearman Rank Order Coefficients for N_2 Uptake and HoA between 0.1 – 1.0 bar and 77K – 373 K, computed using the Calero et al., REPEAT-AMP and SQE_{AB}-AMP force fields.*

a. N_2 Uptake Analysis

model	$\overline{PE}$ (%)	$\overline{PE_r}$ (%)	Pearson (R^2)	Spearman (R^2)	Standard Deviation (±%)
Calero *et al.*	34.08	+31.88	0.976	0.943	37.97
SQE_{AB}-AMP	22.38	-14.76	0.993	0.989	4.86
REPEAT-AMP	15.27	-13.79	0.996	0.996	2.65

b. N_2 HoA Analysis

model	$\overline{PE}$ (%)	$\overline{PE_r}$ (%)	Pearson (R^2)	Spearman (R^2)	Standard Deviation (±%)
Calero *et al.*	35.30	+22.72	0.515	0.588	34.09
SQE_{AB}-AMP	10.21	+6.64	0.818	0.828	8.60
REPEAT-AMP	3.21	+3.2	0.954	0.987	2.95

Using the REPEAT-AMP and SQE_{AB}-AMP force fields, the average PEs and associated standard deviations were found to be 15.27 ± 2.65% and 22.38 ±4.86%, respectively. Relative to the Calero *et al.* force field for N_2 adsorption, which was found to deviate from experiment by 34.08 ±

37.97%, the AMP force field with either charge method offer significant improvements in accuracy (>10% decrease in the PE for the AMP force field). The Pearson correlation (least-squares) coefficient for the REPEAT-AMP and SQE_{AB}-AMP are both ~0.98, closely followed by Calero *et al.* at ~0.97. Similar results were obtained with the Spearman rank correlation coefficients.

In Figure 4.8b, the simulated HoAs are plotted against the experimental data, given in Table 4-S8 in Appendix A. The force fields evaluated are shown to over predict the energetics of N_2 adsorption. The relative percent errors were determined as +22.72%, +6.64%, and +3.2% for Calero *et al.*, SQE_{AB}-AMP, and REPEAT-AMP force fields, respectively. Of the methods consider, the AMP force field is shown to give excellent agreement with experiment with an average PE of 10.21 ± 8.60% when using the optimized SQE_{AB} charges and 3.21 ± 2.95% when using REPEAT charges on the framework atoms. This presents a significant enhancement over the Calero *et al.* force field, where the average PE was found to be 35.30 ± 34.09%. Looking at the results of a linear regression analysis, the Pearson correlation coefficient was found to be 0.818 for the SQE_{AB}-AMP force field and 0.954 for the REPEAT-AMP force field. In comparison with the Calero *et al.* force field, where the Pearson correlation coefficient was found to be 0.515, the AMP force field is shown to provide a better model for the energetics of N_2 adsorption in aluminosilicates.

The performance of the AMP force field with respect to CO_2 adsorption in aluminosilicates is next assessed relative to experiment, and compared with the Calero *et al.* and Fang *et al.* force fields. As described in the methods section of this chapter, these two force fields differ in their optimization strategies, where Calero *et al.* fit parameters to experimental data, and Fang *et al.* utilized first principles derived interaction energies. Combining both of these procedures, the AMP force field optimization followed a similar schema to that proposed by Fang *et al.* while utilizing experimental data analogous to Calero *et al.* Presented in Figure 4.9a are the simulated CO_2 uptakes

plotted relative to the experimental data, where fairly similar behavior is observed for the four unique force fields investigated with all of the models considered demonstrating a tendency to over predict the CO_2 uptakes in aluminosilicates, with the exception of SQE_{AB}-AMP. Looking at the simulated CO_2 HoAs in Figure 4.9b, however, the AMP model is much more evidently seen reproducing experiment within a smaller range of error relative to the other force fields, were significant overestimations are observed.

In Table 4.8, the average PEs show, that the AMP force field offers a minimum improvement over the others by ~5.0% when using SQE_{AB} charges, where the average PE was found to be 25.89±16.54%. Even better results are obtained when using DFT derived REPEAT charges, where the average PE was found to be 19.92±9.90%. Though the improvement when using SQE_{AB} charges is relatively small, further consideration of the quantitative analysis summarized in Table 4.8a shows that the AMP force field is generally more reliable than the others. For example, the max PEs found for both Calero *et al.* and Fang *et al.* are more than 20% larger than for the SQE_{AB}-AMP force field, and more than double the REPEAT-AMP force field. This is also reflected in the smaller standard deviations observed for the AMP for the force field relative to the other models. In this respect it should be noted that the small standard deviations found with the AMP force field for both N_2 and CO_2 provide a strong indication of the robustness of the model across the framework types and conditions, which are important to post-combustion CO_2 capture, or CO_2/N_2 separations.

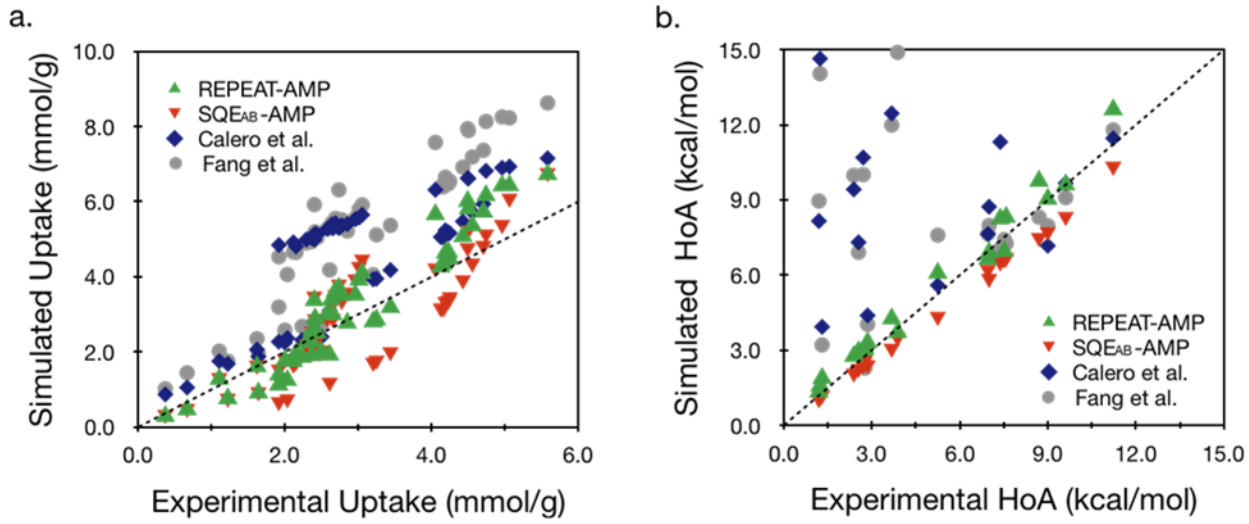

Figure 4.9: *Comparison of adsorption properties calculated with the AMP force field using either DFT derived REPEAT charges or optimized SQE_{AB} charges, and the Calero et al. and Fang et al. force fields, between 0.1–10.0 bar and 297K – 333 K for* ***(a)*** *CO_2 uptake and* ***(b)*** *heats of adsorption. The dotted black line represents the ideal correspondence between the simulated and experimental results.*

In Table 4.8, the average PEs show, that the AMP force field offers a minimum improvement over the others by ~5.0% when using SQE_{AB} charges, where the average PE was found to be 25.89±16.54%. Even better results are obtained when using DFT derived REPEAT charges, where the average PE was found to be 19.92±9.90%. Though the improvement when using SQE_{AB} charges is relatively small, further consideration of the quantitative analysis summarized in Table 4.8a shows that the AMP force field is generally more reliable than the others. For example, the max PEs found for both Calero *et al.* and Fang *et al.* are more than 20% larger than for the SQE_{AB}-AMP force field, and more than double the REPEAT-AMP force field. This is also reflected in the smaller standard deviations observed for the AMP for the force field relative to the other models. In this respect it should be noted that the small standard deviations found with the AMP force field for both N_2 and CO_2 provide a strong indication of the robustness of the model across the framework types and conditions, which are important to post-combustion CO_2 capture, or CO_2/N_2 separations.

Table 4.8: *Average PEs, Average relative PEs, Pearson and Spearman Rank Order Coefficients for* ***(a)*** *CO_2 Uptake and* ***(b)*** *HoA between 0.1 – 10.0 bar and 297K – 333 K, computed using the Calero et al., Fang et al., REPEAT-AMP and SQE_{AB}-AMP force fields.*

a. CO_2 Uptake

model	$\overline{PE}$ (%)	$\overline{PE_r}$ (%)	Pearson (R^2)	Spearman (R^2)	Standard Deviation (±%)
Calero *et al.*	31.05	+27.94	0.677	0.735	25.41
Fang *et al.*	34.49	+58.89	0.649	0.665	27.68
SQE_{AB}-AMP	25.89	-10.38	0.797	0.762	16.54
REPEAT-AMP	19.92	+2.64	0.922	0.922	9.90

b. CO_2 HoA

model	$\overline{PE}$ (%)	$\overline{PE_r}$ (%)	Pearson (R^2)	Spearman (R^2)	Standard Deviation (±%)
Calero *et al.*	64.73	+61.04	0.310	0.309	87.85
Fang *et al.*	33.28	+25.87	0.461	0.895	72.34
SQE_{AB}-AMP	14.91	-14.97	0.973	0.986	3.06
REPEAT-AMP	13.26	+10.90	0.996	0.996	9.97

When considering the heats of CO_2 adsorption, Table 4.8b shows that the AMP force field again performs better than the other models considered. The average PE for SQE_{AB}-AMP was found to be 14.91±3.06%, while for the REPAT-AMP force field we see a smaller error of 13.26±9.97%. Both of these are less than half of the errors observed when using the Fang *et al.* force field, and well under a third of the errors obtained when using the Calero *et al.* force field. The Pearson correlation coefficient for the REPEAT-AMP and SQE_{AB}-AMP are 0.996, and 0.973 respectively, while for Calero *et al.* and Fang *et al.* the correlation between the simulated and experimental HoAs were found to be 0.310 and 0.461, respectively. For the AMP model, similar results were obtained with the Spearman rank correlation coefficients. For the other force fields, a larger change between the Pearson and the Spearman Rank order coefficients is observed, generally suggesting the existence of outliers in the data, which can be seen in Figure 4.9b for both the Calero *et al.* and Fang *et al.* force fields.

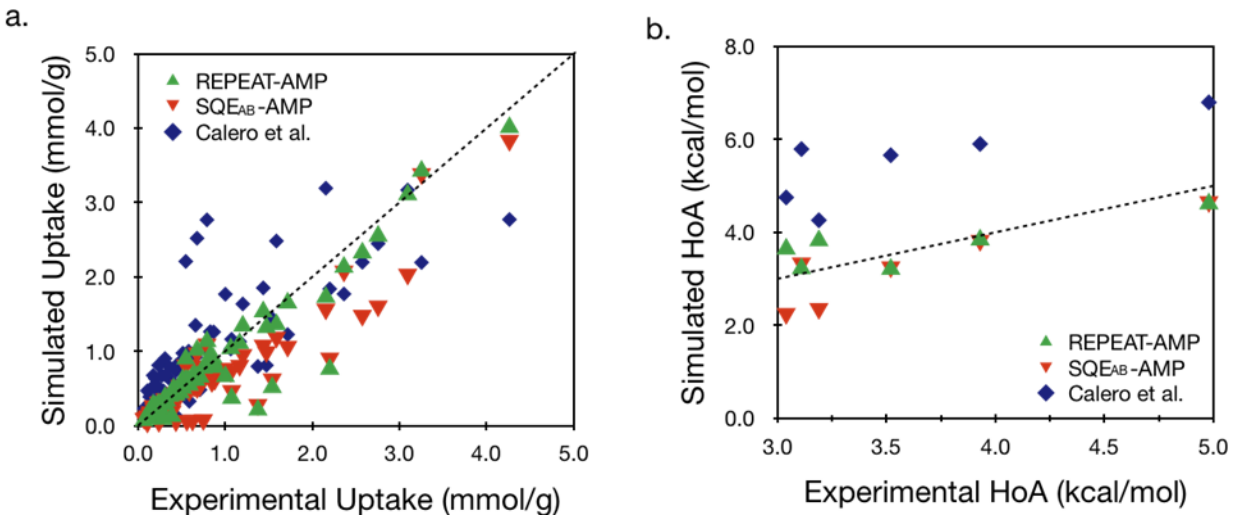

Figure 4.10: *Comparison of adsorption properties calculated with the AMP force field using either DFT derived REPEAT charges or optimized SQEAB charges, and the Calero et al. force fields, between 0.1– 10.0 bar and 297K – 333 K for* ***(a)*** *CH_4 uptake and* ***(b)*** *heats of adsorption. The dotted black line represents the ideal correspondence between the simulated and experimental results.*

For modeling methane interactions with aluminosilicate zeolite frameworks, the AMP force field is evaluated against experiment and compared with the Calero *et al.* force field for small molecule gases. In Figure 4.10a the simulated uptakes are shown, where the models appear to overestimate the uptake at lower loadings and underestimate the uptake slightly at higher loadings. Figure 4.10b shows the simulated HoAs, where a similar trend is observed as was for the uptakes.. Relative to the Calero *et al.* force field, the AMP model again achieves a greater degree of accuracy in predicting CH_4 adsorption characteristics within aluminosilicate materials. The results of the numerical analysis associated with Figure 4.10 are summarized in Table 4.9. When simulating uptakes, the average PEs for the AMP force field when using either REPEAT or SQE_{AB} charges are 17.33±16.44% and 26.50±14.11, respectively, while the average PE for the Calero *et al.* force field was found to be 40.87±33.95%. In terms of linear correlation (Pearson) or ranking (Spearman), the AMP force field is again seen to offer an increase in accuracy relative to the Calero model, with Pearson correlations of ~0.88 and ~0.81 for REPEAT and SQE_{AB} relative to ~0.59 for Calero. Similar results are obtained with the Spearman rank order coefficients.

Table 4.9: *Average PEs, Average relative PEs, Pearson and Spearman Rank Order Coefficients for* ***(a)*** *CH_4 Uptake* ***and (b)*** *HoA between 0.1 – 10.0 bar and 248K – 343 K, computed using the Calero et al., REPEAT-AMP and SQE_{AB}-AMP force fields.*

a. CH_4 Uptake

model	$\overline{PE}$ (%)	$\overline{PE_r}$ (%)	Pearson (R^2)	Spearman (R^2)	Standard Deviation (±%)
Calero *et al.*	40.87	+34.68	0.593	0.617	33.95
SQE_{AB}-AMP	26.50	-17.75	0.812	0.737	14.11
REPEAT-AMP	17.33	+0.98	0.879	0.798	16.44

b. CH_4 HoA

model	$\overline{PE}$ (%)	$\overline{PE_r}$ (%)	Pearson (R^2)	Spearman (R^2)	Standard Deviation (±%)
Calero *et al.*	62.93	+58.76	0.246	0.098	29.53
SQE_{AB}-AMP	26.45	+0.48	0.285	0.156	17.39
REPEAT-AMP	24.98	+14.75	0.600	0.588	17.19

When simulating the CH_4 HoAs, the average PEs for the AMP force field when using either REPEAT or SQE_{AB} charges were found to be 24.98±17.19% and 26.45±17.39%, in contrast with 62.93±29.53% for Calero *et al.* It is interesting to note that the SQE_{AB}-AMP and REPEAT-AMP force fields have fairly similar performance here. Looking at Table 4.8b similar performance is again observed when simulating the CO_2 HoAs around similar conditions. Due to the large quadrupole moment of CO_2, the calculated adsorption properties are strongly affected by the framework charges. Methane, on the other hand, is a non-polar molecule for which we would expect the simulated uptakes to be unaffected by the framework charges. Although in Figure 4.10b it is apparent that at lower energies REPEAT charges tend to over predict the HoA, while when using SQE_{AB} charges we observe a slight under prediction, the two charge methods generally reproduce the same results using the same set of LJ parameters. This result is provides a further validation of the optimized SQE_{AB} charges to reproduce the QM ESP as well as periodic DFT derived REPEAT charges.

Looking at the result of the linear regression, we see lower Pearson correlation coefficients for the AMP force field relative to the other guests, with 0.6 for REPEAT-AMP and ~0.29 for SQE_{AB}-AMP. The few deviations from the 1:1 line seen in Figure 4.10b for REPEAT- and SQE_{AB}-AMP, which are on the order of magnitude of ~1.0 kcal/mol in some cases, have a large effect on the calculated correlation coefficients. The same can been seen with the Calero *et al.* force field, where deviations can be on the order of ~2.0 kcal/mol, and the corresponding Pearson R^2 value is ~0.25. Analysis of the Si/Al ratios of the frameworks considered here did not reveal any relationship between the observed error. Rather, as stated, the force fields seem to deviate most at lower energies, which would more likely be related to the temperature and pressure at which adsorption occurs.

As a final test of the efficacy of the AMP force field, we apply the it towards determining the locations of the extra-framework sodium cations in the LTA-4A framework (Si/Al = 1.0), originally resolved by Pluth *et al.*[70] Although it is usually not possible to determine the Al-sitings and extra-framework cation locations from traditional XRD methods, the high symmetry of the LTA framework along with the equal number of Si and Al atoms occupying the T-sites when fully saturated with Al, make this possible when following Lowenstein's rule. In essence, there is only one possible configuration of Al atoms that allow for LTA to exist with a Si/Al ratio of 1. The single crystal X-ray structure of the LTA framework is shown in Figure 4.11a, with the extra framework Na^+ cations depicted in purple. For the unit cell depicted in Figure 4.11a, only 48 Na^+ ions are required to neutralize the unit cell, whereas 72 counter-ions locations are identified from the XRD experiment. (From the perspective shown in Figure 4.11a there are only 36 sites that are 'on top' of one another and hidden from view). In other words, there are only 48 Na^+ counter-ions on average per unit cell, which are spread across 72 sites within the crystal due to disorder in the cation

positions. For example, in the 4 central channels that are seen in Figure 4.11a, the 4 locations that are shown are the perfect archetype of dynamic or static disorder commonly observed in aluminosilicates. Using a well sampled Monte Carlo simulation of the Na^+ positions , one should pick up all of these binding sites as well as the disorder in the extra framework cation positions, given a good model of the guest-host interactions.

In Figure 4.11b, the results obtained using the Fang *et al.* force field are shown. Since this force field was optimized with the counter ions placed according to the experimental positions and occupancies,[70] the predicted positions are expected to match experimental very well, and serve to ensure that we have sufficiently sampled MC simulation. Using either REPEAT charges, fit to reproduce the QM ESP, or SQE_{AB} charges that were optimized to reproduce the QM ESP, we have run single point Monte Carlo simulations to determine the most probable positions of the Na^+ ions. A comparison between Figure 4.11c and 4.11d with 4.11a reveals that there is excellent agreement between the predicted and experimental extra-framework ion positions for the AMP force field when using either REPEAT or SQE_{AB} charges. Furthermore, we can see the disorder in the cation positions within the pore apertures being reproduced well by either model.

Although the AMP model offers significant improvements over current force fields that utilize generic charge sets optimized on neutral zeolites, the flexibility of the parameters could be further improved. As discussed in the methods section, the optimization schemed adopted from Fang *et al.* limits the flexibility of the model by reducing the number of free variables during the optimization procedure. Allowing each of the cross-species terms to vary independently of one another during the GA, should naturally result in a more robust model, as well as improve the accuracy of the model. Using the results of this work, appropriate constraints that ensure physically meaningful parameters can be rationalized.

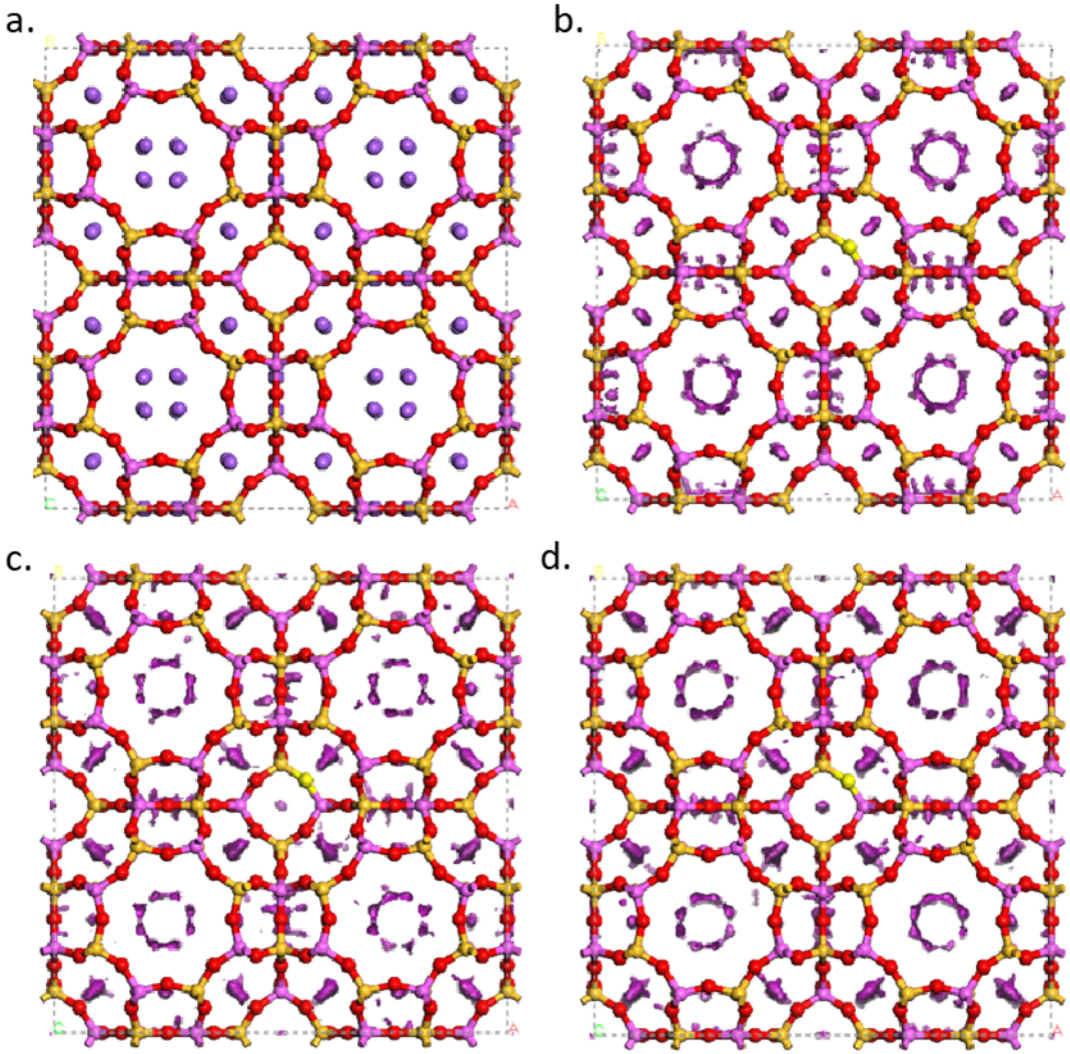

Figure 4.11 (a) *Experimental XRD crystal structure of LTA-4A (Si/Al = 1.0), with the extra-framework Na^+ ions shown in purple. Isosurface(purple) of the extra-framework Na^+ probability distribution determined from a Monte Carlo simulation using* ***(b)*** *the Fang et al. force field,* ***(c)*** *the REPEAT-AMP force field, and* ***(d)*** *the SQE_{AB}-AMP force field. Overlap of the isosurfaces on the plane of the page with those behind the page, on the opposite face of the unit cells shown, results in what appears to be 8 regions of concentrated probability in each of the 4 central pores where there should only be four (due to disorder in the cation positions).*

4.5 Conclusion

In the first part of this work, we developed a robust set of parameters for the SQE_{AB} method introduced by Krykunov *et al.*[24], termed SQE_{AB}-AMP, as an alternative to generic charge sets that are commonly used to model the electrostatic interactions in zeolite systems. These parameters were fit to reproduce the ESP derived from first principles QM calculations for each unique system

contained in the training set. A total of 111 unique framework types were taken from the IZA-SC database, from which 63 were selected to populate the training set while the remaining 48 were used to validate the model. The training and validation sets were carefully tailored to reflect the structural diversity found across all siliceous and aluminosilicate framework types identified in the IZA-SC database. The resulting SQE_{AB}-AMP charges were then shown to reproduce the ESP within the pores of the validation frameworks significantly better than the generic charge sets published by Calero *et al.* and Fang *et al.,* as well as SQE charges optimized by T. Verstraelen *et al.* to reproduce Hirchfeld I charges. Across the validation set, the MAD between the QM ESP and those resulting from the SQE_{AB}-AMP charges was approximately 50% smaller than those resulting from the other methods examined, with the exception of REPEAT charges. We then showed that by incorporating a simple atom typing scheme that distinguished between O atoms bond to Al and those bonded exclusively to Si atoms, SQE_{AB}-AMP charges accurately captured the chemical structure of the system. This was explicitly examined for a charged zeolite framework, where it was found that the SQE_{AB} charges on oxygen atoms covalently bonded to an Al atom were more polarized than those that were not – a result that was in agreement with the QM derived REPEAT charges.

The SQE_{AB}-AMP method was further validated by evaluating the CO_2 uptake from GCMC simulations using SQE_{AB}-AMP charges and compared with the results obtained when using REPEAT charges on the framework atoms and UFF parameters to model the VdW interactions. The adsorption properties determined with SQE_{AB}-AMP charges were found to be in excellent agreement with those determined using DFT derived REPEAT charges, with a Pearson and Spearman Rank correlation of ~0.99. This is a significant result since the SQE_{AB}-AMP model was trained to reproduce the DFT derived ESP, and not the adsorption properties. Furthermore, in a study carried out by Fischer and Bell, the CO_2 uptakes calculated using generic charges to model the

electrostatic interactions in siliceous zeolites were found to differ by an RMSPE of more than 10% in half the cases considered, with a maximum RMSPE of ~23%. Conversely, the RMSPE does not exceed 5% for any of the 29 siliceous and aluminosilicate frameworks that were assessed in this work. On average, the SQE_{AB}-AMP charges were shown to reproduce the simulated CO_2 adsorption isotherms obtained using REPEAT charges within an RMSPE of 1.47%, with a maximum RMSPE of 4.38%. We also noted that this result suggests that for the same LJ potential parameters, the two charge methods should yield similar predictions of the adsorption properties of a given zeolite. The CO_2 uptakes determined using SQE_{AB}-AMP charges were found to slightly over predict the CO_2 uptake at lower loadings, and slightly under estimate the CO_2 uptake at higher loadings relative to the results obtained using REPEAT charges. SQE_{AB}-AMP was shown to be a robust method for rapidly generating partial atomic charges that accurately reproduce the DFT derived ESP in siliceous and aluminosilicate zeolites that would be ideal for high-throughput screening, or in the identification of Al-sitings in zeolites through the computational method proposed by Calero *et al.*[71]

In the following section, we turned our attention to the development of a force field for modeling gas adsorption in zeolites, for which both SQE_{AB}-AMP and the REPEAT charges can be used in place of the generic charges sets that are typically used for modeling the electrostatic interactions in these systems. In total, 39 unique adsorption isotherms were used to train the model, with the data equally distributed between the N_2, CO_2, CH_4. In order to demonstrate the validity of the model parameters using either charge model, we assessed their performance in predicting three different physical observables directly related to accuracy of the simulated guest-host interactions and compared these results with other force fields found in the literature. Firstly, the simulated uptakes were compared to the experimental uptakes for the REPEAT-AMP and the SQE_{AB}-AMP force fields for each of the guests modeled, and further evaluated against the Calero *et al.* and Fang *et*

al. force fields. The results of this comparison showed that for N_2, CO_2, and CH_4 uptakes, that the REPEAT-AMP force field outperforms the other models investigated by up to 20%, while the faster SQE_{AB}-AMP model outperforms the Calero and Fang force fields by up to 15%. Similarly, the AMP force field was also shown to better describe the energetics of adsorptions, with REPEAT-AMP surpassing the other models by a maximum of ~50%, and SQE_{AB} performing at least ~36.5% better than the force fields taken from the literature. However, we note that the flexibility of the model parameters presented here is limited, and future refinement should be carried by allowing all cross-species terms to varied independently.

Since charged aluminosilicate materials are currently one of the most highly utilized porous materials in industry, having fast and accurate methods for modeling the guest-host interactions within the pores and channels of these systems is of great value. As we have shown, SQE_{AB} charges optimized to reproduce the DFT derived ESP in these systems offers a more accurate alternative to the generic charge sets, without any substantial increase in the computational demand for calculating the charges. A further benefit, as well a future application of the AMP force field, results from the its easily expandable nature. In this respect, new guest molecules or a greater diversity of extra-framework cations can easily be included into the force field, without the need to re-parameterize the SQE_{AB} charge model This is because it was optimized to reproduce the gauge-corrected ESP of charge aluminosilicates with a constant background charge to neutralize the unit cell rather than counter-ions. Therefore, the framework atomic charges are independent of the extra-framework cation charges. Thus, we envision an AMP force field in the future which is capable of modeling a wider variety of small molecule gases such as noble gases and oxygen, as well as a variety of additional counter ions, including some divalent counter-ions such as Cs^{2+}. Additionally, it may be interesting to include and combine different types of data from both experiment and theory such as

heats of adsorption and DFT derived interactions energies, which could be used to supplement lacking data in some cases.

4.6 Appendix

4.6.1 Appendix A: Supporting Information

Table 4-S10: *Training Set for SQEAB Optimization. All unique structures are listed with their Framework Type Codes (FTCs) and the corresponding Si/Al ratio(s) that were used.*

FTC	Si/Al	FTC	Si/Al	FTC	Si/Al
ABW	1.0, 3.0	**EPI**	1.0, 3.0	**MRE**	1.0, 3.0, ∞
AFG	1.0, 3.0	**ERI**	1.0, 3.0	**MSE**	1.0, 3.0, ∞
ANA	1.0, 3.0	**ESV**	1.0	**MSO**	1.0, 3.0, ∞
BEA	1.0, 3.0	**EUO**	3.0, ∞	**MTF**	1.0, 3.0, ∞
BIK	∞	**FAU**	1.0	**MTT**	1.0, 3.0, ∞
BRE	1.0, 3.0	**FER**	1.0, 3.0	**MTW**	1.0, 3.0
BOG	1.0	**GME**	1.0, 3.0	**NAT**	1.0, 3.0
CAN	1.0, ∞	**GOO**	1.0, 3.0	**NES**	1.0, 3.0
CAS	1.0, 3.0	**HEU**	1.0	**NON**	1.0, 3.0, ∞
CDO	1.0, 3.0, ∞	**IFR**	1.0, 3.0, ∞	**PAU**	3.0
CFI	1.0, 3.0, ∞	**ITE**	1.0, 3.0, ∞	**RHO**	1.0, 3.0
CHA	1.0, 3.0	**ITH**	1.0, 3.0	**RTE**	1.0, ∞
CSV	1.0, 3.0, ∞	**ITW**	1.0, 3.0, ∞	**STI**	1.0, 3.0
DAC	1.0, 3.0	**JBW**	1.0, 3.0	**STO**	1.0, 3.0, ∞
DDR	1.0, 3.0, ∞	**LTA**	1.0, 3.0	**SFW**	1.0
DOH	1.0, 3.0, ∞	**LTL**	1.0, 3.0	**SOD**	1.0
DON	1.0, 3.0	**LTN**	1.0	**TOL**	1.0, 3.0
EAB	1.0 3.0, ∞	**MAZ**	1.0, 3.0	**TON**	1.0
EDI	1.0, 3.0	**MEI**	1.0, 3.0, ∞	**TUN**	1.0, 3.0, ∞
EMT	1.0, 3.0	**MEL**	1.0, 3.0, ∞	**UFI**	1.0, 3.0, ∞
EON	1.0, 3.0	**MFI**	1.0, 3.0, ∞	**VET**	1.0, 3.0, ∞

Table 4-S11: *Validation Set for SQEAB Optimization. All unique structures are listed with their Framework Type Codes (FTCs) and the corresponding Si/Al ratio(s) that were used.*

FTC	Si/Al	FTC	Si/Al
DON	∞	**MFS**	1.0, 3.0
EEI	∞	**MON**	1.0, 3.0
FAR	1.0, 3.0	**MOR**	1.0, 3.0
FRA	1.0, 3.0	**MOZ**	1.0, 3.0
GIS	1.0, 3.0	**MTN**	1.0, 3.0
GIU	1.0, 3.0	**NSI**	1.0, 3.0, ∞
GIU	1.0, 3.0	**OFF**	1.0
GON	1.0, 3.0, ∞	**OKO**	∞
IFY	∞	**PCR**	∞
IHW	∞	**PHI**	1.0, 3.0
IMF	1.0, 3.0, ∞	**RRO**	1.0, ∞
ISV	1.0, 3.0, ∞	**RWR**	1.0, 3.0, ∞
ITH	∞	**SFE**	1.0, ∞
IWV	1.0, 3.0	**SFF**	1.0, 3.0, ∞
KFI	1.0, 3.0	**SFV**	∞
LAU	1.0, 3.0	**SGT**	1.0, 3.0, ∞
LEV	1.0	**STF**	∞
LIO	1.0, 3.0	**STT**	1.0, ∞
LOS	1.0, 3.0	**SZR**	1.0, 3.0
LTF	∞	**TER**	1.0, 3.0
LTJ	1.0, 3.0	**THO**	1.0
MAR	1.0, 3.0	**TSC**	1.0, 3..0
MEP	1.0	**UOV**	1.0, 3.0, ∞

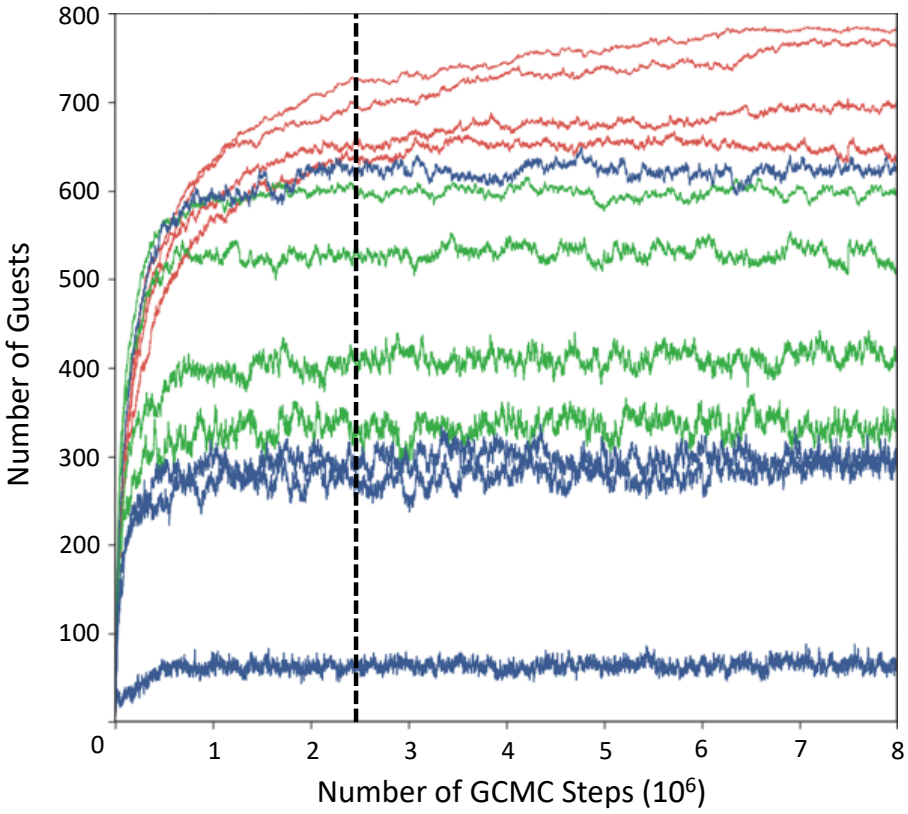

Figure 4-S2: *Justification of the Number of Steps used for GCMC Simulations Carried out During LJ Parameter optimization. Shown here is the number of guests as a function of the number of GCMC steps for a select group of isotherms that are included in the training set. A total of 4 GCMC simulations are shown for each guest: CO_2 (red), CH_4(green), and N_2 (blue). Only 4 of the 13 isotherms included in the training set for each guest are shown in the above figure to maintain clarity. The dashed line at 2.5 million steps indicates the end of the equilibration phase and the start of the production phase of the GCMC simulation. At this point the number of guests is fairly stable. At 7.5 million steps the number of guests begins to converge, which is reflected in the graph as the number of guests in all cases begins to level off, and fluctuate consistently about a mean value. Thus, for all simulations carried out during the LJ parameter optimization, 2.5 million equilibration steps, and 5 million production steps were used in accordance with the above result.*

Table 4-S12: *Optimized QEq Parameters without Atom-typing.*

atom	χ_i (eV)	κ_i (eV)
O	41.431	28.522
Si	0.39622	18.451
Al	0.21122	31.818

Table 4-S13: *Optimized SQE_{AB}-AMP Parameters without Atom-Typing*

atom	χ_i (eV)	κ_i (eV)	bond	χ_{ij} (eV)	κ_{ij} (eV)
O	50.0	45.029	O-Al	-2.3155	0.0243442
Si	0.24458	48.838	O-Si	-4.1041	0.135120
Al	0.02016	26.558			

Table 4-S14: *Optimized QEq-AMP Parameters with Atom-Typing*

atom	χ_i (eV)	κ_i (eV)
O	49.265	45.34
O_{Al}	48.384	30.952
Si	0.014133	31.406
Al	0.016617	32.133

Table 4-S15: *Force field parameters used in this work to describe guest interactions. The dispersion interactions are represented through a Lennard-Jones potential.*

Guest		Calero *et al.*	AMP	Fang *et al.*	References
Na^+	**q/e**	q_{Na} = +0.3834	q_{Na} = +1.00	q_{Na} = +0.99	Calero *et al.*[20]
	ε/K	ε_{Na^+} = 15.1	ε_{Na^+}= ___	ε_{Na^+} = 149.0	This work
	σ/Å	σ_{Na^+} = 2.658	σ_{Na^+}= ___	σ_{Na^+} = 2.37	Fang *et al.*[21] Grimme[42]
CO_2	**q/e**	q_C = +0.6512 q_O = -0.3256	q_C = +0.6512 q_O = -0.3256	q_C = +0.6512 q_O = -0.3256	
	ε/K	ε_C = 29.93 ε_O = 85.66	ε_C = 29.93 ε_O = 85.66	ε_C = 29.93 ε_O = 85.66	
	σ/Å	σ_C= 2.80 σ_O= 3.05	σ_C= 2.80 σ_O= 3.05	σ_C= 2.80 σ_O= 3.05	
CH_4	**q/e**	q_{COM} = +0.3834	q_C = -0.484 q_H = +0.116		TraPPE Siepmann *et al.*[44]
	ε/K	ε_{COM} = 148.00	ε_C = 52.84 ε_H = 7.66		
	σ/Å	σ_{COM}= 3.73	σ_C= 3.43 σ_H= 2.846		5-site Kollman *et al.*[72]
N_2	**q/e**	q_N = -0.482 q_{COM} = +0.964	q_N = -0.482 q_{COM} = +0.964		TraPPE Siepmann *et al.*[73]
	ε/K	ε_N = 36.0	ε_N = 36.0		
	σ/Å	σ_N = 3.31	σ_N = 3.31		

Table 4-S16: *Lennard-Jones Parameter Optimization Reference Gas Uptake Data.*

Guest	Temp. (K)	Isotherms	FTC (Si/Al)	References
CO_2	297 – 333	13	LTA (1.0) MFI (25, 30) FAU (NaX, NaY)	○ L. Hauchhum *et al.*, *Int. J. Energy Environ. Eng.,* **2014**, (5), 349. ○ J. Dunne *et al.*, *Langmuir*, **1996**, *12* (24), 5896. ○ V. B. Mortola *et al.*, *Microporous and Mesoporous Mater.,* **2016**, 222, 209. ○ W. Shao *et al.*, *Adsorption,* **2009**, 15, 497. ○ Cavenati *et al.*, *J. Chem. Eng. Data.,* **2004**, 49 (4), 1095.
CH_4	248 – 343	13	LTA (1.0, 2.0, 5.0) MFI (30) FAU (NaX, NaY)	○ M. Palomino *et al.*, *Langmuir*, **2009**, 26 (3), 1910. ○ C. A. Grande *et al.*, *Energy Fuels*, **2014**, 28 (10), 6688. ○ J. Dunne *et al.*, *Langmuir*, **1996**, *12* (24), 5896. ○ M. Mofarahi *et al.*, *J. Chem. Eng. Data*, **2015**, 60 (3), 683. ○ O. Talu *et al.*, *J. Phys. Chem.*, **1993**, 97 (49), 12894.
N_2	77 – 373	13	LTA (2.0) MFI (30) FAU (NaX, NaY)	○ J. C. Groen *et al.*, *J. Am. Chem. Soc.*, 129, 355. ○ Cavenati *et al.*, *J. Chem. Eng. Data.,* **2004**, 49 (4), 1095. ○ M. Mofarahi *et al.*, *J. Chem. Eng. Data*, **2015**, 60 (3), 683. ○ X. Gongkul *et al.*, *Adsorption*, 23 (1), 131. ○ J. Dunne *et al.*, *Langmuir*, **1996**, *12* (24), 5888. ○ W. Shao *et al.*, *Adsorption,* **2009**, 15, 497. ○ R. J. Harper *et al.*, *Canadian Journal of Chemistry*, **1969**, 47 (24), 4661.

Table 4-S17: *HoA Data used in LJ Parameter Validation.*

Guest	Temp. (K)	Isotherms	FTC (Si/Al)	References
CO_2	303 - 333	4	LTA (1.0) MFI (30) FAU (NaX, NaY)	○ J. Dunne *et al.*, *Langmuir*, **1996**, *12* (24), 5896. ○ W. Shao *et al.*, *Adsorption,* **2009**, 15, 497. ○ M. Palomino *et al.*, *Langmuir*, **2009**, 26 (3), 1910. ○ Zukal *et al.*, *Topics in Catalysis*, **2010**, 53(19), 1361.
CH_4	293	4	LTA (1.0, 2.0, 3.4, 5.0)	○ M. Palomino *et al.*, *Langmuir*, **2009**, 26 (3), 1910.
N_2	303 - 306	2	FAU (NaX, NaY)	○ W. Shao *et al.*, *Adsorption,* **2009**, 15, 497.

Table 4-S18: *Resulting Optimized Scaling Factors used to calculate the cross-species LJ parameters using Grimme's semi-empirical expression for dispersion interactions.*

model		**CH_4**	**CO_2**	**N_2**
Neutral (Si, O)	s_6	0.7417	0.8517	0.8416
	s_{12}	3.7281	4.534	4.516
Charged (Al, Na^+)	s_6	0.3969	0.2819	0.2567
	s_{12}	1.2273	1.2452	1.2273

4.6.2 Appendix B: Genetic Algorithm used for SQE_{AB} and LJ Parameter Optimization

A custom GA, developed by Sean Collins of the Woo research group, was used to fit all of the parameters simultaneously for each method separately (QEq and SQE_{AB}). The GA is initialized with the creation of multiple sets of randomly generated parameters, collectively known as a generation, that were then evaluated for how closely they reproduced the QM ESP. The new generation was formed by using a roulette wheel selection algorithm, which chooses two individuals from the generation to act as parents to new individuals by a mating algorithm. The mating algorithm selects a random value for each parameter which is between the values of each corresponding parameter for both the parents'. Subsequent mutations were allowed that would alter a given parameter by ± 30% of the parameter value. The GA was considered converged when the top performer in subsequent generations remained the same for ten generations.

The same underlying GA was used in the optimization of LJ parameters, with some minor alterations. Firstly, the fitness of the parameters was evaluated in terms of how closely they were able to reproduce experimental gas uptake data. The GA is again initialized with the creation of multiple sets of randomly generated parameters, however, in this case Latin-Hypercube Sampling (LHS) was implemented over normal random sampling. Due to the increased computational expense of this parameter optimization, LHS was implemented principally to ensure a more uniform sampling of the search space after the first generation. In this way, the algorithm would be more efficient by reducing the total number of generations required for the GA to converge. After evaluating the fitness of the chromosomes in the first generation, the GA then follows

the same routine as described previously. The only other modification made to the GA is within the submit script, which was altered to work with GCMC.

4.7 References

(1) Pires, J.; de Carvalho, M. B.; Ribeiro, F. R.; Derouane, E. G. Carbon Dioxide in Y and ZSM-20 Zeolites: Adsorption and Infrared Studies. *J. Mol. Catal.* **1993**, *85* (3), 295–303.

(2) Zhang, J.; Singh, R.; Webley, P. A. Alkali and Alkaline-Earth Cation Exchanged Chabazite Zeolites for Adsorption Based CO 2 Capture. **2007**.

(3) Hudson, M. R.; Queen, W. L.; Mason, J. A.; Fickel, D. W.; Lobo, R. F.; Brown, C. M. Unconventional, Highly Selective CO $_2$ Adsorption in Zeolite SSZ-13. *J. Am. Chem. Soc.* **2012**, *134* (4), 1970–1973.

(4) Bae, T.-H.; Hudson, M. R.; Mason, J. A.; Queen, W. L.; Dutton, J. J.; Sumida, K.; Micklash, K. J.; Kaye, S. S.; Brown, C. M.; Long, J. R. Evaluation of Cation-Exchanged Zeolite Adsorbents for Post-Combustion Carbon Dioxide Capture. *Energy Environ. Sci.* **2013**, *6* (1), 128–138.

(5) Foster, M. D.; Friedrichs, O. D.; Bell, R. G.; Paz, F. A. A.; Klinowski, J. Chemical Evaluation of Hypothetical Uninodal Zeolites. *J. Am. Chem. Soc.* **2004**, *126* (31), 9769–9775.

(6) Simperler, A.; Foster, M. D.; Friedrichs, O. D.; Bell, R. G.; Almeida Paz, F. A.; Klinowski, J. Hypothetical Binodal Zeolitic Frameworks. *Acta Cryst* **2005**, *61*, 263–279.

(7) Kim, J.; Lin, L.-C.; Swisher, J. A.; Haranczyk, M.; Smit, B. Predicting Large CO $_2$ Adsorption in Aluminosilicate Zeolites for Postcombustion Carbon Dioxide Capture. *J. Am. Chem. Soc.* **2012**, *134* (46), 18940–18943.

(8) Lin, L.; Berger, A. H.; Martin, R. L.; Kim, J.; Swisher, J. A.; Jariwala, K.; Rycroft, C. H.; Bhown, A. S.; Deem, M. W.; Haranczyk, M. In Silico Screening of Carbon-Capture Materials. *Nat. Mater.* **2012**, *11* (7), 633–641.

(9) Simon, C. M.; Kim, J.; Gomez-Gualdron, D. A.; Camp, J. S.; Chung, Y. G.; Martin, R. L.; Mercado, R.; Deem, M. W.; Gunter, D.; Haranczyk, M.; Sholl, D. S.; Snurr, R. Q.; Smit, B. The Materials Genome in Action: Identifying the Performance Limits for Methane Storage. *Energy Environ. Sci.* **2015**, *8* (4), 1190–1199.

(10) Campañá, C.; Mussard, B.; Woo, T. K. Electrostatic Potential Derived Atomic Charges for Periodic Systems Using a Modified Error Functional. *J. Chem. Theory Comput.* **2009**, *5* (10), 2866–2878.

(11) Ray, K. G.; Olmsted, D. L.; Burton, J. M. R.; Houndonougbo, Y.; Laird, B. B.; Asta, M. Gas Membrane Selectivity Enabled by Zeolitic Imidazolate Framework Electrostatics. *Chem. Mater.* **2014**, *26* (13), 3976–3985.

(12) Sutrisno, A.; Terskikh, V. V; Shi, Q.; Song, Z.; Dong, J.; Ding, S. Y.; Wang, W.; Provost, B.

R.; Daff, T. D.; Woo, T. K.; Huang, Y. Characterization of Zn-Containing Metal-Organic Frameworks by Solid-State 67 Zn NMR Spectroscopy and Computational Modeling. *Chem. Eur. J.* **2012**, *18* (39), 12251–12259.

(13) Morris, W.; He, N.; Ray, K. G.; Klonowski, P.; Furukawa, H.; Daniels, I. N.; Houndonougbo, Y. A.; Asta, M.; Yaghi, O. M.; Laird, B. B. A Combined Experimental-Computational Study on the Effect of Topology on Carbon Dioxide Adsorption in Zeolitic Imidazolate Frameworks. *J. Phys. Chem. C* **2012**, *116* (45), 24084–24090.

(14) Burtch, N. C.; Jasuja, H.; Dubbeldam, D.; Walton, K. S. Molecular-Level Insight into Unusual Low Pressure CO $_2$ Affinity in Pillared Metal–Organic Frameworks. *J. Am. Chem. Soc.* **2013**, *135* (19), 7172–7180.

(15) Vaidhyanathan, R.; Iremonger, S. S.; Shimizu, G. K. H.; Boyd, P. G.; Alavi, S.; Woo, T. K. Direct Observation and Quantification of CO2 Binding Within an Amine-Functionalized Nanoporous Solid. *Science (80-.).* **2010**, *330* (6004), 650–653.

(16) Taylor, J. M.; Vaidhyanathan, R.; Iremonger, S. S.; Shimizu, G. K. H. Enhancing Water Stability of Metal−Organic Frameworks via Phosphonate Monoester Linkers.

(17) Hirotani, A.; Mizukami, K.; Miura, R.; Takaba, H.; Miya, T.; Fahmi, A.; Stirling, A.; Kubo, M.; Miyamoto, A. Grand Canonical Monte Carlo Simulation of the Adsorption of CO2 on Silicalite and NaZSM-5. *Applied Surface Science*. 1997, pp 81–84.

(18) Faux, D. A.; Smith, W.; Forester, T. R. Molecular Dynamics Studies of Hydrated and Dehydrated Na + -Zeolite-4A. **1997**, *5647* (96), 1762–1768.

(19) Jaramillo, E.; Auerbach, S. M. New Force Field for Na Cations in Faujasite-Type Zeolites. *J. Phys. Chem. B* **1999**, *103* (44), 9589–9594.

(20) García-Sánchez, A.; Ania, C. O.; Parra, J. B.; Dubbeldam, D.; Vlugt, T. J. H.; Krishna, R.; Calero, S. Transferable Force Field for Carbon Dioxide Adsorption in Zeolites. *J. Phys. Chem. C* **2009**, *113* (20), 8814–8820.

(21) Fang, H.; Kamakoti, P.; Ravikovitch, P. I.; Aronson, M.; Paur, C.; Sholl, D. S. First Principles Derived, Transferable Force Fields for CO2 Adsorption in Na-Exchanged Cationic Zeolites. *Phys. Chem. Chem. Phys.* **2013**, *15* (31), 12882.

(22) Fischer, M.; Bell, R. G. Modelling CO 2 Adsorption in Zeolites Using DFT-Derived Charges : Comparing System-Specific and Generic Models.

(23) Sun, L.; Yang, L.; Zhang, Y.-D.; Shi, Q.; Lu, R.-F.; Deng, W.-Q. Accurate van Der Waals Force Field for Gas Adsorption in Porous Materials. *J. Comput. Chem.* **2017**, *38* (23), 1991–1999.

(24) Krykunov, M.; Demone, C.; Lo, J. W.-H.; Woo, T. K. A New Split Charge Equilibration Model and REPEAT Electrostatic Potential Fitted Charges for Periodic Frameworks with a Net Charge. *J. Chem. Theory Comput.* **2017**, *13* (6), 2858–2869.

(25) Degnan, T. F. *Recent Progress in the Development of Zeolitic Catalysts for the Petroleum Refining and*

Petrochemical Manufacturing Industries; Elsevier B.V., 2007; Vol. 170.

(26) Rappe, A. K.; Goddard, W. A. Charge Equilibration for Molecular-Dynamics Simulations. *J. Phys. Chem.* **1991**, *95* (8), 3358–3363.

(27) Müser, M. H.; Mosey, N.; Mosey, N. J. A Generalization of the Charge Equilibration Method for Nonmetallic Materials. **2006**, *125*.

(28) Smirnov, K. S.; van de Graaf, B. Study of Methane Adsorption in {MFI} and {MEL} Zeolites by Combination of the Electronegativity Equalization Method and Molecular Dynamics. *J. Chem. Soc. Faraday Trans.* **1996**, *92* (13), 2475–2480.

(29) Verstraelen, T.; Sukhomlinov, S. V.; Van Speybroeck, V.; Waroquier, M.; Smirnov, K. S. Computation of Charge Distribution and Electrostatic Potential in Silicates with the Use of Chemical Potential Equalization Models. *J. Phys. Chem. C* **2012**, *116* (1), 490–504.

(30) Calero, S.; Dubbeldam, D.; Krishna, R.; Smit, B.; Vlugt, T. J. H.; Denayer, J. F. M.; Martens, J. A.; Maesen, T. L. M. Understanding the Role of Sodium during Adsorption: A Force Field for Alkanes in Sodium-Exchanged Faujasites. *J. Am. Chem. Soc.* **2004**, *126* (36), 11377–11386.

(31) García-Pérez, E.; Parra, J. B.; Ania, C. O.; García-Sánchez, a.; van Baten, J. M.; Krishna, R.; Dubbeldam, D.; Calero, S. A Computational Study of CO2, N2, and CH4 Adsorption in Zeolites. *Adsorption* **2007**, *13* (5–6), 469–476.

(32) Martin-Calvo, A.; Gutiérrez-Sevillano, J. J.; Parra, J. B.; Ania, C. O.; Calero, S. Transferable Force Fields for Adsorption of Small Gases in Zeolites. *Phys. Chem. Chem. Phys.* **2015**, *17* (37), 24048–24055.

(33) Beerdsen, E.; Dubbeldam, D.; Smit, B.; Vlugt, T. J. H.; Calero, S. Simulating the Effect of Nonframework Cations on the Adsorption of Alkanes in MFI-Type Zeolites. *J. Phys. Chem. B* **2003**, *107* (44), 12088–12096.

(34) Dubbeldam, D.; Krishna, R.; Calero, S.; Yazaydın, A. Ö. Computer-Assisted Screening of Ordered Crystalline Nanoporous Adsorbents for Separation of Alkane Isomers. *Angew. Chem. Int. Ed. Engl.* **2012**, *51* (47), 11867–11871.

(35) (IZA-SC), S. C. of the I. Z. A. Database of Zeolite Structures http://www.iza-structure.org/databases/ (accessed May 13, 2016).

(36) Jarman, R. H. Application of Powder X-Ray Diffraction Data to the Determination of Framework Composition in Zeolites. *Zeolites* **1985**, *5* (4), 213–216.

(37) B.Nagy, J.; Derouane, E. G. NMR Spectroscopy and Zeolite Chemistry; 1988; pp 2–32.

(38) Löwenstein, W. The Distribution of Aluminum in the Tetrahedra of Silicates and Aluminates. *Am. Mineral.* **1954**, *39*, 92–96.

(39) Collins, S. P.; Woo, T. K. Split-Charge Equilibration Parameters for Generating Rapid Partial Atomic Charges in Metal–Organic Frameworks and Porous Polymer Networks for High-Throughput Screening. *J. Phys. Chem. C* **2017**, *121* (1), 903–910.

(40) Wolf, D.; Keblinski, P.; Phillpot, S. R.; Eggebrecht, J. Exact Method for the Simulation of Coulombic Systems by Spherically Truncated, Pairwise R-1summation. *J. Chem. Phys.* **1999**, *110* (17), 8254–8282.

(41) Ewald, P. P. Die Berechnung Optischer Und Elektrostatischer Gitterpotentiale. *Ann. Phys.* **1921**, *369* (3), 253–287.

(42) Grimme, S. Semiempirical GGA-Type Density Functional Constructed with a Long-Range Dispersion Correction. *J. Comput. Chem.* **2006**, *27* (15), 1787–1799.

(43) Harris, J. G.; Yung, K. H. Carbon Dioxide's Liquid-Vapor Coexistence Curve And Critical Properties as Predicted by a Simple Molecular Model. *J. Phys. Chem.* **1995**, *99* (31), 12021–12024.

(44) Martin, M. G.; Siepmann, J. I. Transferable Potentials for Phase Equilibria. 1. United-Atom Description of N -Alkanes. *J. Phys. Chem. B* **1998**, *102* (97), 2569–2577.

(45) Sui, Y.; Spellmeyer, D.; Pearlman, D. A.; Kollman, P.; Spellmeyer, D.; Pearlman, D. A. Simulation of the Solvation Free Energies for Methane, Ethane, and Propane and Corresponding Amino Acid Dipeptides: A Critical Test of the "Bond-PMF" Correction, a New Set of Hydrocarbon Parameters, and the Gas Phase-Water Hydrophobicity Scale. *J. Am. Chem. Soc.* **1992**, *114* (17), 6798–6801.

(46) Potoff, J. J.; Siepmann, J. I. Vapor–liquid Equilibria of Mixtures Containing Alkanes, Carbon Dioxide, and Nitrogen. *AIChE J.* **2001**, *47* (7), 1676–1682.

(47) Martin-Calvo, A.; Gutiérrez-Sevillano, J. J.; Parra, J. B.; Ania, C. O.; Calero, S. Transferable Force Fields for Adsorption of Small Gases in Zeolites. *Phys. Chem. Chem. Phys.* **2015**, *17* (37), 24048–24055.

(48) Do, D. D.; Do, H. D. Evaluation of 1-Site and 5-Site Models of Methane on Its Adsorption on Graphite and in Graphitic Slit Pores. *J. Phys. Chem. B* **2005**, *109* (41), 19288–19295.

(49) Graham, C.; Imrie, D. A.; Raab, R. E. Measurement of the Electric Quadrupole Moments of CO2, CO, N2, Cl2 and BF3. *Mol. Phys.* **1998**, *93* (1), 49–56.

(50) Wang, M.; Lawal, A.; Stephenson, P.; Sidders, J.; Ramshaw, C. Post-Combustion CO2 Capture with Chemical Absorption: A State-of-the-Art Review. *Chem. Eng. Res. Des.* **2011**, *89* (9), 1609–1624.

(51) Mason, J. A.; Sumida, K.; Herm, Z. R.; Krishna, R.; Long, J. R. Evaluating Metal–organic Frameworks for Post-Combustion Carbon Dioxide Capture via Temperature Swing Adsorption. *Energy Environ. Sci.* **2011**, *4* (8), 3030.

(52) Wilmer, C. E.; Farha, O. K.; Bae, Y.-S.; Hupp, J. T.; Snurr, R. Q. Structure–property Relationships of Porous Materials for Carbon Dioxide Separation and Capture. *Energy Environ. Sci.* **2012**, *5* (12), 9849.

(53) Sircar, S. Separation of Methane and Carbon Dioxide Gas Mixtures by Pressure Swing Adsorption. *Sep. Sci. Technol.* **1988**, *23* (6–7), 519–529.

(54) Kapoor, A. Kinetic Separation of Methane-Carbon Dioxide Mixture by Adsorption on Molecular Sieve Carbon. *Chem. Eng. Sci.* **1989**, *44* (8), 1723–1733.

(55) Bae, Y. S.; Snurr, R. Q. Development and Evaluation of Porous Materials for Carbon Dioxide Separation and Capture. *Angewandte Chemie - International Edition*. December 2, 2011, pp 11586–11596.

(56) E. García-Pérez, †; D. Dubbeldam, ‡; T. L. M. Maesen, § and; S. Calero*, †. Influence of Cation Na/Ca Ratio on Adsorption in LTA 5A: A Systematic Molecular Simulation Study of Alkane Chain Length. **2006**.

(57) Fang, H.; Kamakoti, P.; Zang, J.; Cundy, S.; Paur, C.; Ravikovitch, P. I.; Sholl, D. S. Prediction of CO 2 Adsorption Properties in Zeolites Using Force Fields Derived from Periodic Dispersion-Corrected DFT Calculations. **2012**.

(58) Kresse, G.; Furthmüller, J. Efficient Iterative Schemes for *Ab Initio* Total-Energy Calculations Using a Plane-Wave Basis Set. *Phys. Rev. B* **1996**, *54* (16), 11169–11186.

(59) Blöchl, P. E. Projector Augmented-Wave Method. *Phys. Rev. B* **1994**, *50* (24), 17953–17979.

(60) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865–3868.

(61) Duren, T.; Bae, Y. S.; Snurr, R. Q. Using Molecular Simulation to Characterise Metal-Organic Frameworks for Adsorption Applications. *Chem. Soc. Rev.* **2009**, *38* (5), 1237–1247.

(62) Smith, W.; Forester, T. R. DL_POLY_2.0: A General-Purpose Parallel Molecular Dynamics Simulation Package. *J. Mol. Graph.* **1996**, *14* (3), 136–141.

(63) Vaidhyanathan, R.; Iremonger, S. S.; Shimizu, G. K. H.; Boyd, P. G.; Alavi, S.; Woo, T. K. Direct Observation and Quantification of CO2 Binding Within an Amine-Functionalize Nanoporous Solid. *Science (80-.).* **2010**, *330*, 650–653.

(64) Vaidhyanathan, R.; Iremonger, S. S.; Shimizu, G. K. H.; Boyd, P. G.; Alavi, S.; Woo, T. K. Competition and Cooperativity in Carbon Dioxide Sorption by Amine-Functionalized Metal-Organic Frameworks. *Angew. Chemie - Int. Ed.* **2012**, *51* (8), 1826–1829.

(65) Nandi, S.; De Luna, P.; Daff, T. D.; Rother, J.; Liu, M.; Buchanan, W.; Hawari, A. I.; Woo, T. K.; Vaidhyanathan, R. A Single-Ligand Ultra-Microporous MOF for Precombustion CO2 Capture and Hydrogen Purification. *Sci. Adv.* **2015**, *1* (11), e1500421.

(66) Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. UFF, a Full Periodic Table Force Field for Molecular Mechanics and Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **1992**, *114* (25), 10024–10035.

(67) O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An Open Chemical Toolbox. *J. Cheminform.* **2011**, *3* (10), 33.

(68) Verstraelen, T.; Bultinck, P.; Van Speybroeck, V.; Ayers, P. W.; Van Neck, D.; Waroquier, M. The Significance of Parameters in Charge Equilibration Models. *J. Chem. Theory Comput.* **2011**, *7* (6), 1750–1764.

(69) Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. Uff, a Full Periodic-Table Force-Field for Molecular Mechanics and Molecular-Dynamics Simulations. *J. Am. Chem. Soc.* **1992**, *114* (25), 10024–10035.

(70) Pluth, J. J.; Smith, J. V. Accurate Redetermination of Crystal Structure of Dehydrated Zeolite A. Absence of near Zero Coordination of Sodium. Refinement of Silicon,aluminum-Ordered Superstructure. *J. Am. Chem. Soc.* **1980**, *102* (14), 4704–4708.

(71) García-Pérez, E.; Dubbeldam, D.; Liu, B.; Smit, B.; Calero, S. A Computational Method To Characterize Framework Aluminum in Aluminosilicates. *Angew. Chemie Int. Ed.* **2007**, *46* (1–2), 276–278.

(72) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges - the Resp Model. *J. Phys. Chem.* **1993**, *97* (40), 10269–10280.

(73) Potoff, J. J.; Siepmann, J. I. Vapor–liquid Equilibria of Mixtures Containing Alkanes, Carbon Dioxide, and Nitrogen. *AIChE J.* **2001**, *47* (7), 1676–1682.

5 Conclusion

5.1 Summary

The work presented in this thesis comprises two distinct projects, each with their own respective goals. In the first part of the thesis, our goal was to address a few gaps in the literature with respect to deriving partial atomic charges for periodic frameworks with a net charge. In this context, the REPEAT method was validated for charged periodic systems, and the limitation of the original SQE method to charge neutral systems was circumvented with the introduction of the SQE_{AB} method. In the latter part of the thesis, these methods were then applied toward developing a force field for modeling gas adsorption in aluminosilicates, which have charged frameworks. In addition to having physically relevant charges, the force field was then shown to provide improved accuracy relative to previously published force fields available for modeling CO_2, CH_4, and N_2 adsorption in aluminosilicates.

5.2 Research Summary

In Chapter 3, two methods are presented for calculating partial atomic charges in charged periodic systems. The first of these methods, REPEAT, was initially introduced in 2009 by Campañá *et al.*[1] as the first method to extract meaningful atomic charges directly from periodic DFT calculations of the QM ESP in charge neutral systems. In the work presented, the REPEAT method is shown to be valid in generating partial atomic charges in charged periodic systems. Traditionally, charged periodic systems have been treated by placing counter-ions within the simulation cell to neutralize the framework charge, and prevent the energy of the periodically repeated system from diverging. In the context of MOFs and zeolites this presents an issue since, in many cases, the extra-framework counter-ion positions are not known experimentally. Furthermore, *ad hoc* placement of ions in the pores of these materials could result in polarization of the framework charges. The latter

presents a problem since the counter-ions must be mobile during simulations, and such polarization would result in a bias of the counter-ions to the pre-determined locations. Therefore, we advocate obtaining ESP-fitted charges of charged frameworks by performing periodic QM calculations where the charge is neutralized by a constant background charge as proposed by Leslie *et al.*[2] To substantiate this, we performed a series of calculations to evaluate REPEAT charges on both charge and neutral systems.

Firstly, we determined the ESP charges of a neutral large pore MOF, IRMOF-16[3] from periodic DFT calculations of a 2x1x1 supercell which contains two pores. We then placed F^-, Na^+, Mg^{2+} ions into the center of the pores, imparting the unit cell with a net charge of -2, -1, +2, and +4e and derived REPEAT ESP charges for each of the charged systems. IRMOF-16 was specifically selected for this test due to its large pores, which provide 15 Å of separation between the ions and the nearest framework atoms so as to minimize polarization effects. To compare between the ESP charges derived from QM calculations of the neutral framework, and those derived from the same 'charged' frameworks, we evaluated the mean absolute deviations (MAD) in the framework atom charges. The maximum MAD was determined as 0.020e with a maximum deviation of 0.047e. Overall, the results indicate that the framework charges in both the neutral and 'charged' frameworks are very similar, and further graphical analysis revealed that the charges were neither systematically higher or lower than the other charges. A similar test was then performed using the MOF POST-1 developed by Seo *et al.*[4] POST-1 was selected since it has a net charge (-4 e per unite cell) in addition to having relatively large, triangular shaped channels with sides of ~13.4 Å in length. To neutralize the unit cell, a Ti^{4+}ion was inserted into the center of one of the channels. Comparing the REPEAT charges obtained with and without counter-ions placed in the unit cell, again there was excellent agreement between the framework atomic charges. In this case the MAD between the charges was determined to be 0.027 e with a maximum deviation of 0.154 e.

To demonstrate that there can be problems with the ESP fitting procedure when neutralizing the framework charge using counter-ions, we next evaluated the ESP-fitted charges in a small pore zeolite of type CAS. In this case, the pores are small enough that when an ion such as K^+ is added, it nearly fills the free space within the pores, ensuring polarization of the framework atom charges. To balance the charge of the CAS framework (Si/Al = 5) 4 K^+ ions were placed in each of the four pores present in the unit cell. Comparing the REPEAT charges of the CAS framework with and without neutralizing K^+ ions placed within the pores, the difference in the derived charges was found to be significant with a MAD of 0.25 e and a maximum deviation of 0.75 e. Further, the fitted charges on the K^+ ions were found to differ by as much as 0.18 e from an expected 1.0 e. Although this was an extreme example, the results from this test suggests that it is not always practical to use counter-ions to neutralize the unit cell. Additionally, the results of the first test suggest that even in large systems there will be similar charge bleeding, albeit to a lesser degree, which will influence the accuracy of subsequent simulations.

As a final means of validating the REPEAT method for deriving ESP-fitted charges in periodic systems with a net charge, we used REPAT charges to evaluate a physical observable. Dincă *et al.*[5] were able to determine the location of the Mn^{2+} extra-framework counter ions in a charged framework benzene-tris-tetrazolate MOF, **1**.[6] From the SCXRD experiment, 24 possible counter-ion locations were identified, with 3 Mn^{2+}counter-ions on average per unit cell spread across these sites in the crystal. Applying MC simulations, we set out to evaluate the quality of REPEAT charges in representing the ESP of the framework by assessing the ability of REPEAT charges to reproduce the experimental counter-ion positions determined by Dincă *et al.* Using REPEAT charges in the MC we were able to identify all 24 distinct binding sites that were determined experimentally. The results obtained with REPEAT charges were also compared with the results obtained when using QTAIM charges, which resulted in drastically different counter-ions locations compared to those

obtained with REPEAT charges and those determined experimentally. We therefore advocated in obtaining REPEAT charges of charged framework by performing periodic QM calculations where the charge is neutralized by a constant background charge rather than using counter-ions.

While REPEAT provides robust first-principles derived atomic charges, in some situations the costly electronic structure calculations that are required are prohibitive, and faster methods are needed. The second half of Chapter 3 focuses on the development and validation of an extension to the new split-charge equilibration model, which we call SQE_{AB}. Whereas the original SQE method can only be applied to neutral systems due to the antisymmetry property of the split-charges, the SQE_{AB} model can be applied to both charged and neutral systems. To circumvent this limitation, we followed approach similar to that used in the QTPIE method,[7] and developed the SQE_{AB} model by employing a pseudo inverse to express the split-charges in an atomic basis. Although an extra matrix diagonalization is required relative to the original SQE method, SQE_{AB} is still able to generate charges in a seconds with modest computing resources, for even large systems.

To validate SQE_{AB}, we implemented and tested the model on a variety of neutral and charged zeolites and MOFs. To obtain SQE_{AB} charges for these systems, we employed either a set of *ad hoc* parameters or parameters fit to reproduce REPEAT charges on a small set of neutral and charged zeolites. For the neutral frameworks, we showed that the SQE_{AB} model elegantly reproduces the charges of the original SQE method. Comparing between the Hessian of SQE_{AB} energy expression and that of QEq, we showed mathematically that the incorporation of adjustable parameters into the off-diagonal matrix elements makes the SQE_{AB} model more flexible than QEq. In this regard, we proposed that SQE_{AB} should more accurately capture the chemical structure of charged frameworks compared with QEq. Comparing between QEq, SQE_{AB} and REPEAT charges on a charged zeolite framework, both the SQE_{AB} and REPEAT methods showed a similar pattern

for the O atom charges, where O atoms covalently bonded to an Al atom were more polarized then those that were not. Conversely, with the QEq model the charges on the oxygen atoms covalently bound to aluminum and those bound solely to silicon were found to be intermixed, confirming the original proposal.

With charged aluminosilicate zeolites being the most highly utilized porous materials in industry, having accurate and fast methods for generating charges in these systems is of great value. In Chapter 4 we turn the focus to aluminosilicate systems, and modeling their adsorption properties. A thorough literature search revealed that many of the currently used force fields for modeling guest-host interactions in zeolites use generic charge sets, wherein each atom of the same type is assigned the same charge throughout the entire framework. Although these charges can provide accurate results in specific cases, they do not account for the structural diversity of different zeolite frameworks. Furthermore, we found that there was significant variation in the value of the charges across the different zeolite force fields considered, with the derivation of most of the charges not being physically motivated. Taking advantage of the work presented in Chapter 3, we developed a force field using physically relevant charges, derived from the periodic QM ESP of charged aluminosilicate zeolites for the first time.

Recently, it was shown by Fischer *et al.*[8] that generic charges tend to underpredict the electrostatic interactions in a majority of the cases that were investigated. In particular, comparing between the simulated uptake when using either REPEAT or their own set of generic charges for neutral siliceous zeolites, they found that in half the cases examined the RMSPE between predicted uptakes was non-negligible, exceeding 10%. While REPEAT is a robust first-principles based method that is preferred for accuracy, the costly electronic structure calculations that are required for each unique framework, makes it impractical in many cases. In such instances, faster alternatives

are required. Although generic charges have been traditionally used for this reason, finding a balance between speed and accuracy would be highly desirable for those modeling the adsorption properties of aluminosilicates. With the introduction of the SQE_{AB} method, the research project presented in Chapter 4 applies this model toward finding such a balance.

In the first part of this project, SQE_{AB} parameters are fit to reproduce the gauge-corrected QM ESP on real-space grid points. In line with the results of Fischer *et al.,* we carefully tailored the training set to be structurally representative of all siliceous and aluminosilicate zeolites found in the IZA-SC database, where 111 unique framework types containing either Si, or a combination of Si and Al were identified. We further generated structures with varying Si/Al ratios to ensure transferability of the model to a variety of compositions, in addition to a diversity of topologies. All structures that were used in both parameter training and validation were assigned an Si/Al ratio of 1, 3 or ∞ after analyzing the relative proportion of zeolites that exist with different compositions. The final training set was comprised of a total of 127 siliceous and aluminosilicate zeolites, made up of 63 unique framework types. The validation set is then comprised of the remaining 48 unique framework types identified in the IZA-SC database as being siliceous of aluminosilicate zeolites. After different Si/Al ratios to these frameworks different Si/Al ratios, the validation set included a total of 89 frameworks.

The SQE_{AB} parameters were then simultaneously optimized to minimize all frameworks' MAD between the ESP resulting from SQE_{AB} charges and the QM ESP, on a set of real-space grid points, using a custom genetic algorithm developed by Sean Collins of the Woo research group. To reflect how O atoms that are bound to Al in aluminosilicates tend to be more polarized than those bound exclusively to Si, we adopted an atom typing scheme regularly employed with generic charge models. The overall fitness of the parameters was evaluated with respect to the arithmetic mean of

the individual MADs calculated for each framework. In this way, each framework in the training set was weighted equally, preventing any bias towards larger structures with unit cells containing more ESP-fitting points. The GA was considered fully converged after the most fit solution remained the same for ten successive generations. The overall best set of parameters were then selected from five independent runs of the GA, and subsequently refined using a custom GD algorithm. The final set of optimized parameters for SQE_{AB} were termed the SQE_{AB}-AMP (split-charge equilibration in the atomic basis Aluminosilicate MicroPorous) parameter set.

To validate the parameters, we performed a series of test to evaluate how well the generated charges reproduce the QM ESP for a collection of zeolite framework types not included in the training set. In the first test, we compare directly between different methods for modeling the electrostatics in aluminosilicates to see how well the SQE_{AB}-parameters perform to current methods and models. Of particular interest for this work is the performance of the SQE_{AB}-AMP model with respect to generic charge models that are typically employed for modelling zeolites. Comparison with the generic charge sets developed by Fang *et al.*[9] and Calero *et al.*[10] showed that the system specific SQE_{AB}-AMP charges better reproduced the QM ESP by 22.16% and 35.98% on average, respectively.

We then considered the maximum MAD for the top 90% of the frameworks in the training and validation sets found using the different charge methods investigated. The top 90% of frameworks corresponds to frameworks ranking in the 90^{th} percentile with respect to the MAD in the QM and SQE_{AB}-AMP ESPs. For SQE_{AB}-AMP, the maximum MAD was found to be 20.6, whereas for the generic charge sets published by Fang *et al.* and Calero *et al.* we found that the maximum MADs were 40.6 and 45.0 mHartree, respectively. Assessment of top 90% of the validation set showed similar results, where we determined the maximum MAD for the SQE_{AB}-AMP

model to be 35.9 mHartree. This was in contrast to the generic charges sets of Fang and Calero, where the maximum MAD was found to be 60.8 and 47.7 mHartree, respectively. Therefore, of the empirical methods assessed in this work, we conclude that the SQE_{AB}-AMP model provides the best description of the QM ESP over all siliceous and aluminosilicate zeolite frameworks considered.

To ensure that the parameters were transferable across a variety of framework types, and that the parameters weren't biased towards the training set structures, we compared the mean of the MADs from the QM ESPs across the training and validation sets. The REPEAT method provides the ideal scenario for this comparison, where the MAD remains constant between the two sets of frameworks. For the SQE_{AB}-AMP model we showed that there is relatively small change of 1.1 mHartree in the mean deviations across the training and validation sets. On the other hand, for the generic charge models the shift in the mean of the MADs from the QM ESP from the training set frameworks to the validation set framework was found to be as high as 3.4 mHartree.

We next set out to validate the inclusion of additional parameters to distinguish between two types of oxygen atoms in aluminosilicates, by comparing the partial atomic charges calculated with SQE_{AB} and QEq, with and without atom-typing, to the framework REPEAT charges for the MFI-type zeolite (Si/Al = 3). The additional atom-typing is shown to provide a better description of the chemical structure of charged aluminosilicate systems, where Al-neighboring O atoms are more negatively polarized than the non-neighboring O atoms. This behavior was evident in the charge distribution where the O atom charges were clearly segregated into two groups. Examining the REPEAT charge distribution, a very similar distribution was exhibited. Beyond providing justification for implementing an atom-typing scheme that distinguishes between two types of oxygen atoms, this result further validates the quality of the SQE_{AB}-AMP model for generating meaningful atomic charges, similar to those obtained with REPEAT.

As a final assessment of the performance of the SQE_{AB}-AMP model, we evaluated the CO_2 uptake in a series of 29 siliceous and aluminosilicate frameworks, using both SQE_{AB}-AMP and REPEAT to model the electrostatic interactions. This analysis is meant to be similar to that carried out by Fischer and Bell,[8] where the CO_2 uptakes predicted using either generic charges of REPEAT charges to model the electrostatics were compared. By comparing the simulated uptakes obtained with either charge model, we are able to indirectly evaluate how well the SQE_{AB} charges represent the ESP of a given system. In particular, since CO_2 has a non-negligible quadrupole moment of -13.4 $\times 10^{-40}$ C m^2,[11] we expect that the electrostatic interactions to play an important role in determining the CO_2 adsorption properties of siliceous systems, and even more so for charged aluminosilicate frameworks.

In the study carried out by Fischer and Bell, the CO_2 uptakes calculated using generic charges to model the electrostatic interactions in siliceous zeolites were found to differ by a root-mean-square percent error (RMSPE) of more than 10% in half the cases considered, with a maximum RMSPE of ~23%. This is in contrast to the uptakes calculated using SQE_{AB}-AMP charges, where the RMSPE did not exceed 5% for a representative cross-section of 29 siliceous and aluminosilicate frameworks taken from the validation set. On average, we found that the SQE_{AB}-AMP charges reproduced the simulated CO_2 adsorption isotherms obtained using REPEAT charges with a RMSPE of 1.47%, and a maximum RMSPE of 4.38%. By plotting the SQE_{AB}-AMP uptakes against those obtained using REPEAT charges, we found indication of systematic over- and under-estimations in the CO_2 loading at low and high pressures, respectively, albeit negligible. To substantiate these results, we then calculated both the Pearson and Spearman rank-order correlation coefficients. The Pearson correlation was determined to be 0.989, while the Spearman correlation coefficient was found to be 0.992. In either case, we concluded that the SQE_{AB}-AMP model provided an approximation to the ESPs in charged and neutral zeolites that was in excellent

agreement with the REPEAT method. From this result we noted that, for the same LJ potential parameters, the two charge methods should yield very similar prediction of the adsorption properties of a given zeolite framework.

In the final part of Chapter 4, we shift our focus to the development of a force field for modeling gas adsorption in zeolites, for which both SQE_{AB}-AMP and the REPEAT charges can be used in place of generic charges sets that are typically used for modeling the electrostatic interactions. In this work we compared our results against force fields developed by Calero *et al.*[10,12] and Fang *et al.*[9] The principle aim of the AMP force field is to be able to accurately model the guest-host interactions involved in post-combustion CO_2 separations and natural gas upgrading. To ensure that the parameters are compatible with popular guest force fields, the guest parameters were held constant during the optimization procedure. Rather, we employed force fields for CO_2, CH_4, and N_2 that have been used extensively in modeling gas uptake in porous materials, and that were shown to accurately reproduce phase equilibria properties and the multipole moments of these molecules. The AMP force field was optimized using the same force fields to model the guest molecules as the other models considered in this work, with the exception of CH_4. Here we adopted a 5-site model in favor of the single-site model most commonly used because the 5-site model was shown to better describe the dense packing of methane in the small pores of zeolites.[13]

As is common practice when simulating adsorption in cation exchanged zeolites,[10,14,15] the zeolite frameworks were assumed to be rigid and the cations were allowed to move freely within the simulation cell. To achieve a robust model, we optimized the cross-species parameters representing the distinct intramolecular interactions present, fitting these simultaneously to 36 experimental isotherms. This data was taken from independent sources, with 12 isotherms for each guest being modeled (CO_2, CH_4, N_2). In total we identified and optimized 20 cross-species interactions (17

unique parameters), after $Na^+\cdots Na^+$, $Na^+\cdots Si$, and $Na^+\cdots Al$ interactions were excluded as, these species are prevented from approaching each other due to repulsive Coulombic interactions. For the AMP force field, we adopted a +1.0e charge on the Na^+ cations, so as to match framework Al densities during simulation, and we use REPEAT to model the framework atomic charges during parameter optimization.

For our LJ parameter optimization, we adopted an approach similar to that employed by Fang *et al.*[9,16] This scheme relies on the assumption that the guest-host interactions in zeolites are represented by pairwise VdW and Coulombic terms, which can be expressed in terms of semi-empirical attractive and repulsive coefficients, developed by Grimme[17], and corresponding scaling factors. In the optimization procedure the scaling factors act as the free variables. In the work carried out by Fang *et al.* [9,16] two sets of scaling factors were used, one to model the interactions with siliceous zeolite frameworks (Si···, O··· based interactions), and another set to model the interactions with aluminosilicate frameworks (Al···, $Na^+\cdots$ based interactions). Though reducing the number of free-variables in the fitting procedure naturally results in less flexible parameters, this scheme has the advantage of requiring only a singled constraint to be enforced during the parameter optimization. Since the AMP force field employs physically motivated charges, we expected that even with fewer degrees of freedom in the parameter optimization, the AMP model would outperform the other available force fields, which employ generic charges derived for charge neutral zeolites.

To demonstrate the validity of the model parameters using either charge model, we evaluated the force fields performance in predicting three different physical observables directly related to accuracy of the simulated guest-host interactions, using both charge models. We then compared our results with other force fields from the literature. Firstly, we examined the simulated N_2 uptakes and heats of adsorptions obtained with the different force fields investigated. All of the models were

found to be in excellent agreement with experiment over the conditions and framework types considered. In general, the AMP force field was shown to under predict the N_2 uptakes at higher loadings. In contrast, the Calero *et al.* force field was seen consistently overestimating the gas uptakes across all of the conditions and framework types considered. Looking at the average percent error (PE) and associated standard deviation in the predicted uptakes, which were found to be 15.27 ± 2.65% using REPEAT charges, 22.38 ±4.86% when using SQE_{AB} charges. Relative to the Calero *et al.* force field for N_2 adsorption, which was found to deviate from experiment by 34.08 ± 37.97%, the AMP force field with either charge method offer significant improvements in accuracy (>10% decrease in the PE for the AMP force field). The Pearson correlation (least-squares) coefficient for the REPEAT-AMP and SQE_{AB}-AMP are both ~0.98, closely followed by Calero *et al.* at ~0.97. Similar results were obtained with the Spearman rank correlation coefficients.

The force fields evaluated are shown to over predict the energetics of N_2 adsorption. The relative percent errors were determined as +22.72%, +6.64%, and +3.2% for Calero *et al.*, SQE_{AB}-AMP, and REPEAT-AMP force fields, respectively. Of the methods consider, the AMP force field was shown to give the best agreement with experiment with an average PE of 10.21 ± 8.60% when using the optimized SQE_{AB} charges and 3.21 ± 2.95% when using REPEAT charges on the framework atoms. This presents a significant enhancement over the Calero *et al.* force field, where the average PE was found to be 35.30 ± 34.09%. The Pearson correlation coefficient was found to be 0.818 for the SQE_{AB}-AMP force field and 0.954 for the REPEAT-AMP force field. In comparison with the Calero *et al.* force field, where the Pearson correlation coefficient was found to be 0.515, the AMP force field is shown to provide a better model for the energetics of N_2 adsorption in aluminosilicates.

The performance of the AMP force field with respect to CO_2 adsorption in aluminosilicates was next assessed relative to experiment, and compared with the Calero *et al.* and Fang *et al.* force fields. Fairly similar behavior was observed in the plotted simulated uptakes for the four different force fields investigated, with all of the models considered demonstrating a tendency to over predict the CO_2 uptakes in aluminosilicates. The sole exception to this trend were the results obtained using the SQE_{AB}-AMP model. Looking at the simulated CO_2 HoAs, however, the AMP model is much more evidently seen reproducing experiment within a smaller range of error relative to the other force fields, were significant overestimations are observed. From a comparison of the average PEs, it was found that the AMP force field offers a minimum improvement over the others by ~5.0% when using SQE_{AB} charges, where the average PE was found to be 25.89±16.54%. Even better results were obtained when using DFT derived REPEAT charges, where the average PE was found to be 19.92±9.90%. Though the improvement when using SQE_{AB} charges is relatively small, further consideration of the quantitative analysis indicated that the AMP force field was more reliable than the others. For instance, we noted that the small standard deviations found with the AMP force field for both N_2 and CO_2 provided a strong indication of the robustness of the model across the framework types and conditions, which are important to post-combustion CO_2 capture, or CO_2/N_2 separations.

When considering the heats of CO_2 adsorption, the AMP force field again is shown to outperform the other models considered. The average PE for SQE_{AB}-AMP was found to be 14.91±3.06%, while for the REPAT-AMP force field a smaller error of 13.26±9.97% was achieved. Both of these are less than half of the errors observed when using the Fang *et al.* force field, and well under a third of the errors obtained when using the Calero *et al.* force field. The Pearson correlation coefficient for the REPEAT-AMP and SQE_{AB}-AMP were determined as 0.996, and 0.973 respectively, while for Calero *et al.* and Fang *et al.* the correlation between the simulated and

experimental HoAs were found to be 0.310 and 0.461, respectively. For the AMP model, similar results were obtained with the Spearman rank correlation coefficients. For the other force fields, a larger change between the Pearson and the Spearman Rank order coefficients was observed, generally suggesting the existence of outliers in the data.

For modeling methane interactions with aluminosilicate zeolite frameworks, the AMP force field was evaluated against experiment and compared with the Calero *et al.* force field for small molecule gases. From the simulated uptakes, all of the models appeared to overestimate the uptake at lower loadings, while underestimating the uptake slightly at higher loadings. A similar trend was observed when the CH_4 HoAs were evaluated. Relative to the Calero *et al.* force field, the AMP model again achieved a greater degree of accuracy in predicting CH_4 adsorption characteristics within aluminosilicate materials. When simulating uptakes, the average PEs for the AMP force field when using either REPEAT or SQE_{AB} charges were determined as 17.33±16.44% and 26.50±14.11, respectively, while the average PE for the Calero *et al.* force field was found to be 40.87±33.95%. In terms of linear correlation (Pearson) or ranking (Spearman), the AMP force field was again seen to offer an increase in accuracy relative to the Calero model, with Pearson correlations of ~0.88 and ~0.81 for REPEAT and SQE_{AB} relative to ~0.59 for Calero. Similar results were obtained with the Spearman rank order coefficients.

When simulating the CH_4 HoAs, the average PEs for the AMP force field when using either REPEAT or SQE_{AB} charges were found to be 24.98±17.19% and 26.45±17.39%, in contrast with 62.93±29.53% for Calero *et al.* The results obtained using the SQE_{AB}-AMP and REPEAT-AMP force fields were found to have very fairly similar performance here, just as was observed when modeling CO_2 interactions. . Due to the large quadrupole moment of CO_2, the calculated adsorption properties are strongly affected by the framework charges. Methane, on the other hand,

is a non-polar molecule for which we would expect the simulated uptakes to be unaffected by the framework charges. For this work, however, we have adopted a five-site model with atomic charges and LJ potential parameters assigned to each atom in CH_4. With this in mind, we concluded that these results were in fact are a further validation of the optimized SQE_{AB} charges to reproduce the QM ESP as well as periodic DFT derived REPEAT charges. Although it was apparent that at lower energies REPEAT charges tend to over-predict the HoA, while when using SQE_{AB} charges we observe a slight under-prediction, the two charge methods were generally found to reproduce the same results using the same set of LJ parameters. From the linear regression analysis, we saw lower Pearson correlation coefficients for the AMP force field relative to the other guests, with 0.6 for REPEAT-AMP and ~0.29 for SQE_{AB}-AMP. The few deviations from experiment for the results obtained using the REPEAT- and SQE_{AB}-AMP models, which were found to be ~1.0 kcal/mol in some cases, have a large effect on the calculated correlation coefficients due to the small number of data points included. The same can been seen with Calero *et al.* where deviations can be on the order of ~2.0 kcal/mol, and the corresponding Pearson R^2 value is ~0.25. Although the framework topology or composition was thought to possibly have an effect on these results, no relationship between the observed error and the composition was found. In general, as was stated, the force fields seem to deviate most at lower energies, which would more likely be related to the temperature and pressure at which adsorption occurs.

As final test of the efficacy of the AMP force field in modeling guest-interactions in aluminosilicate zeolites, we applied the force field towards determining the location of the extra-framework sodium ions in the LTA-4A framework (Si/Al = 1.0), originally resolved by Pluth *et al.*[18] Although it is generally not possible to determine the Al-sitings and extra-framework cation locations from traditional XRD methods, the high symmetry of the LTA framework along with the equal number of Si and Al atoms occupying the T-sites when fully saturated with Al, make this

possible when following Lowenstein's rule. In essence, there is only one possible configuration of Al atoms that allow for LTA to exist with a Si/Al ratio of 1. The single crystal X-ray structure of the LTA framework was shown, where only 48 Na^+ ions were required to neutralize the unit cell, whereas 72 counter-ions locations were identified from the XRD experiment. In other words, there are only 48 Na^+ counter-ions on average per unit cell, which are spread across 72 sites within the crystal due to disorder in the cation positions. In the 4 central channels that are seen in the LTA crystal structure, the 4 locations that are shown are the perfect archetype of dynamic or static disorder commonly observed in aluminosilicates. Thus, using a well sampled Monte Carlo simulation of the Na^+ positions, we were able to identify all of these binding sites as well as the disorder in the extra framework cation positions, using either the more accurate REPEAT-AMP force field or the faster SQE_{AB}-AMP model.

Although the AMP model offers significant improvements over current force fields that utilize generic charge sets optimized on neutral zeolites, the flexibility of the parameters could be further improved. As discussed, the optimization schemed adopted from Fang *et al.* limits the flexibility of the model by reducing the number of free variables during the optimization procedure. To correct this we have proposed allowing each of the cross-species terms to vary independently of one another during the GA. By doing so, a more robust model with improved accuracy should be obtained. Using the results of this work, appropriate constraints that ensure physically meaningful parameters are obtained, can be rationalized.

5.3 Future Work

5.3.1 The SQE_{AB}-AMP Force Field

The primary objective in development of the SQE_{AB}-AMP force field was to create a model for simulating gas uptake in aluminosilicate zeolites with a charge generation method that offers a

balance between the speed of generic charges and the accuracy of system specific QM-derived charges. With an infinite number of possible zeolite frameworks, the SQE_{AB}-AMP force field would be of great value in screening these materials for post-combustion CO_2 capture and natural gas purification. Furthermore, the force field could be used to predict possible Al-sitings in zeolites using the computational method proposed by Calero *et al.*[19] The proposed method involves determining possible Al-sites by generating multiple Al configurations for a given framework type, and evaluating the simulated CO_2 adsorption against experimental data. The Al distribution that results in the isotherm most similar to experiment is considered the most likely configuration. This procedure for identifying Al-sitings could easily be automated, and would be of particular value to the Woo group in the future for modeling a greater diversity of frameworks, since the exact Al positions are not known in almost all aluminosilicates. Using the SQE_{AB}-AMP force field in conjunction with this method, uptake data for CH_4 and N_2 uptake in aluminosilicates could be used in addition to CO_2 adsorption data. Many of the algorithms that would be required to automate this process have already been developed for the research that was presented in this thesis and, therefore, would be easily implemented. Future inclusion of additional criteria related to the energetics of the structure could further improve the reliability with which the Al-siting are determined. Additionally, it may be of interest to relate synthetic conditions of the zeolites for which experimental uptake data is taken to the most likely Al-distributions that are identified computationally. If enough of this data was accumulated, it may aid in understanding the interplay of kinetics and thermodynamics that results in a given Al configuration.

Beyond implementing the force field towards specific applications, it could be extended to include a greater diversity of counter-ions. It has been demonstrated, both experimentally and computationally, that the extra-framework cations present in aluminosilicates play an important role in the binding of guests within charged frameworks.[14] In a relatively recent study, Shang *et al.*[20] used

DFT to study the effect of Cs^+ ions in the pore aperture of the chabazite aluminosilicate. The authors found that what appears to be molecular sieving based on shape selectivity, is actually based on the electrostatic interactions of the guest with the Cs^+ ions in the pore aperture. This new mechanism for selectively adsorbing guests relies on the ability of the guest molecule to induce a reversible deviation of the extra-framework cation from the center of the pore apertures. In essence, cations which tend to localize to the pore apertures will act as electrostatic 'gates', allowing guests that interact more favorably to pass more easily (with a lower energetic penalty). This new mechanism for discriminating between different guests also presents a new means of tuning the adsorption properties of aluminosilicates. Hence, having the ability to screen a diverse set of structure-cation combinations would be of great value for identifying new candidate materials for selective gaseous separations involving CO_2, CH_4, and N_2.

5.3.2 SQE_{AB} Parameter Optimization for Charged MOFs

Another future project related to the research presented in this thesis, is to optimize SQE_{AB} parameters for charged MOFs. Since one of the primary interest of the Woo group is in modeling and screening MOF materials for different applications, having a fast and accurate charge generation method would be of great value, particularly as research of charged materials gains in interest. Employing a procedure similar to that outlined in Chapter 4, the first step would be the development of a diverse and structurally representative training set of charged MOFs. In this regard, many structural databases containing charged MOFs are available, including the CoRE database[21] as well as the growing database of hypothetical structures developed in the Woo group. This diverse database currently contains millions of MOFS generated from combination of 23 inorganic and 175 organic SBUs, as well as 50 functional groups, and a total of 116 network topologies. In particular, frameworks should be included from the Computation-Ready Experimental (CoRE) database since it contains more than 4700 experimentally realized structures,[21]

while the other databases include hypothetical MOFs. In addition to accurately representing the structural diversity amongst MOFs, it is important to appropriately account for the diversity in the chemical composition of MOFs. To do so, a subsequent elemental analysis would be performed on the structurally representative set of MOFs selected in the first step. It is important that the training set included all possible atom types in relative proportions that reflects their natural occurrence amongst all framework considered. This will further ensure that all relevant bond-types are sampled during the parameter optimization. An example of such analyses are given in a recent parameter optimization of SQE for neutral MOFs carried out by Sean Collins of the Woo research group.[22]

To further improve the transferability of these parameters, including an atom-typing scheme such as was used for aluminosilicates in Chapter 4. However, a more sophisticated atom-typing model should be employed, which discriminates between atoms of the same element based on hybridization. For example, carbon atoms in MOFs can exists in both sp^2 and sp^3 hybridization states. Differentiating between these two types of carbon is of particular importance in charged MOF materials where the framework charge can arise from either unsaturated metal centers, or charged functional groups that can be added to the organic SBUs. In either case, the high degree of conjugation in the organic SBUs will result in delocalization of the added charge. By including parameters for different hybridizations states, the resulting model will more effectively describe the charge distribution across a diversity of environments, such the delocalized aromatic systems present in the organic SBUs. Having such an accurate and fast method for modeling the electrostatics in charged MOF materials would enable an extensive screening of hypothetical and experimentally realized structures to identify high-performing materials as well as new strategies for designing charged materials. Furthermore, this charge model could be used in combination with UFF parameters, which have recently been updated to include all moieties present in the CoRe database,[23] for rapid generation of hypothetical charged MOFs from SBU databases. This would in turn not

only be of great value to the Woo research group, which could expand its current database of charged MOFs, but to the MOF research community in general.

5.4 References

(1) Campañá, C.; Mussard, B.; Woo, T. K. Electrostatic Potential Derived Atomic Charges for Periodic Systems Using a Modified Error Functional. *J. Chem. Theory Comput.* **2009**, *5* (10), 2866–2878.

(2) Leslie, M.; Gillan, N. J. The Energy and Elastic Dipole Tensor of Defects in Ionic Crystals Calculated by the Supercell Method. *J. Phys. C Solid State Phys.* **2000**, *18* (5), 973–982.

(3) Yaghi, O. M.; O'Keeffe, M.; Ockwig, N. W.; Chae, H. K.; Eddaoudi, M.; Kim, J. Reticular Synthesis and the Design of New Materials. *Nature* **2003**, *423* (6941), 705–714.

(4) Seo, J.; Whang, D.; Lee, H.; Jun, S.; Oh, J.; Jeon, Y.; Kim, K. A Homochiral Metal-Organic Porous Material for Enantioselective Separation and Catalysis. *Nature* **2000**, *404* (April), 982–986.

(5) Dincă, M.; Dailly, A.; Liu, Y.; Brown, C. M.; Neumann, D. a; Long, J. R. Hydrogen Storage in a Microporous Metal-Organic Framework with Exposed Mn2+ Coordination Sites. *J. Am. Chem. Soc.* **2006**, *128* (51), 16876–16883.

(6) Dinca, M.; Long, J. R. Hydrogen Storage in Microporous Metal-Organic Frameworks with Exposed Metal Sites. *Angew. Chemie-International Ed.* **2008**, *47* (36), 6766–6779.

(7) Chen, J., Hundertmark, D. & Martínez, T. A. A Unified Theoretical Framework for Fluctuating Charge Models in Atomspace and in Bondspace. *J.Chem.Phys* **2008**, *129*, 214113.

(8) Fischer, M.; Bell, R. G. Modelling CO 2 Adsorption in Zeolites Using DFT-Derived Charges : Comparing System-Specific and Generic Models.

(9) Fang, H.; Kamakoti, P.; Ravikovitch, P. I.; Aronson, M.; Paur, C.; Sholl, D. S. First Principles Derived, Transferable Force Fields for CO2 Adsorption in Na-Exchanged Cationic Zeolites. *Phys. Chem. Chem. Phys.* **2013**, *15* (31), 12882.

(10) García-Sánchez, A.; Ania, C. O.; Parra, J. B.; Dubbeldam, D.; Vlugt, T. J. H.; Krishna, R.; Calero, S. Transferable Force Field for Carbon Dioxide Adsorption in Zeolites. *J. Phys. Chem. C* **2009**, *113* (20), 8814–8820.

(11) Graham, C.; Imrie, D. A.; Raab, R. E. Measurement of the Electric Quadrupole Moments of CO2, CO, N2, Cl2 and BF3. *Mol. Phys.* **1998**, *93* (1), 49–56.

(12) Martin-Calvo, A.; Gutiérrez-Sevillano, J. J.; Parra, J. B.; Ania, C. O.; Calero, S. Transferable Force Fields for Adsorption of Small Gases in Zeolites. *Phys. Chem. Chem. Phys.* **2015**, *17* (37), 24048–24055.

(13) Do, D. D.; Do, H. D. Evaluation of 1-Site and 5-Site Models of Methane on Its Adsorption on Graphite and in Graphitic Slit Pores. *J. Phys. Chem. B* **2005**, *109* (41), 19288–19295.

(14) Calero, S.; Dubbeldam, D.; Krishna, R.; Smit, B.; Vlugt, T. J. H.; Denayer, J. F. M.; Martens, J. A.; Maesen, T. L. M. Understanding the Role of Sodium during Adsorption: A Force Field for Alkanes in Sodium-Exchanged Faujasites. *J. Am. Chem. Soc.* **2004**, *126* (36), 11377–11386.

(15) E. García-Pérez, †; D. Dubbeldam, ‡; T. L. M. Maesen, § and; S. Calero*, †. Influence of Cation

Na/Ca Ratio on Adsorption in LTA 5A: A Systematic Molecular Simulation Study of Alkane Chain Length. **2006**.

(16) Fang, H.; Kamakoti, P.; Zang, J.; Cundy, S.; Paur, C.; Ravikovitch, P. I.; Sholl, D. S. Prediction of CO 2 Adsorption Properties in Zeolites Using Force Fields Derived from Periodic Dispersion-Corrected DFT Calculations. **2012**.

(17) Grimme, S. Semiempirical GGA-Type Density Functional Constructed with a Long-Range Dispersion Correction. *J. Comput. Chem.* **2006**, *27* (15), 1787–1799.

(18) Pluth, J. J.; Smith, J. V. Accurate Redetermination of Crystal Structure of Dehydrated Zeolite A. Absence of near Zero Coordination of Sodium. Refinement of Silicon,aluminum-Ordered Superstructure. *J. Am. Chem. Soc.* **1980**, *102* (14), 4704–4708.

(19) García-Pérez, E.; Dubbeldam, D.; Liu, B.; Smit, B.; Calero, S. A Computational Method To Characterize Framework Aluminum in Aluminosilicates. *Angew. Chemie Int. Ed.* **2007**, *46* (1–2), 276–278.

(20) Shang, J.; Li, G.; Singh, R.; Gu, Q.; Nairn, K. M.; Bastow, T. J.; Medhekar, N.; Doherty, C. M.; Hill, A. J.; Liu, J. Z.; Webley, P. A. Discriminative Separation of Gases by a "Molecular Trapdoor" Mechanism in Chabazite Zeolites. *J. Am. Chem. Soc.* **2012**, *134* (46), 19246–19253.

(21) Chung, Y. G.; Camp, J.; Haranczyk, M.; Sikora, B. J.; Bury, W.; Krungleviciute, V.; Yildirim, T.; Farha, O. K.; Sholl, D. S.; Snurr, R. Q. Computation-Ready, Experimental Metal–Organic Frameworks: A Tool To Enable High-Throughput Screening of Nanoporous Crystals. *Chem. Mater.* **2014**, *26* (21), 6185–6192.

(22) Collins, S. P.; Woo, T. K. Split-Charge Equilibration Parameters for Generating Rapid Partial Atomic Charges in Metal–Organic Frameworks and Porous Polymer Networks for High-Throughput Screening. *J. Phys. Chem. C* **2017**, *121* (1), 903–910.

(23) Coupry, D. E.; Addicoat, M. A.; Heine, T. Extension of the Universal Force Field for Metal–Organic Frameworks. *J. Chem. Theory Comput.* **2016**, *12* (10), 5215–5225.