

# Computational studies of hydrogen storage materials and the development of related methods

by

Timothy K. Mueller

A. B. Applied Mathematics  
Harvard University (1998)

Submitted to the Department of Materials Science and Engineering  
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY IN MATERIALS SCIENCE AND ENGINEERING

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

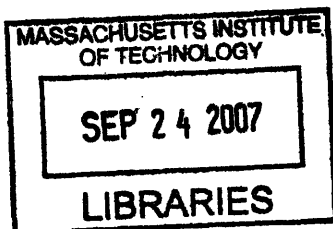
September 2007

© 2007 Massachusetts Institute of Technology. All rights reserved.

Author.....  
Department of Materials Science and Engineering  
July 25, 2007

Certified by.....  
Gerbrand Ceder  
R.P. Simmons Professor of Computational Materials Science  
Thesis Supervisor

Accepted by.....  
Samuel M. Allen  
POSCO Professor of Physical Metallurgy  
Chair, Departmental Committee on Graduate Students



ARCHIVES





# Computational studies of hydrogen storage materials and the development of related methods

by

Timothy K. Mueller

Submitted to the Department of Materials Science and Engineering  
on July 25<sup>th</sup>, 2007 in partial fulfillment of the requirements for the degree of  
Doctor of Philosophy in Materials Science and Engineering

## ABSTRACT

Computational methods, including density functional theory and the cluster expansion formalism, are used to study materials for hydrogen storage.

The storage of molecular hydrogen in the metal-organic framework with formula unit  $\text{Zn}_4\text{O}(\text{O}_2\text{C}-\text{C}_6\text{H}_6-\text{CO}_2)_3$  is considered. It is predicted that hydrogen adsorbs at five sites near the metal-oxide cluster, in good agreement with recent experimental data. It is also shown that the metal-oxide cluster affects the electronic structure of the organic linker, qualitatively affecting the way in which hydrogen binds to the linker.

Lithium imide ( $\text{Li}_2\text{NH}$ ), a material present in several systems being considered for atomic hydrogen storage, is extensively investigated. A variation of the cluster expansion formalism that accounts for continuous bond orientations is developed to search for the ground state structure of this material, and a structure with a calculated energy lower than any known is found. Two additional discrete cluster expansions are used to predict that the experimentally observed phase of lithium imide is metastable at temperatures below approximately 200 K and stabilized primarily by vibrational entropy at higher temperatures. A new structure for this low-temperature phase that agrees well with experimental data is proposed.

A method to improve the predictive power of cluster expansions through the application of statistical learning theory is developed, as are related algorithms. The Bayesian approach to regularization is used to show that by taking advantage of the prior expectation that cluster expansions are local, the convergence and prediction properties of cluster expansions can be significantly improved. A variety of methods to generate cluster expansions are evaluated on three different binary systems. It is suggested that a good method to generate cluster expansions is to use a prior distribution that penalizes the ECI for larger clusters more and has few parameters. It is shown that the generalized cross-validation score can be an efficient and effective substitute for the leave-one-out cross-validation score when searching for a good set of parameters for the prior distribution. Finally it is shown that the Bayesian approach can also be used to improve the convergence and prediction properties of cluster expansions for surfaces, nanowires, nanoparticles, and certain defects.

Thesis advisor: Gerbrand Ceder

Title: R. P. Simmons Professor of Computational Materials Science



# I. Overview



---

## I.1. Acknowledgements

---

I would like to thank my advisor, Professor Gerbrand Ceder. He was a valuable source of insight and research advice while at the same time allowing me sufficient freedom to pursue new ideas. This is a difficult balance to strike, but he does it very well.

I would also like to thank the members of the Ceder group, my housemates, and my other friends at MIT. They made my graduate career enjoyable and they have contributed greatly to my understanding of materials science.

I would like to thank my family members, who have been very supportive and have never stopped trying to understand exactly what it is that I do.

Most importantly, I would like to thank my fiancée Elisa. She has been loving, supportive, patient, and someone who I could always depend upon for intelligent and honest discussion.

---

## I.2. Table of contents

---

<b>I. Overview .....</b>	<b>5</b>
<b>I.1. Acknowledgements.....</b>	<b>7</b>
<b>I.2. Table of contents .....</b>	<b>8</b>
<b>I.3. Introduction.....</b>	<b>12</b>
<b>II. Materials for hydrogen storage .....</b>	<b>15</b>
<b>II.1. Background and motivation.....</b>	<b>17</b>
II.1.1. Department of Energy targets.....	18
II.1.2. Potential hydrogen storage systems.....	19
II.1.2.1. Compressed and cooled hydrogen.....	19
II.1.2.2. Chemical hydrogen storage .....	19
II.1.2.3. Adsorption of molecular hydrogen.....	19
II.1.2.4. Storage of atomic hydrogen .....	20
II.1.3. Thesis research.....	21
<b>II.2. Hydrogen storage in metal-organic frameworks.....</b>	<b>22</b>
II.2.1. Methodology .....	23
II.2.2. Structural results .....	25
II.2.3. Hydrogen adsorption sites .....	26
II.2.3.1. Comparison with other studies.....	36
II.2.3.2. Discussion.....	38
II.2.4. The effect of the framework on the BDC linker.....	44
II.2.5. Summary and Conclusion .....	47
<b>II.3. Lithium imide .....</b>	<b>48</b>
II.3.1. An antiferite model of lithium imide.....	49
II.3.1.1. Methodology.....	50
II.3.1.1.1. The effective Hamiltonian .....	50
II.3.1.1.2. Enthalpies of formation.....	55
II.3.1.2. Results .....	55
II.3.1.2.1. Effective interactions.....	55

II.3.1.2.2. A new low-energy structure.....	59
II.3.2. The low temperature phase.....	67
II.3.2.1. Methodology.....	70
II.3.2.2. Results .....	72
II.3.2.2.1. Local interactions.....	72
II.3.2.2.2. Ground state search.....	75
II.3.2.2.3. Finite-temperature behavior .....	83
II.3.2.3. Discussion.....	87
<b>III. A Bayesian approach to building cluster expansions.....</b>	<b>89</b>
<b>III.1. Background and motivation .....</b>	<b>91</b>
III.1.1. The cluster expansion .....	92
III.1.1.1. Effective cluster interactions (ECI).....	96
III.1.2. A new approach to developing cluster expansions.....	97
<b>III.2. Bayesian Regression.....</b>	<b>98</b>
III.2.1. Derivation of Bayesian regression.....	98
III.2.2. Prior distributions .....	101
III.2.2.1. The mean of the prior distribution.....	101
III.2.2.2. The shape of the prior distribution .....	102
III.2.2.2.1. The Gaussian prior.....	102
III.2.2.2.2. The Laplace prior .....	106
III.2.2.3. The width of the prior distribution .....	112
III.2.2.3.1. Cluster selection.....	113
III.2.2.3.2. Constant width.....	113
III.2.2.3.3. Independent widths.....	114
III.2.2.3.4. The exponential width function .....	114
III.2.2.3.5. The hierarchical width function .....	118
III.2.2.4. Parameterizing the width-generating functions .....	119
III.2.2.4.1. Leave-one-out cross-validation .....	119
III.2.2.4.2. Generalized cross-validation.....	121
<b>III.3. Sample generation .....</b>	<b>124</b>

III.3.1. Estimating the domain matrix .....	126
III.3.1.1. Direct enumeration .....	127
III.3.1.1.1. Enumerating structure lattices .....	129
III.3.1.1.2. Enumerating structure basis decorations .....	132
III.3.1.2. Analytical methods .....	135
III.3.1.2.1. All possible structures .....	135
III.3.1.2.2. Fixed composition .....	136
III.3.1.3. Statistical methods .....	138
III.3.2. Selecting the training set .....	140
<b>III.4. Experiments .....</b>	<b>142</b>
III.4.1. Sample data .....	142
III.4.1.1. Calculating input values .....	142
III.4.1.1.1. Generating the sample structures .....	142
III.4.1.1.2. Selecting training structures .....	143
III.4.1.2. Calculating output values .....	143
III.4.1.2.1. The embedded atom potential .....	143
III.4.1.2.2. The Tersoff potential .....	145
III.4.2. Model selection .....	145
III.4.3. Cluster expansion options .....	147
III.4.4. Results .....	149
III.4.4.1. Prior distributions .....	149
III.4.4.2. Regularized variable .....	151
III.4.4.3. Data sets .....	152
III.4.4.4. Weighting .....	156
III.4.4.5. Convergence .....	157
III.4.4.6. Discussion .....	177
<b>III.5. Further Considerations .....</b>	<b>181</b>
III.5.1. The posterior distribution .....	181
III.5.2. Continuous potential energy surfaces .....	182
III.5.3. Applying the Bayesian cluster expansion to surfaces .....	182



III.5.3.1. Example: Si-Ge nanowire .....	186
<b>IV. References .....</b>	<b>191</b>

---

## I.3. Introduction

---

This thesis has been divided into four chapters. The first chapter contains an overview of the thesis, and the last is a list of references. The middle two chapters contain the bulk of the research. In this introduction, I will briefly, and informally, provide the story behind the research in these middle chapters.

The second chapter is an exploration of materials for hydrogen storage, with a focus on mobile fuel-cell applications. This is a pressing problem that remains one of the key obstacles that must be overcome for the commercialization of mobile fuel-cell technology. Based on initial reports of positive experimental results, the storage of molecular hydrogen in nanoporous metal-organic frameworks was investigated. Later experiments showed that these materials do not have as much promise for hydrogen storage as initially thought, which is consistent with the calculations in this thesis.

Subsequently the focus of my research shifted to atomic hydrogen storage in systems based on lithium, nitrogen, and hydrogen (Li-N-H). A glaring problem in the study of Li-N-H-based systems is the fact that the structure of a key material, lithium imide ( $\text{Li}_2\text{NH}$ ), was unknown. Numerous attempts to identify the structure of this material have all arrived at different conclusions. In the end, I found two new structural phases for this material: one that is a likely candidate for the ground state at 0K, and another that is expected to be entropically stabilized at room temperature and metastable at lower temperatures. These results resolve discrepancies between calculations and experiments in previous studies and should help researchers better model hydrogen storage systems that contain lithium imide.

The cluster expansion methodology was repeatedly useful in the study of lithium imide. Cluster expansions have been used to successfully address a variety of materials science problems, but it is difficult to parameterize cluster expansions that have a large number of degrees of freedom. This is essentially a problem in function learning, a mathematical problem that has been studied in depth in the field of statistical learning theory. In an attempt to improve the parameterization of cluster expansions, I applied some of the general concepts in statistical learning theory to develop new methods for building cluster expansions. Because these methods are relatively unproven, I evaluated them on several sets of test data. The insights provided by statistical learning theory proved useful, as the new

methods repeatedly outperformed more traditional methods of building cluster expansions. A bonus is that the new methods are fairly easy to implement. Although many of the new methods presented in this thesis were developed in the context of the study of lithium imide, I am hopeful that researchers find them useful for other problems in materials science.



## II. Materials for hydrogen storage



---

## II.1. Background and motivation

---

As the global demand for hydrocarbon-based fuels rises and the supply of natural hydrocarbons falls, there has been a growing effort to develop alternatives to traditional power sources. Fuel cells, in which fuel is electrochemically converted into electricity, have emerged as leading contenders due to their high theoretical efficiency, low pollution relative to combustion technologies, and quiet operation. A large number of fuel cell systems have been developed based on a variety of fuels, electrodes, electrolytes, and operating environments. It is beyond the scope of this thesis to review these systems, but the curious reader can find thorough reviews in references [1-4].

A potential application of fuel cell technology is in the transportation sector, where cleaner, quieter, and more efficient alternatives to traditional internal combustion engines are sought [2-4]. If fuel cells are to be used for transportation (“on board” fuel cells) they will need to meet challenging technical requirements, including but not limited to the following [5]:

- High power density
- Short refueling time
- Long lifetime
- Quick startup

Based on these requirements, fuel cells that use proton exchange membrane (PEM) electrolytes are thought to be well suited for use in transportation [1-4]. PEM fuel cells operate at temperatures of about 80 °C, but no higher than approximately 100 °C, so they can be quickly started in normal atmospheric conditions. At the anode of a PEM fuel cell hydrogen is separated into protons and electrons. The electrons travel through an electrical circuit to the cathode, and protons diffuse to the cathode through a polymer membrane. At the cathode, the protons and electrons are typically combined with oxygen from the atmosphere to form water, which is the only waste product if pure hydrogen is used as a fuel [1-3]. For a PEM fuel cell to function, the fuel must contain hydrogen atoms to provide the required protons.

Before PEM fuel cells can be considered a viable alternative to combustion engines there are several technical challenges that must be overcome. One of the most significant of these challenges is the development of a means of storing hydrogen for use in these fuel cells. At atmospheric pressure, pure hydrogen is a gas above 21 K. If hydrogen is to be stored as a fuel for on-board fuel cells, it must be physically and/or chemically compressed. This storage system will add weight and cost to the overall fuel cell, making the choice of storage system critical for the overall fuel cell efficiency. In the following section, the current state of hydrogen storage research will be summarized.

### II.1.1. Department of Energy targets

The United States Department of Energy (DOE) has set forth technical targets for on-board hydrogen storage [5]. Some of the key targets are as follows:

Target Year	2007	2010	2015
Gravimetric density (Usable H <sub>2</sub> weight as a percentage of total storage system weight)	4.5%	6%	9%
Volumetric capacity (kg usable H <sub>2</sub> per liter of system volume)	0.036	0.045	0.081
Operating ambient temperature range (°C)	-20 to 50	-30 to 50	-40 to 60
H <sub>2</sub> delivery temperature from tank range (°C)	-20 to 85	-30 to 85	-40 to 85
System fill time for 5 kg H <sub>2</sub> (minutes)	10	3	2.5
Cost (\$US / kg usable H <sub>2</sub> )	200	133	67

**Table 1** Some of the technical targets for on-board hydrogen storage set by the U.S. Department of Energy

It is important to recognize that these are system targets, not material targets. It is estimated that depending on the final system, the hydrogen storage material capacities should be 1.2-2.0 times the given system target capacities [6].



## II.1.2. Potential hydrogen storage systems

---

To meet the DOE targets, researchers are investigating a wide variety of hydrogen storage methods. These different approaches will be briefly discussed here, and a more thorough discussion of the alternatives can be found in referenced review articles [7-10].

### II.1.2.1. Compressed and cooled hydrogen

The hydrogen could be stored as pure liquid H<sub>2</sub>, but it is difficult to maintain hydrogen in its liquid form without boil off, and there is typically a 30% energy loss during to the cooling process [2, 9]. Alternatively, hydrogen could be stored in compressed tanks, but current solutions are expensive [8] and unable to achieve the DOE 2010 goals for volumetric storage capacity [2, 9]. This is not surprising, considering the 2010 target for volumetric density is 64% the density of liquid hydrogen at its condensation temperature and 1 atm pressure. The 2015 system target is 114% the density of liquid hydrogen, and it is unlikely that either liquid hydrogen or compressed hydrogen will be able to efficiently achieve this goal.

### II.1.2.2. Chemical hydrogen storage

It is possible to chemically store hydrogen as part of a molecule, such as methanol (CH<sub>3</sub>OH), that is a liquid at fuel cell operating conditions. An on-board reformer would extract the hydrogen from this liquid when it is needed by the fuel cell. The disadvantages to this approach are the cost and additional weight of the on-board reformer, the release of gaseous pollutants, and the need for off-board regeneration of the liquid fuel if it is to be sustainable (although this may or may not be more efficient than the off-board generation of pure hydrogen) [3].

### II.1.2.3. Adsorption of molecular hydrogen

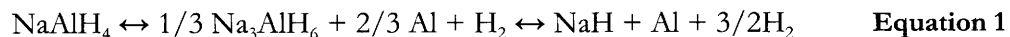
The adsorption of molecular hydrogen in porous materials and carbon-based nanostructures has also been extensively investigated. Although the kinetics of hydrogen storage in these structures is very good, with full uptake occurring in minutes, hydrogen uptake is generally not as exothermic as it is in materials that store atomic hydrogen, and the gravimetric densities tend to be too low. Early positive reports on hydrogen storage in

carbon nanotubes and nanofibers have not been reproducible, and it is believed that these materials hold a maximum of about 1 wt% hydrogen [8, 9]. Activated carbons have been more successful, achieving gravimetric storage densities of about 2% at 77 K and 1 bar [8, 11]. Gravimetric densities of up to 4.2% have been reported for collapsed BN nanotubes at room temperature 10 MPa pressure, but the hydrogen is only released at temperatures above 300 °C [12]. Initial reports of high hydrogen storage in a class of nanoporous materials known as metal-organic frameworks have been discounted in more recent experiments, and the best results to date are for less than 3 wt% gravimetric storage density at 77 K and 1 atm [13-15]. At pressures of 20 atm, storage capacities of over 6 wt% have been reported [15].

None of the solutions described above have reached the 2010 DOE targets for either gravimetric or volumetric storage densities, and remains unclear which, if any, type of hydrogen storage system will be able to reach these targets [2].

#### II.1.2.4. Storage of atomic hydrogen

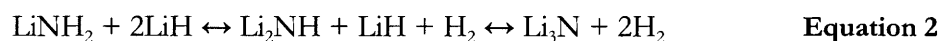
The remaining alternatives for hydrogen storage are based on solid-state materials. Metal hydrides can reversibly store hydrogen in the interstitial sites of the host metal, but the hydrides used to date store hydrogen at insufficient gravimetric densities (e.g.  $\text{LaNi}_5\text{H}_6$  at 1.37 wt%) and / or thermodynamically release hydrogen at temperatures that are too high (e.g.  $\text{MgH}_2$  at 1 bar and 300 °C) [7-9]. Higher gravimetric densities can be achieved through the use of complex metal hydrides. Some of these, such as sodium borohydride ( $\text{NaBH}_4$ ) do not store hydrogen reversibly, and produce solid waste products that must be removed for off-board recycling [16, 17]. Sodium alanate ( $\text{NaAlH}_4$ ) doped with Zr- and Ti- based catalysts, can reversibly store 3.7 wt% hydrogen at temperatures as low as 33 °C, although with very slow kinetics. At higher temperatures (around 100° C), an additional 1.8 wt% can be desorbed [17]. The two-stage reaction is as follows:



There are a wide variety of combinations of reactants and products that may combine to release hydrogen. The entropy of hydrogen gas at room temperature is approximately 130 J / mol K, meaning that for hydrogen release to occur thermodynamically at room temperature the reaction enthalpy should be approximately 40 kJ / mol  $\text{H}_2$  [7]. The

approximate reaction enthalpies of a large number of possible reactions have been calculated by Alavati et al. [18, 19], yielding a variety of candidates for new hydrogen storage systems that are awaiting experimental verification. However, they limited their calculations to systems for which the structures of all reactants and products are known.

A two-stage reversible adsorption process was discovered for lithium amide ( $\text{LiNH}_2$ ) [20]:



The first step in this process releases a theoretical maximum of 6.5 wt% hydrogen (ignoring the unmodified LiH), and the second step releases a theoretical maximum of 5.5 wt% (ignoring the already released  $\text{H}_2$ ). The total process could release 10.4 wt% hydrogen, although in practice only about 6 wt% can be desorbed at temperatures below 300 °C.

### II.1.3. Thesis research

---

In this chapter, computational methods have been used to study two of the more promising solid-state hydrogen storage systems: metal-organic frameworks, and Li-N-H based systems. Some of this work has been published in references [21, 22].

---

## II.2. Hydrogen storage in metal-organic frameworks

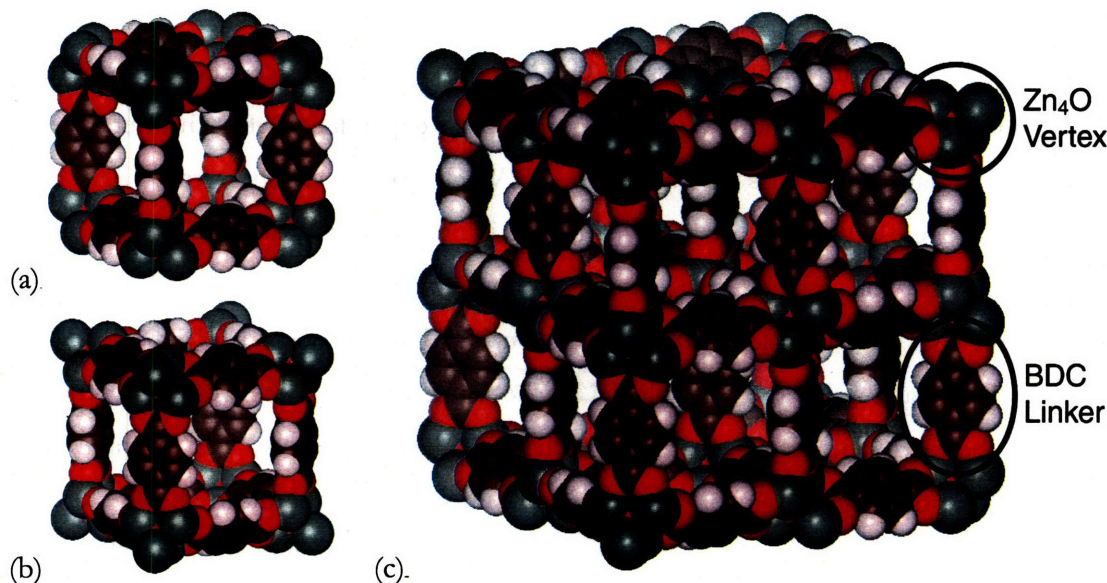
---

It is difficult to achieve the required storage densities in porous materials that adsorb molecular hydrogen due to the weak interaction of  $H_2$  with most surfaces and the difficulty in creating stable high-surface area materials [23, 24]. To increase the density of molecular hydrogen that can be stored in porous materials, it is necessary to maintain a sufficiently strong interaction between hydrogen and the pore surfaces and keep the ratio of the overall volume and weight of the material to the pore surface area to a minimum. In other words, it is necessary to find a material with high gravimetric and volumetric surface area densities. In addition, if the pores are sufficiently small, the attractive interactions between hydrogen and the opposing surfaces of the pore can facilitate multi-layer adsorption of hydrogen within the pore. For these reasons, although achievable volumetric hydrogen storage densities may be low, a class of materials known as metal-organic frameworks have been investigated for hydrogen storage [13, 25-28].

Metal-organic frameworks are formed by using organic ligands to connect small metal-based clusters of atoms, forming a periodic framework. The resulting frameworks are typically nanoporous materials with high surface area densities. Some are stable with empty pores up to several hundred degrees Celsius [25, 29, 30]. Frameworks of similar morphologies can be synthesized using a variety of organic linkers, providing the ability to tailor the nature and size of the pores. Several frameworks have been experimentally investigated for their abilities to store hydrogen, but to date none are able to do so at high enough densities to be useful for fuel cell technology [13, 26-28]. By learning more about how hydrogen interacts with these frameworks it may be better determined if their limitations are intrinsic or if frameworks can be designed that store  $H_2$  at higher densities.

In this section *ab initio* calculations are used to examine the storage of hydrogen in the metal-organic framework known as MOF-5 [29]. MOF-5 is formed by using 1,4-benzenedicarboxylate (BDC) to link together  $Zn_4O$  clusters. The resulting framework with formula unit  $Zn_4O(BDC)_3$  consists of cubic pores, where BDC forms the edges of the cubes and the  $Zn_4O$  clusters form the vertices. The width of each pore is approximately 13 Å. In half of the pores the face of the carbon rings faces towards the center of the pore, and in the remaining pores the edges of the carbon rings face the center, so that there are two formula

units per primitive cell. (See Figure 1) Experiments indicate that MOF-5 is capable of achieving a gravimetric hydrogen storage density of 1.3% at 77 K and 1 atm [13]. Although this is significantly less than the minimum target of 6% set by the United States Department of Energy, it is worth exploring how MOF-5 stores the hydrogen to determine if it is possible to design other framework structures with higher H<sub>2</sub> capacity.



**Figure 1** (a) The pore type that is surrounded by the edges of the BDC rings. (b) The pore type that is surrounded by the faces of the BDC rings. (c) The two pore types combined to form an eight-pore section of the framework. The atoms are colored as follows: Grey = Zn, Red = O, Brown = C, White = H.

### II.2.1. Methodology

MOF-5 has a highly symmetric periodic structure with 106 atoms per primitive unit cell. Because it is computationally expensive to perform calculations on systems of this size, it is tempting to model MOF-5 by looking only at a finite piece of the framework. For example, the computational cost of studying the BDC linker alone is significantly smaller than that of studying the BDC linker in the context of the framework. Later it will be argued that such an approach can give misleading results. In addition, studying only a part of the framework may bias the results to specific adsorption sites and overlook other important interactions between the framework and H<sub>2</sub>. For these reasons the full periodic structure has been modeled.

Modeling the interaction between molecular hydrogen and MOF-5 is a difficult task because a significant portion of the interaction may be caused by non-local electronic correlation. *Ab initio* methods that accurately evaluate this type of interaction typically scale poorly with system size, making them computationally expensive for systems such as MOF-5 with large unit cells. On the other hand, methods that scale well with system size, such as Density Functional Theory [31] (DFT) with the Local Density Approximation (LDA) or the Generalized Gradient Approximation (GGA), often fail to accurately calculate the magnitude of weak interactions because of the strong dependence of such interactions on electron correlation. Studies on weak van der Waals complexes show that LDA usually significantly overestimates the magnitude of these interactions [32-35]. The results for GGA depend strongly on which exchange-correlation functional is used [32, 34, 36, 37]. Functionals that obey the Lieb-Oxford bound, such as Perdew-Wang 1991 [38] (PW91) and Perdew, Burke and Ernzerhof [39] (PBE) are best able to model the weak interaction [32, 34]. Of these two the PBE functional usually gives the best results. For example, in two studies of the binding energies of rare-gas atoms into a dimer, the mean absolute error for PW91 was 7.7 meV [32] and 7.1 meV [35], whereas the mean absolute error for PBE was 2.5 meV [32] and 3.6 meV [35]. For these reasons DFT with the PBE exchange-correlation functional has been used for all calculations in this section.

All calculations were performed using the plane-wave DFT code from the Vienna *Ab-initio* Simulation Package (VASP). The Projector Augmented Wave [40, 41] (PAW) method was used, simplifying the problem to solving for the wavefunctions of 468 valence electrons per unit cell. Real-space projectors were used to evaluate the projected wavefunction character. For static calculations and relaxations an energy cutoff of 520 meV was used, and for molecular dynamics calculations, for which accurate energy evaluations are less important, a cutoff of 400 meV was used. Due to the large size of the unit cell, the only k-point used was the gamma point. Using an evenly spaced 2x2x2 gamma-centered k-point grid reduces the relaxed energy per 106-atom unit cell by 3.1 meV, or 0.03 meV per atom, supporting the use of only one k-point. The FFT mesh used is sufficiently large to prevent wrap-around errors.

To find the optimal structure for the MOF-5 framework with empty pores, the ions were relaxed at a series of lattice parameters with increasing and decreasing increments of

0.1% of the experimental lattice parameter of 25.9109 Å [42]. The shape of the unit cell was fixed for these calculations. The graph of the five lowest structural energies as a function of the lattice parameter was fit to a parabola using a least squares fit. One final relaxation was performed in which the lattice parameter was fixed to the value corresponding to the minimum of this parabola. The ionic positions for the resulting structure were used for all calculations in which hydrogen was adsorbed to the pores. Because this assumption does not account for the relaxation of the framework atoms due to H<sub>2</sub> adsorption, the interaction strengths reported in this section should be slightly weaker than those that would be obtained if full relaxation were allowed.

To calculate the interaction energies of hydrogen molecules three different reference energies were considered. The first is the energy of the system in which a single hydrogen molecule is placed in the center of the pore surrounded by the faces of the carbon rings. The second is the energy of the system in which a hydrogen molecule is in the center of the other pore. The third reference state is the sum of the energy of the relaxed empty framework and the energy of an array of hydrogen molecules placed on the same lattice as the framework but without the framework ions. All of these reference states are within 0.1 meV of each other, and for the values given in this section the third option is used. All interaction energies given in this section are per hydrogen molecule.

## II.2.2. Structural results

---

The calculated lattice parameter of the framework with empty pores is 26.137 Å, which is less than 1% greater than the experimental lattice parameter of 25.911 Å. Similarly, the calculated values for the Zn-Zn bond length, the C-C bond lengths, and the H-H bond length in the H<sub>2</sub> molecule are approximately 1% greater than their experimental values (see Table 2). Errors of this type are typical when using a GGA functional. The O-C-O bond angle is calculated at 126.1°, which is 0.2% greater than the experimentally measured value of 125.8° [42].

	Experimental	Calculated	Ratio
Lattice parameter	25.911 Å	26.137 Å	1.009
Zn-Zn bond length	3.181 Å	3.220 Å	1.012
C-C distance in carbon ring	1.396 Å	1.404 Å	1.006
O-C-O bond angle	125.8°	126.1°	1.002
H <sub>2</sub> bond length	0.741 Å [43]	0.750 Å	1.012

**Table 2** Experimental and calculated values for structural parameters of MOF-5 [42] and H<sub>2</sub>.<sup>1</sup>

### II.2.3. Hydrogen adsorption sites

The pore surfaces of MOF-5 contain numerous distinct sites at which hydrogen might adsorb. Rather than guessing which sites hydrogen might interact with most strongly, a molecular dynamics simulation was performed to determine the areas of high average hydrogen density. Seventeen hydrogen molecules per formula unit were included, which is approximately the density of hydrogen originally thought to adsorb in MOF-5 at 1 atm and 77 K [26]. This result has since been corrected, and it is currently believed that under these conditions MOF-5 adsorbs 5 hydrogen molecules per formula unit [13]. The molecular dynamics simulation was performed with a timestep of 2 fs using velocity rescaling to maintain the temperature at 77 K.

To initialize the molecular dynamics simulation, the pores were randomly populated by hydrogen molecules so that the centers of any pair of molecules were no closer than 2.3 Å from each other. The molecules were assigned random orientations and initial velocities. After four picoseconds, the simulation had achieved dynamical equilibration as measured by the frequency and magnitude of the fluctuations in total energy. The simulation continued for another 17.5 picoseconds, during which time the average hydrogen density was recorded. A smoothed hydrogen density distribution  $\rho(x)$  was calculated at each timestep using Equation 3, where  $n_i$  is the center of the  $i$ th hydrogen nucleus and  $\sigma = 0.5$  Å.

<sup>1</sup> The values for MOF-5 were measured at 30 K.

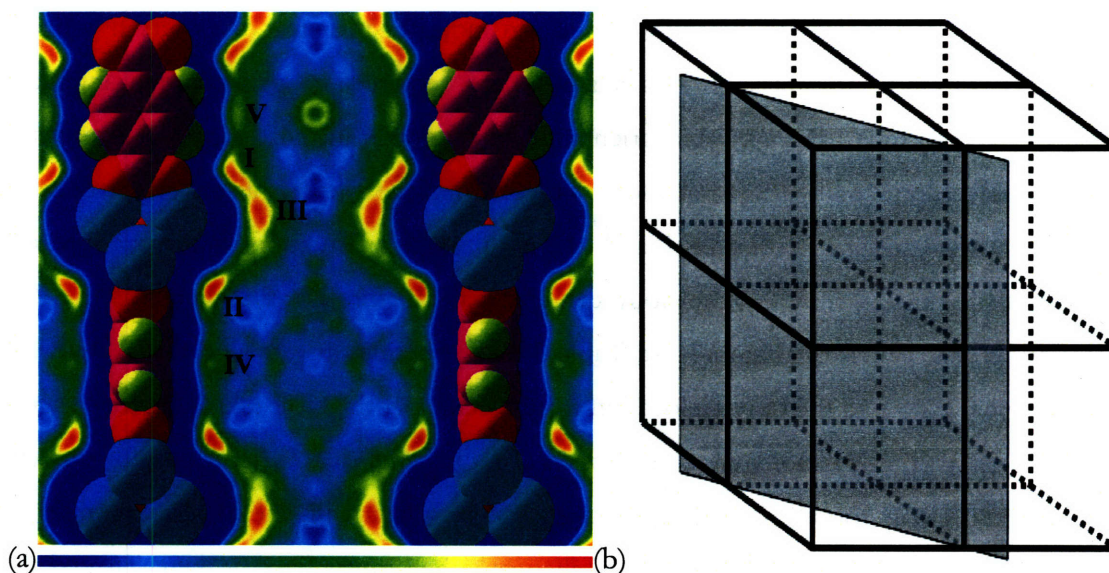


$$\rho(x) = \sum_i \rho_i(x)$$

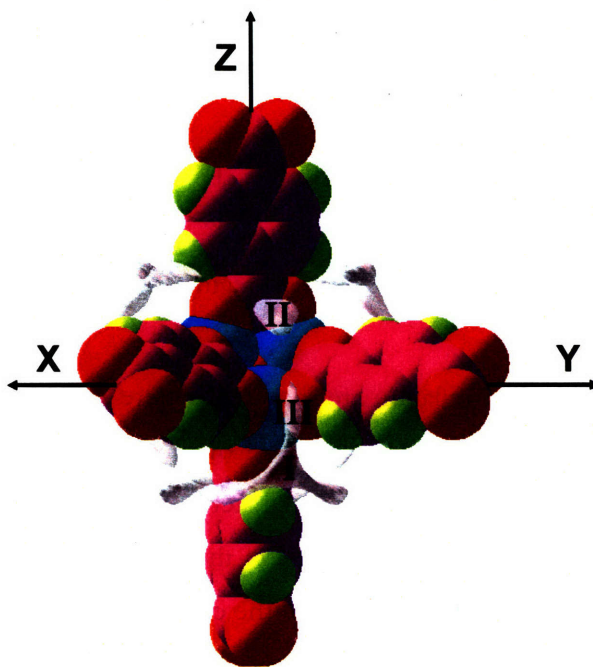
$$\rho_i(x) = \begin{cases} e^{-\frac{1}{2}\left(\frac{x-n_i}{\sigma}\right)^2} & |x-n_i| \leq 3\sigma \\ 0 & |x-n_i| > 3\sigma \end{cases} \quad \text{Equation 3}$$

The average density was calculated by averaging  $\rho(x)$  over all timesteps after the initial four picoseconds.

The results of this simulation are shown in Figure 2. The density is greatest near the  $\text{Zn}_4\text{O}$  clusters. Figure 3 shows the isosurfaces of the highest hydrogen density. There are three distinct local maxima near the  $\text{Zn}_4\text{O}$  clusters. Two of these are where they might be expected, at the high-symmetry sites located in the corners of the cubic pores. These sites are labeled as sites I and II, where site I is in the corner of the pore surrounded by the edges of the BDC linkers and site II is in the corner of the other pore. The third, labeled as site III, is not as intuitive; it is a low-symmetry site located in between the other two. This is potentially the most significant of the three sites because there are three times as many of sites in the framework of type III as there are of sites of type I or type II. The locations of these sites are given in Figure 3.



**Figure 2** (a) Hydrogen density on a plane in MOF-5 as calculated using first-principles molecular dynamics. The color scale that linearly corresponds to density is shown below the plot, with blue corresponding to zero hydrogen density. The five adsorption sites investigated in this thesis are marked in this plane. (b) The plane shown in the context of the three-dimensional MOF-5 structure. The dark lines correspond to the BDC linkers.



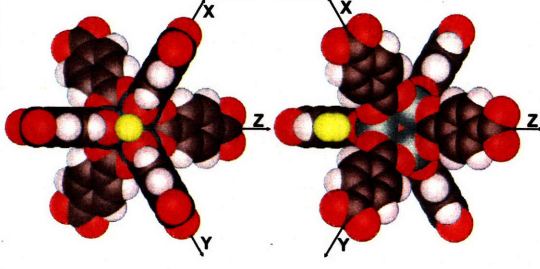
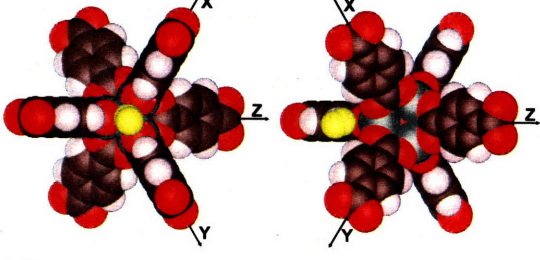
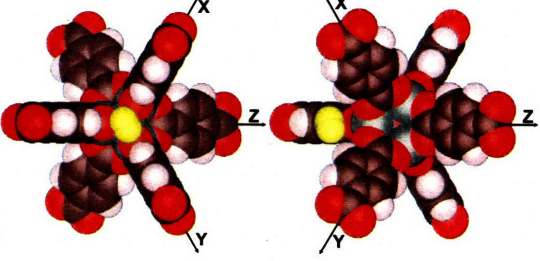
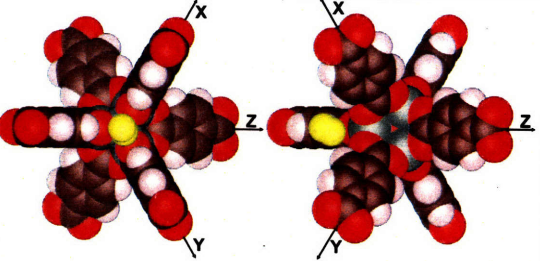
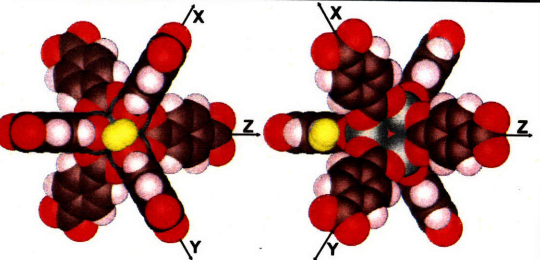
**Figure 3** Isosurfaces of high hydrogen density near the  $Zn_4O$  cluster. The three adsorption sites I, II, and III are labeled. In the coordinate axes shown, where the origin is at the center of the  $Zn_4O$  cluster, the coordinates of the three sites are approximately as follows, in angstroms: Site I: (3.4, 3.4, -3.4). Site II: (2.6, 2.6, 2.6). Site III: (4.4, 4.4, -1.3).

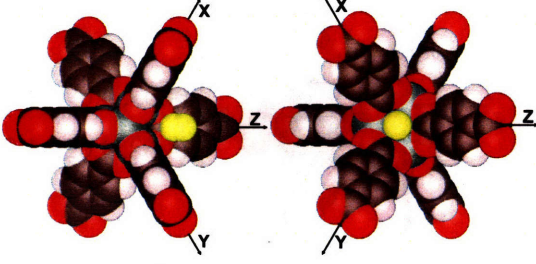
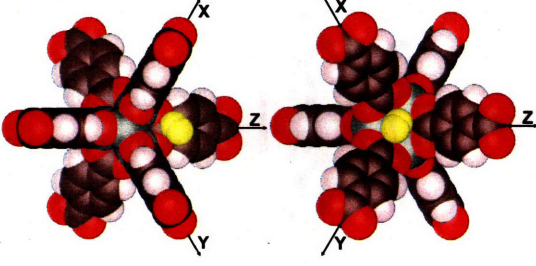
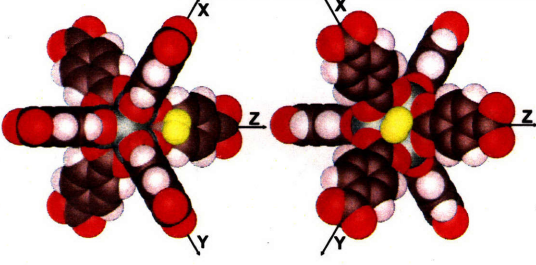
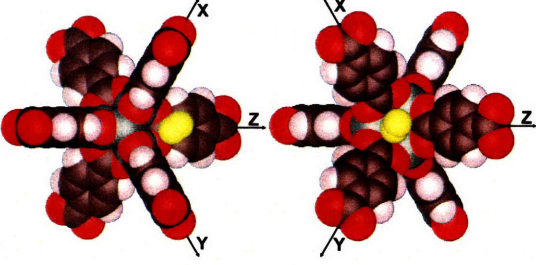
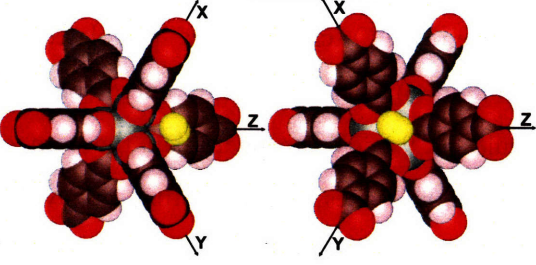
Whereas the molecular dynamics simulation indicates the areas to where hydrogen is attracted, ground-state relaxation calculations are required to get accurate interaction energies and to reveal the preferred orientation of hydrogen in these regions. To calculate the binding energy of hydrogen at each of these three sites it was necessary to try several different orientations. For sites I and II, an intuitive orientation is for the hydrogen molecule to be aligned along the body diagonal passing through the center of the  $Zn_4O$  cluster and the center of the pore. In Table 3 these orientations are labeled I(A) and II(A) respectively. To find the optimal location for the center of the molecule along this diagonal, a series of static calculations were performed at increments of 0.1 Å. Along this direction the minimum was found by fitting a parabola to the points with the lowest energies. The hydrogen bond length was not allowed to change; rather, it was fixed at the calculated value of the relaxed isolated hydrogen molecule, which is 0.7501 Å.

Other orientations at sites I, II, and III were also considered. (See Table 3) Unlike for orientations I(A) and II(A), the molecular centers for these orientations do not necessarily lie

on a fixed axis of symmetry. Because of this, the molecule was dynamically relaxed by calculating the forces on the atoms instead of doing a series of static calculations. Relaxation was stopped when the energy difference between two successive relaxation steps was less than 0.2 meV per unit cell. This same method was used to examine the interaction energy at various orientations near site III. In all these calculations the final bond length of the hydrogen molecule is within  $10^{-3}$  Å of the bond length of an isolated molecule, supporting the use of the frozen bond length for the static calculations. The results for these calculations are shown in Table 3. The adsorption energy is lowest (most negative) at site I, followed closely by site III. Site II has the highest (least negative) adsorption energy of the three sites, and is the only one for which the two hydrogen atoms are equidistant from the center of the  $Zn_4O$  cluster in the lowest energy orientation.



Site and orientation	Diagram	Hydrogen locations (Å)	Adsorption energy (meV)
I(A)		(3.65, 3.65, -3.65) (3.22, 3.22, -3.22)	-21.7
I(B)		(3.11, 3.18, -3.11) (3.42, 3.79, -3.42)	-19.9
I(C)		(3.11, 3.79, -3.11) (3.42, 3.18, -3.42)	-12.5
I(D)		(3.00, 3.44, -3.00) (3.53, 3.53, -3.53)	-20.2
I(E)		(3.53, 3.48, -3.00) (3.00, 3.48, -3.00)	-14.8

Site and orientation	Diagram	Hydrogen locations (Å)	Adsorption energy (meV)
II(A)		(2.88, 2.88, 2.88) (2.45, 2.45, 2.45)	-9.5
II(B)		(2.56, 2.24, 2.56) (2.67, 2.98, 2.67)	-13.6
II(C)		(2.46, 2.92, 2.46) (2.77, 2.30, 2.77)	-17.9
II(D)		(2.35, 2.66, 2.35) (2.88, 2.57, 2.88)	-13.3
II(E)		(2.35, 2.61, 2.35) (2.88, 2.61, 2.88)	-17.9

Site and orientation	Diagram	Hydrogen locations (Å)	Adsorption energy (meV)
III(A)		(4.09, 4.09, -1.25) (4.62, 4.62, -1.36)	-20.8
III(B)		(4.62, 4.09, -1.31) (4.09, 4.62, -1.31)	-7.5
III(C)		(4.35, 4.35, -0.94) (4.32, 4.32, -1.68)	-11.9

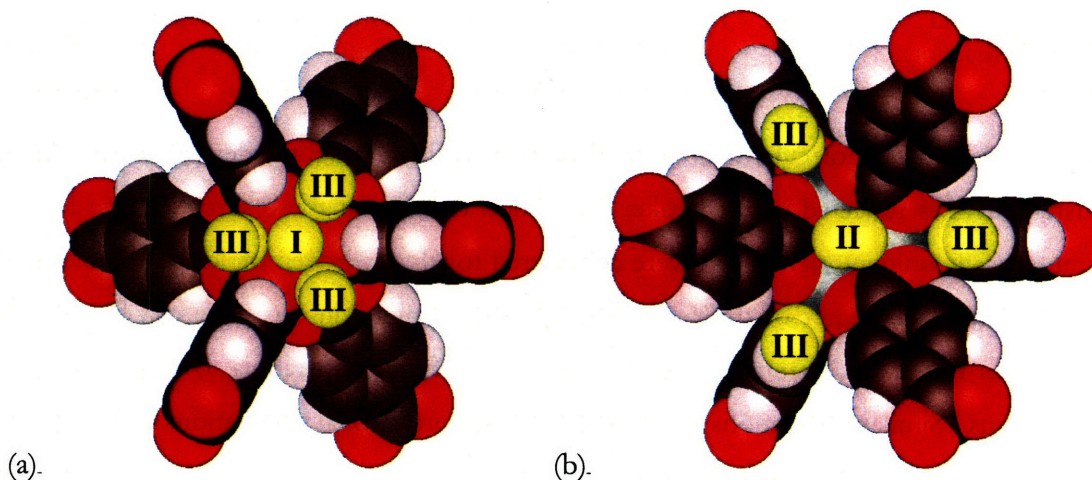
**Table 3** The orientations considered near sites I, II, and III and the calculated adsorption energies for those orientations.<sup>2</sup>

Around each  $Zn_4O$  cluster there are a total of 20 sites that are of type I, II, or III. Given the proximity of some of the sites to each other it may not be energetically favorable to populate all of these sites at the same time. A calculation was performed in which all sites of types II and III around a given  $Zn_4O$  cluster were populated with hydrogen in orientations II(C) and III(A) (see Figure 4). The interaction energy between the framework and the sixteen hydrogen molecules is 9 meV stronger than the sum of the interaction energies with hydrogen at the individual sites, meaning that on average the attraction becomes stronger by a little more than 0.5 meV per site as the vertex saturates with  $H_2$ . This

<sup>2</sup> The colors are the same as those used in Figure 1, and the adsorbed hydrogen is shown in yellow. The coordinates of each adsorbed hydrogen atom are given in the coordinate system shown, which is the same as the coordinate system used in Figure 3.

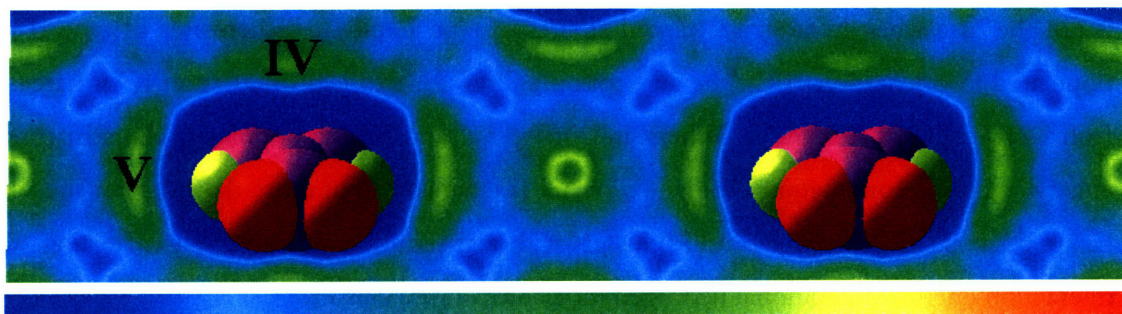


is not the case when hydrogen is placed at sites of type I and II. When sites of type II are populated in orientation II(C), the adsorption energy for site I in orientation I(A) increases from -21.7 meV to -6.3 meV. When sites of type II and III are populated, the adsorption energy at site I in orientation I(A) is -4.5 meV. It is unlikely that more than sixteen sites can be populated around any  $Zn_4O$  cluster without significantly adversely affecting the interaction energy of additional hydrogen molecules on the cluster.



**Figure 4** (a) Sites I and III simultaneously occupied with hydrogen in its lowest-energy orientation. (b) Sites II and III simultaneously occupied with hydrogen in its lowest-energy orientation.

Two symmetrically distinct areas of high hydrogen density are located near the BDC linker. One of these areas is located directly above the face of the carbon ring, and the other is located to the side of the carbon ring. (See Figure 5) These are labeled as sites IV and V, respectively.

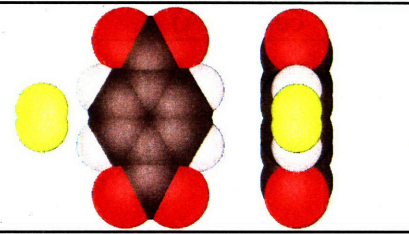
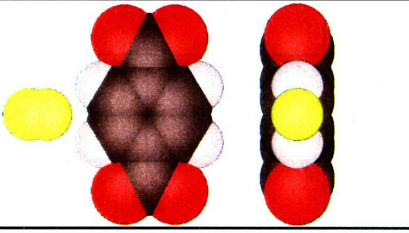


**Figure 5** Hydrogen density along a plane that slices through the centers of two BDC linkers. Sites IV and V are labeled. The color scale that linearly corresponds to density is shown below the plot.

The density of molecular hydrogen at these sites is lower than the density near the  $Zn_4O$  cluster. This is reflected in the weaker interaction energies at these sites. At site IV, five symmetrically unique orientations were considered for the hydrogen molecule. The optimal distance from the center of the carbon ring was determined in the same way the optimal center was found for orientations I(A) and II(A): by fitting a parabola to a series of static calculations. Similar calculations were done for three orthogonal orientations at site V. The results of these calculations are shown in Table 4.



Site and orientation	Diagram	Distance from BDC center (Å)	Rotational angle (°)	Adsorption energy (meV)
IV(A)		3.50	N/A. Orthogonal to BDC plane.	-14.3
IV(B)		3.45	90	-11.0
IV(C)		3.46	60	-10.6
IV(D)		3.50	30	-9.9
IV(E)		3.52	0	-9.5
V(A)		4.85	N/A. Orthogonal to BDC plane.	-12.6

Site and orientation	Diagram	Distance from BDC center (Å)	Rotational angle (°)	Adsorption energy (meV)
V(B)		4.92	0	-10.2
V(C)		5.07	90	-5.2

**Table 4** Adsorption energies and at different orientations at sites IV and V.<sup>3</sup>

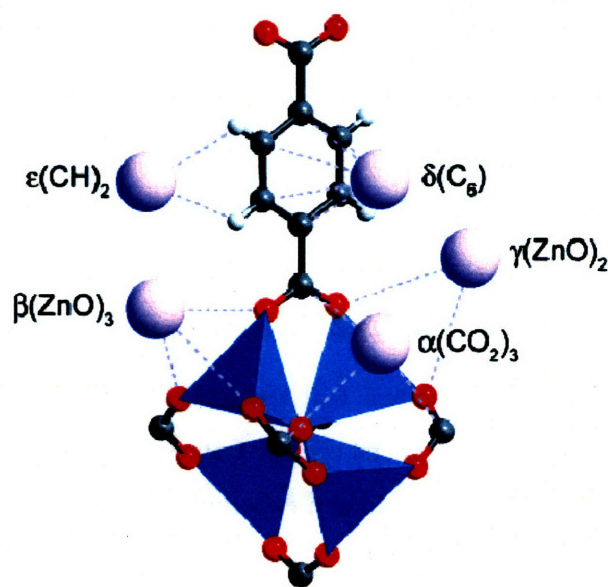
### II.2.3.1. Comparison with other studies

Hydrogen adsorption on the linker is weaker than adsorption near the  $Zn_4O$  cluster. This is consistent with the results of Sagara et al., in which second-order Møller Plesset theory (MP2) was used to compare interaction energies on benzene to interaction energies on  $Zn_4O(HCO_2)_6$  [44].

In an analysis of Inelastic Neutron Scattering (INS) data taken from hydrogen stored in MOF-5, Rosi et al. observe two distinct adsorption sites. They associate one of the sites with Zn, and the other with the BDC linker [26]. The results presented in this thesis suggest an alternative explanation: these could be two of the sites on the  $Zn_4O$  cluster. For example hydrogen might first adsorb at the sites of type I, and as the hydrogen density is increased the hydrogen coverage at sites of type II and III might increase. Any increase in hydrogen coverage at sites of type III would likely come at the expense of the coverage at sites of type I.

<sup>3</sup> The distance between the center of the BDC linker and the center of the  $H_2$  molecule is given along with the orientation of the  $H_2$  molecule. Unless otherwise noted, the  $H_2$  molecule is parallel to the plane of the BDC linker, and the angles given are measured from the axis running between the centers of the two carboxyl groups. The colors are the same as those used in Table 3.

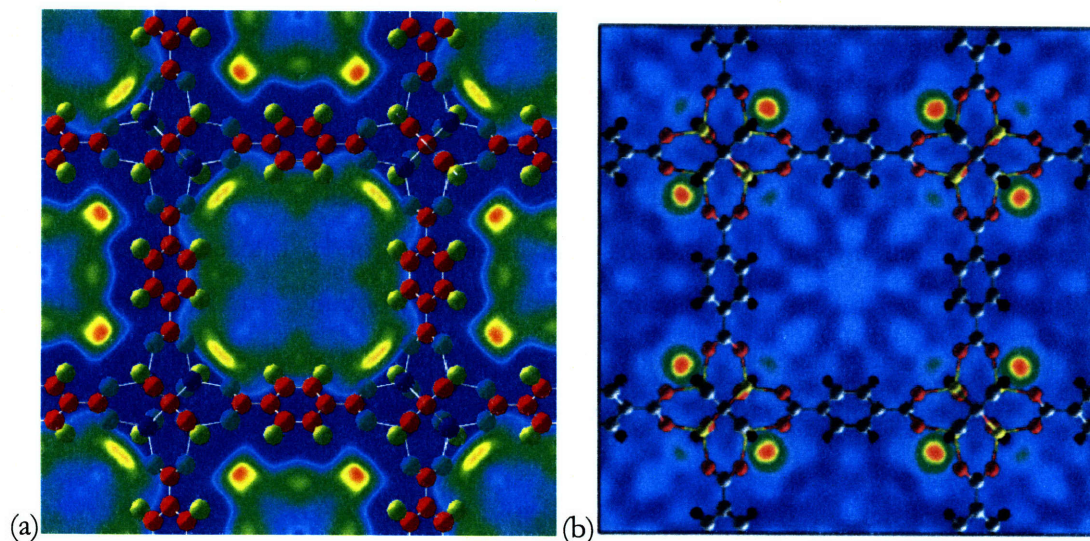
There have recently been several experimental studies that also support this result. Rowsell et al. used x-ray diffraction on argon adsorbed in MOF-5 to identify potential adsorption sites that are remarkably similar to the ones discovered in this thesis [45]. (Figure 6) The ranking of the adsorption energy for the sites was believed to be, from strongest to weakest,  $II \gg I > III > IV \approx V$ , which is similar to the ranking of  $I \approx III > II > IV \approx V$  calculated in this thesis. The notable difference is that Rowsell et al. suggested a stronger relative binding energy for Ar on site II than the energy for  $H_2$  calculated in this thesis. Initially, Rowsell et al. believed that hydrogen preferentially bound to site II, followed by site III due to its higher multiplicity. In a later paper, based on single-crystal Laue neutron diffraction of  $H_2$  adsorbed in MOF-5, this was revised to site II followed by site I [46].



**Figure 6** The five main argon adsorption sites identified by Rowsell et al. The sites correspond to the sites identified in this thesis in the following way:  $\alpha=II$ ,  $\beta=I$ ,  $\gamma=III$ ,  $\delta=IV$ ,  $\epsilon=V$ . Figure from reference [47].

Using neutron powder diffraction, Yildirim and Hartman were able to experimentally map the hydrogen density in MOF-5 [48]. They identified the same adsorption sites near the metal-oxide cluster and the organic linker, and a similar hydrogen density pattern to the one calculated in this thesis (Figure 7).





**Figure 7** (a) The hydrogen density in MOF-5 as calculated in this thesis plotted on a plane. (b) The hydrogen density calculated experimentally using neutron powder diffraction in reference [48]. In both plots, blue represents an area of low density and red is an area of high density, but the scales are not the same.

### II.2.3.2. Discussion

It has been experimentally shown that replacing the BDC linker with other organic linkers influences the ability of the metal organic frameworks to adsorb hydrogen [13]. This indicates that at least some of the hydrogen experimentally adsorbs on the linker. However none of the linkers tested caused the framework to store fewer than four molecules of hydrogen per formula unit at 1 atm and 77K, suggesting the possibility that four molecules of hydrogen adsorb on each of the  $Zn_4O$  clusters and the remaining hydrogen is associated with the organic linker.

The molecular dynamics simulation indicates an area of high hydrogen density near the center of pore surrounded by the edges of the BDC linkers (Figure 2). This may be a result of the pair correlation of molecular hydrogen. The first peak in the pair correlation function indicates the distance that most commonly separates two hydrogen molecules. As hydrogen adsorbs on the side of the pore, other hydrogen molecules in the pore will stay approximately this distance away from the adsorbed hydrogen. For the pore surrounded by

the edges of the BDC linker, the first peaks in the pair correlation functions from the adsorption sites on the edges of the pore overlap at the center of the cell, yielding the observed increase in hydrogen density there. For the larger pore, the peaks do not quite overlap the center of the cell, resulting in the observed ring of locally high hydrogen density. Similar results have been observed experimentally for argon adsorbed in MOF-5 [45].

Pores of the smaller type are surrounded by twelve sites of type V. When each of these sites is populated with hydrogen in orientation V(A), the energy for a hydrogen molecule in the center of the pore decreases by about 12 meV. The adsorption of hydrogen on the edges of the pores creates a new adsorption site near the center of the pore. This type of interaction suggests that under sufficiently high pressures hydrogen could condense to a liquid-like state in the pore.

The adsorption energies can be interpreted in the context of a simple thermodynamic model of adsorption in the pore. Under conditions of constant temperature and pressure, a hydrogen molecule can be expected to enter the framework when the chemical potential of the hydrogen in the pore is no greater than the chemical potential of free hydrogen gas. If it is assumed that the Gibb's free energy is a homogeneous function of the number of particles, the chemical potential of free hydrogen gas is given by the following expression:

$$\mu_{ext} = e_{ext} + \frac{P}{\rho_{ext}} - Ts_{ext} \quad \text{Equation 4}$$

where  $\mu$  represents the chemical potential,  $e$  represents energy,  $\rho$  represents density,  $T$  represents temperature, and  $s$  represents entropy. The subscript *ext* is used to describe values external to the framework, and lower-case letters are used to represent values that have been normalized by the number of molecules.

Likewise, the chemical potential of hydrogen inside the framework is given by the following expression:

$$\mu_{int} = e_{int} + \frac{P}{\rho} - Ts_{int} \quad \text{Equation 5}$$

The condition for adsorption of a gas to a given density within the framework is therefore:

$$e_{\text{int}} + \frac{P}{\rho_{\text{int}}} - Ts_{\text{int}} \leq e_{\text{ext}} + \frac{P}{\rho_{\text{ext}}} - Ts_{\text{ext}} \quad \text{Equation 6}$$

Equation 6 can be re-arranged to produce:

$$e_{\text{int}} - e_{\text{ext}} \leq P \left( \frac{1}{\rho_{\text{ext}}} - \frac{1}{\rho_{\text{int}}} \right) - T(s_{\text{ext}} - s_{\text{int}}) \quad \text{Equation 7}$$

The first term on the right can be thought of as the required work to compress the hydrogen to a given density at a given pressure. The remaining terms describe the change in Helmholtz free energy when the hydrogen molecules enter the framework. Adsorption will occur when the change in Helmholtz free energy offsets the work that must be done.

In this approximate analysis it will be assumed that the electronic and vibrational contributions to the difference in chemical potentials are small relative to the translational and rotational contributions. Under the classical assumption that the kinetic energy and potential energy can be separated, the change in the average kinetic energy upon adsorption is zero, leaving only the difference in average potential energy. For an ideal diatomic gas Equation 7 can thus be written [49]:

$$\Delta u \leq kT - \frac{P}{\rho_{\text{int}}} - \frac{3}{2}kT - kT \ln \left( \left( \frac{2\pi mkT}{h^2} \right)^{3/2} \frac{ekT}{P} \right) - Ts_{\text{rot}} + Ts_{\text{int}} \quad \text{Equation 8}$$

where  $\Delta u$  is the average potential energy of a molecule inside the pore relative to outside the pore,  $m$  is the mass of a hydrogen molecule,  $e$  is Euler's number, and  $s_{\text{rot}}$  is the nuclear and rotational contribution to the entropy for a free  $\text{H}_2$  rotor.

Every term on the right-hand side of Equation 8 is known with the exception of  $s_{\text{int}}$ . Two cases will be considered: one in which  $s_{\text{int}} = 0$ , the minimum possible value, and one in which  $s_{\text{int}}$  achieves its maximum possible value. The maximum value of  $s_{\text{int}}$  would occur if the potential energy surface of  $\text{H}_2$  in the framework were completely flat, and is given by the expression for the entropy of an ideal gas:

$$Ts_{\text{int,max}} = \frac{3}{2}kT + kT \ln \left( \left( \frac{2\pi mkT}{h^2} \right)^{3/2} \frac{e}{\rho_{\text{int}}} \right) + Ts_{\text{rot}} \quad \text{Equation 9}$$

Inserting the minimum and maximum entropy expressions into Equation 8 produces expressions for the upper and lower bound of the average adsorption energy inside the pore required for adsorption to occur:

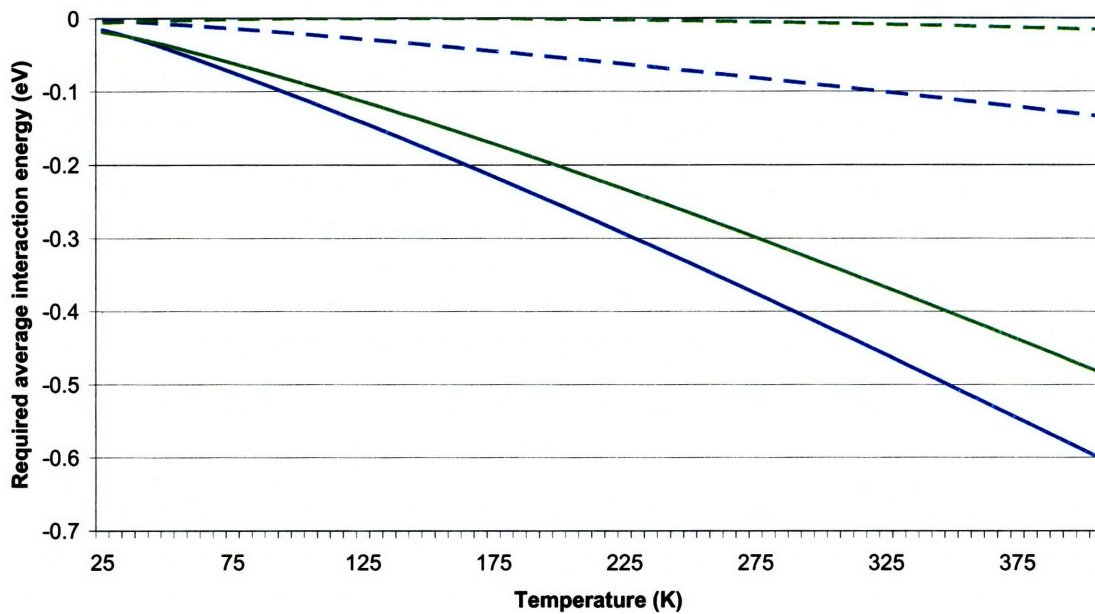
$$\Delta u_{\min} = -\frac{P}{\rho_{\text{int}}} - \frac{3}{2}kT - kT \ln \left( \left( \frac{2\pi mkT}{h^2} \right)^{3/2} \frac{kT}{P} \right) - Ts_{\text{rot}}$$

**Equation 10**

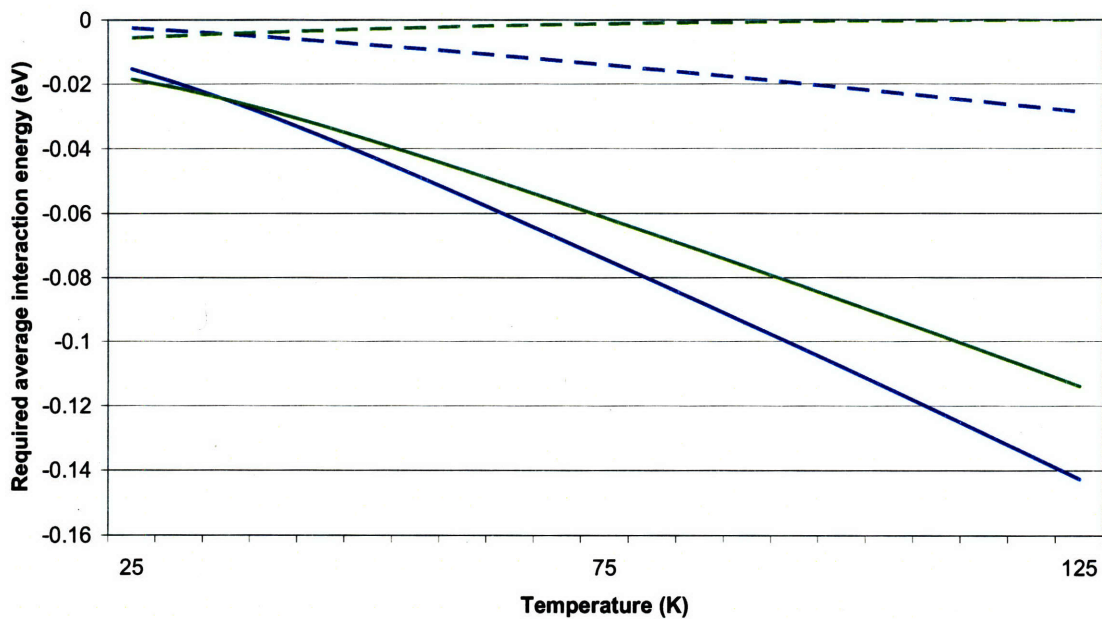
$$\Delta u_{\max} = kT - \frac{P}{\rho_{\text{int}}} - kT \ln \left( \frac{kT\rho_{\text{int}}}{P} \right)$$

The bounds in Equation 10 are plotted in Figure 8 and Figure 9 for both 1 atm of pressure and 40 atm. Figure 8 shows the bounds for the adsorption energy required to adsorb five molecules per formula unit ( $\sim 1.3$  wt%), which is approximately what is observed experimentally at 77 K and 1 atm [13]. It appears in Figure 8 that the maximum entropy bound increases with increasing temperature. This is due to the fact that at these temperatures and pressures, the natural density of free hydrogen gas is greater than five molecules per formula unit of MOF-5.

Because there are few sites where there is significant adsorption energy, it might be expected that the required adsorption energies are closer to the minimum-entropy bound (solid line) than the maximum-entropy bound (dashed line). Given the calculated adsorption energies, it may be surprising that any adsorption at all is experimentally observed at 77 K. There are a couple of explanations for this. The first is that the energies calculated by GGA may underestimate the true strength of the interactions. For example, GGA using the PW91 functional estimates the adsorption energy between molecular hydrogen and graphite to be about -21 meV [36], whereas experimental values range from -36 to -52 meV [50, 51]. It has also been suggested based on experiments that the hydrogen adsorption energy on site II in MOF-5 is in the range of approximately -40 to -52 meV [52, 53]. In addition the hydrogen, even when adsorbed, has not lost all translational entropy and likely moves between adsorption sites. This lowers the required interaction energy. For example, in a simple configurational model in which five hydrogen molecules populate sixteen equivalent and independent adsorption sites, at 77 K the entropic contribution to the free energy would be 11 meV per molecule in addition to the contribution of any entropy due to center-of-mass motion at each adsorption site.



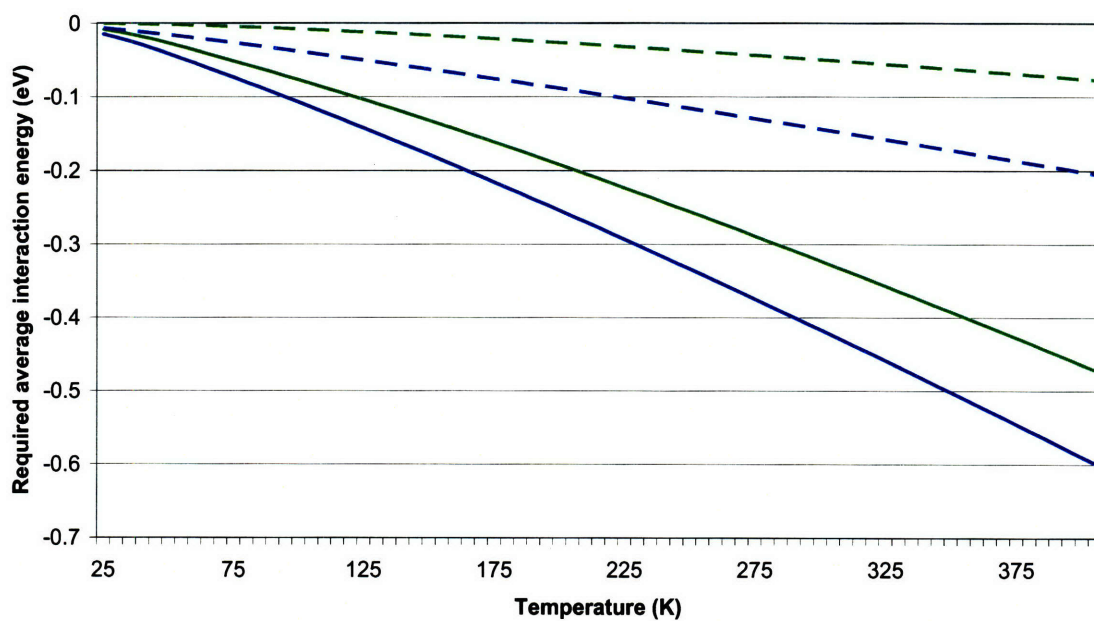
(a)



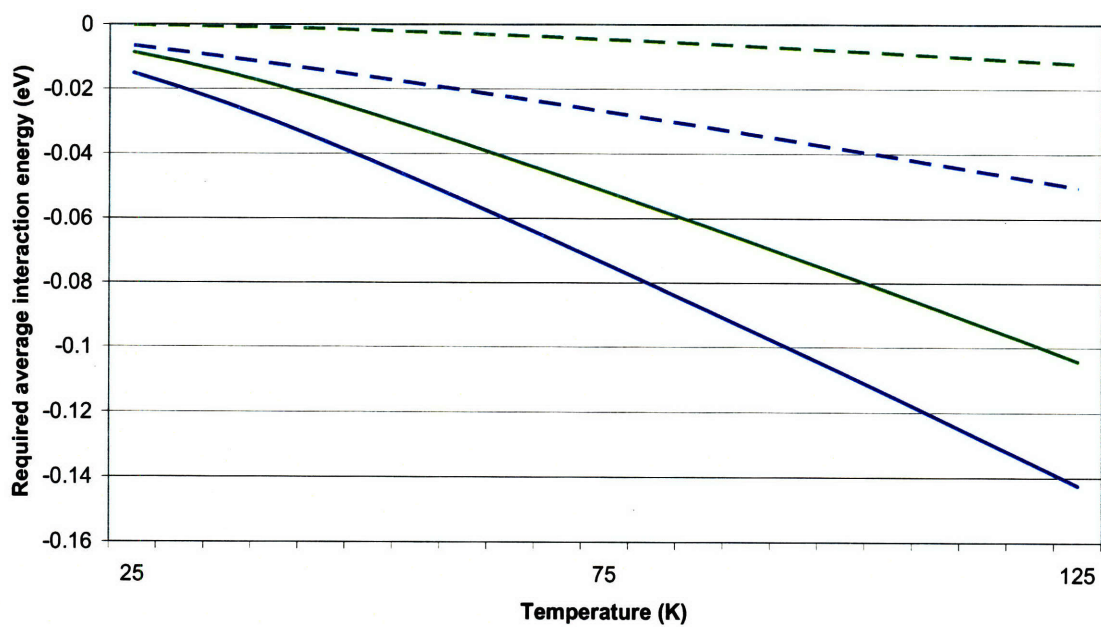
(b)

**Figure 8** Bounds for the required average adsorption energy per molecule for adsorption of 5 molecules per formula unit to occur at 1 atm (blue) and 40 atm (green). The bounds were generated under the minimum entropy assumption (solid) and the maximum entropy assumption (dashed). (a) From 25 to 400 K (b) From 25 to 125 K.





(a)



(b)

**Figure 9** Bounds for the required average interaction energy per molecule for adsorption of ten percent hydrogen by weight ( $\sim 38$  molecules per formula unit) to occur at 1 atm (blue) and 40 atm (green). The bounds were generated under the minimum entropy assumption (solid) and the maximum entropy assumption (dashed). (a) From 25 to 400 K (b) From 25 to 125 K.

A more realistic maximum entropy bound might be constructed if the density of hydrogen in the free volume of the framework is used instead of the density of hydrogen in the full framework. Experimentally, it is estimated that 61% of the framework is free volume [29]. In the molecular dynamics simulation  $H_2$  molecules passed within 1.5 Å of approximately 73% of the volume of the framework. If it is assumed that hydrogen is confined the free volume, the maximum entropy bound gets closer to the minimum entropy bound. For example, at 40 atm and 300 K, the average adsorption energy can be no greater than -49.7 meV for 10 wt% hydrogen to adsorb. If only 73% free volume is considered, the average adsorption energy must be lower than -57.3 meV.

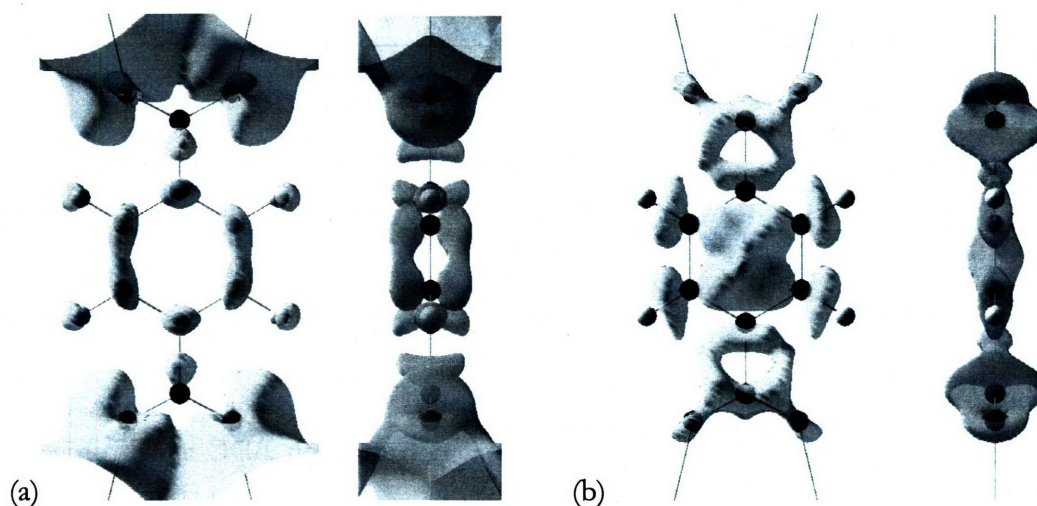
Figure 9 shows the bounds for the adsorption energy required to adsorb 10 wt% hydrogen, which is approximately the DOE material target for 2010 [6]. This density corresponds to approximately 38 hydrogen molecules per formula unit, which is greater than the number of adsorption sites identified. It is clear from Figure 9 that much lower adsorption energies would be required to adsorb 10 wt% hydrogen at 1 atm and 300 K. At high pressures and 77 K, it might be possible to store close to 10 wt% hydrogen in a microporous framework, which is supported by recent experiments [54]. However, the cost and extra weight of a storage vessel that can maintain the material at such lower temperatures and high pressures must be taken into account.

#### II.2.4. The effect of the framework on the BDC linker

It is possible to modify the nature of the pores in a metal-organic framework by changing the linker used to synthesize the framework. If it were possible to predict how these frameworks would interact with hydrogen by using what is known about the standalone linkers, it would facilitate the process of determining which linkers might form the most promising frameworks. In addition, the computational problem could be simplified to one of studying the just the linkers without the need to model the entire periodic framework. This would greatly reduce computational cost. The key is that the linker must have the same properties in the framework that it has outside the framework.

To investigate this, calculations were performed on the BDC linker without the surrounding framework. For the first calculation a linker was selected from the unit cell of the periodic system and all other atoms were removed. The coordinates of the atoms of the

linker were frozen and the electronic structure was calculated using exactly the same parameters used for the entire framework. The charge density of the standalone linker was then subtracted from the charge density of the full framework to determine how the framework influences the electronic structure of the linker. The results are shown in Figure 10. As might be expected, the  $Zn_4O$  clusters contribute charge, which resides principally around the oxygen atoms in the carboxyl groups of the linker. This causes a redistribution of charge over the rest of the linker from the carbon  $sp^2$  orbitals to the carbon  $p_x$  and hydrogen  $s$  orbitals.



**Figure 10** Isosurfaces of the change in the electronic density around the BDC linker when it is placed in the framework. (a) The charge added by the framework. The isosurface is drawn at  $0.0075 q_e / \text{\AA}^3$ . (b) The charge removed by the framework. The isosurface is drawn at  $0.001 q_e / \text{\AA}^3$ .

To test whether these small changes in the charge density have any meaningful effect on the interaction between BDC and the hydrogen molecule, the adsorption energies were calculated for hydrogen located at site IV for the standalone BDC linker in orientations IV(A), IV(B), and IV(E). (See Table 5) The adsorption energy changes by up to 6 meV, and the order of the preferred orientations is reversed. The reason for the change in the preferred orientation can be inferred from Figure 10. Placing the linker in the context of the framework increases the density of hydrogen above the carbon atoms and decreases the density near the center of the ring. This means there is less electronic repulsion for the

hydrogen molecule in orientation IV(A), but more in orientations IV(B) and IV(E). The adsorption energy for orientation IV(E) suffers the most because of the repulsion caused by the extra charge around the carboxyl groups. The change in adsorption energy is small in absolute terms but is significant relative to the already small adsorption energies.

H <sub>2</sub> Site and Orientation	IV(A)	IV(B)	IV(E)
Interaction energy in MOF-5 (meV)	-14.3	-11.0	-9.5
Adsorption energy for standalone BDC with coordinates frozen to match those of BDC in MOF-5 (meV)	-7.9	-11.9	-12.4

**Table 5** The adsorption energies for BDC in and out of the context of the framework.

If instead of freezing the coordinates of the atoms, the standalone BDC linker is allowed to relax, the angle of the O-C-O bond decreases from 126° to 113°. The coordinates of the carbon and hydrogen atoms change slightly, but remain within 0.05 Å of their frozen positions. This shift also influences the hydrogen interaction energy, changing it from 7.9 meV to 10.4 meV in orientation IV(A).

Thus the framework influences the interaction between BDC and hydrogen in two ways: by changing the electronic structure and by changing the physical structure of the linker. With these results in mind, it may be prudent to consider the interaction between molecular hydrogen and the linker in the context of the framework rather than in a standalone manner. If it is necessary to use a standalone molecule to represent the framework, care should be taken to ensure the molecule accurately reflects the relevant properties in the framework.

The way in which the interaction is changed also suggests how the adsorption energy of hydrogen on the linker can be lowered. The carboxyl groups tend to draw charge from the rest of the organic linker, reducing the strength of the hydrogen adsorption on the linker. Putting the linker in the framework counters this effect and increases the strength of the interaction between the linker and hydrogen. The effect of the carboxyl groups may be further reduced if a larger linker is used, or if the linker is doped with an element that injects

electron density into the linker. Hübner et al. have studied the effects of modifying benzene through substitution and arrive at a similar conclusion [55].

## II.2.5. Summary and Conclusion

---

Density functional theory has been used to investigate the adsorption of hydrogen in the metal-organic framework known as MOF-5. A molecular dynamics simulation indicates that there are five distinct adsorption sites on the edges of the pores. Hydrogen adsorbs most strongly at the three sites near the  $Zn_4O$  cluster, and least strongly at the sites near the BDC linker, indicating that observed hydrogen adsorption might occur near the metal-oxide cluster rather than on the organic linker. Although there are a total of twenty adsorption sites on each  $Zn_4O$  cluster, the ability of these clusters to adsorb hydrogen falls off sharply after sixteen of the sites are populated.

The way in which placing the BDC linker in the framework changes the electronic and physical properties of the linker and how these changes affect the interaction between the linker and molecular hydrogen have also been investigated. The framework contributes charge to the linker, increasing the interaction between hydrogen and the linker and changing the preferred orientation of hydrogen adsorbed above the aromatic ring.

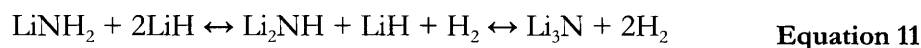
The calculated interaction energies are weak, but this may be due to the fact that GGA is often not able to calculate the exact magnitude of weak interactions accurately. Nonetheless the calculations indicate that there would likely need to be a significant strengthening of the interaction between molecular hydrogen and the pores of the framework for a metal-organic framework to be a viable material for hydrogen storage. It may be possible to do this by using intelligently functionalized linkers or by enhancing the interaction energy of hydrogen on the metal-based clusters that connect the linkers [26]. Despite numerous attempts to design metal-organic frameworks with enhanced hydrogen storage capacity, gravimetric capacities remain below 3% at 1 atm [15, 54]. Gravimetric capacities of up to 7.5% have been reported for frameworks at pressures above 70 bar [54], but the need for a high-pressure storage container must be taken into account when considering such a system for on-board hydrogen storage.

---

## II.3. Lithium imide

---

Chen et al. have observed that the combination of lithium amide ( $\text{LiNH}_2$ ) and lithium hydride ( $\text{LiH}$ ) releases hydrogen in a two-step reaction [20]:



The first step in this reaction, in which lithium amide reacts with lithium hydride to form lithium imide ( $\text{Li}_2\text{NH}$ ), is of interest as a hydrogen storage system because it releases up to 6.5 wt% hydrogen and is reversible at temperatures below 300°C [20]. Lithium imide is also believed to be a reactant in related hydrogen storage systems, such as Li-Mg-N-H [56] and Li-B-N-H [57]. The structures of lithium amide and lithium hydride are well characterized, but despite the recent high levels of interest in lithium imide a consensus on a complete specification of its structure remains lacking.

Using X-ray diffraction, Juza and Opp concluded that lithium imide is most likely anti-fluorite with lithium cations and nitrogen anions, but they were unable to resolve the positions of the hydrogen ions [58]. Several studies since have indicated that each hydrogen nucleus is bonded to a nitrogen nucleus to form an imide (N-H) anion [59-64] though the orientation of these N-H groups is unknown. Using neutron powder diffraction (NPD), Ohoyama et al. proposed a structure in which hydrogen randomly occupies one of four symmetrically equivalent sites around the nitrogen ion [59]. Based on synchrotron X-ray diffraction results, Noritake et al. supported another structure, in which the hydrogen randomly occupies one of twelve sites around the nitrogen ion [60]. Zhang et al. treated the hydrogen nucleus as a quantum mechanical particle and found that when nitrogen and lithium ions are fixed at their anti-fluorite positions the wavefunction of a hydrogen nucleus is centered at the nitrogen nucleus with density maxima along the  $\langle 100 \rangle$  directions [62].

Herbst and Hector, based on experimental work by Balogh et al. [63] and density functional theory [65] (DFT) calculations, have proposed a fully occupied low-symmetry structure in which one in eight lithium ions have moved to the empty octahedral sites between nitrogen ions [61]. They calculated the enthalpy of formation of their proposed structure to be approximately 17 to 38 kJ per mol formula unit (f. u.) higher than the experimentally derived enthalpy of formation for lithium imide, as opposed to a difference

of 3 kJ per mol f.u. for  $\text{LiNH}_2$  [61]. This suggests that although the structure they propose produces diffraction patterns similar to those observed experimentally it may not be the energetic ground state. Recently Magyari-Köpe et al. discovered several structures with significantly lower calculated energies than any previously suggested structure [64].

It is worth considering whether lithium imide might be thermally disordered at room temperature. Ohoyama et al., find no evidence of a structural phase transition between 10 K and room temperature [59], suggesting that if the structure is disordered at room temperature the disorder persists to a very low temperature. A differential thermal analysis (DTA) study of lithium imide indicates that there is a second-order phase transition at about 356 K, which has been attributed to an order-disorder transition [66]. Balogh et al. have found a similar result [63]. These results suggest that the hydrogen positions in lithium imide are fully or partially ordered at room temperature.

Kojima et al. have shown that lithium imide shows two broad N-H stretching peaks in the infrared, unlike lithium amide for which the absorption peaks are sharp [67]. Although the authors suggest that the peak broadening could be due to small crystallite size or thermal disorder, their results might also indicate a wide variety of local environments for N-H groups, implying either partial disorder or a large unit cell for the imide.

### II.3.1. An antifluorite model of lithium imide

---

There is a widely shared belief that the average positions of the Li and N atoms form an antifluorite structure, but very different assignments of the H positions have been proposed [59-64]. The distinction between structural models is significant, as *ab initio* calculations indicate that the energy of the imide depends rather strongly on the H positions [61, 63, 64]. Hence a good structural model and insight into the factors that determine the arrangement of hydrogen nuclei in these imides is required. In this section a model to obtain the *effective* H-H interaction in Li-imide is developed and used to obtain a structure with lower energy than any previously proposed structure.

Under the assumption that lithium imide has an antifluorite-like structure, the challenge in determining the ground state structure is primarily one of determining the lowest-energy orientations for the imide groups. The N-H bond length varies little; its average length in 99 relaxed structures calculated for this section is 1.038 Å and the standard deviation is 0.003 Å.

However the orientation of one N-H bond substantially affects the orientations of N-H bonds in neighboring cells. To perform a thorough search of possible ground states and to clarify the interaction between N-H bond orientations, an effective Hamiltonian has been developed for N-H bond orientations in lithium imide. The objective is to write the energy of the system directly in terms of the orientation of the N-H groups with the other degrees of freedom implicitly relaxed to local minima. The model is parameterized with *ab initio* calculations and used to predict a new structure with lower calculated energy than any structure known to date.

### II.3.1.1. Methodology

#### II.3.1.1.1. The effective Hamiltonian

The effective Hamiltonian for the N-H orientations is based on a modification of the cluster expansion [68, 69] formalism to include continuous variables describing the imide group orientation. The N-H bond orientation on the *ith* imide group can be characterized with a polar and azimuthal coordinate,  $(\theta_i, \phi_i)$ . The objective of the model is to find an expression for the function

$$F(\theta_1, \phi_1, \theta_2, \phi_2, \theta_3, \phi_3, \dots) \quad \text{Equation 12}$$

where  $F$  is the energy of the system with the N-H groups in the specified orientations and all other coordinates (e.g. the N-H bond length and the Li positions) relaxed. This coarse-grained energy function is similar to the cluster expansion for configurational disorder in alloy theory where displacements (and sometimes vibrations) are coarse-grained over to obtain an energy expression solely in terms of site occupation variables [68-70].

For each domain  $(\theta_i, \phi_i)$  a local basis of functions  $f_{n_i}(\theta_i, \phi_i)$  is defined, where  $f_{n_i}$  indicates the *n*th basis function for the *ith* imide group. The tensor product of these local bases forms a basis for the function  $F$ :

$$F(\theta_1, \phi_1, \dots) = \sum_{\vec{n}} \left( V_{\vec{n}} \prod_i f_{n_i}(\theta_i, \phi_i) \right) \quad \text{Equation 13}$$

where  $\vec{n}$  is a vector of basis function indices in which the *ith* element is  $n_i$ . The sum is over all possible  $\vec{n}$ , and  $V_{\vec{n}}$  are coefficients to be determined later.



For example, if  $F$  is square integrable over imide group orientations,  $f_{n_i}$  can be chosen as spherical harmonics. As with a discrete cluster expansion, it is convenient to choose  $f_{n_i}$  such that  $f_{0_i}$  is a constant for all  $i$ . The function  $F$  can then be expanded as a linear combination of functions of the orientations of single imide groups, functions of the orientations of pairs of imide groups, etc.:

$$F(\theta_1, \phi_1, \dots) = V_0 + \sum_i \sum_{n_i \neq 0} V_{n_i} f_{n_i}(\theta_i, \phi_i) + \sum_{i,j} \sum_{n_i \neq 0, n_j \neq 0} V_{n_i, n_j} f_{n_i}(\theta_i, \phi_i) f_{n_j}(\theta_j, \phi_j) + \dots \quad \text{Equation 14}$$

The symmetry properties of the crystal can be used to group terms that share the same coefficient. Hence, it is more convenient to choose basis functions  $f_{n_i}$  that form closed orbits under the operations in the space group of the crystal. One result of this step is that terms that are anti-symmetric with respect to the symmetry operations of the crystal disappear.

The expansion in Equation 14 is in principle an exact representation of Equation 12, but to reduce this to a finite problem it is necessary to make approximations based on physical intuition. As with a discrete cluster expansion, it is assumed that Equation 14 will be dominated by the constant term and terms representing the interactions between physically small clusters of imide groups. In this study, only terms representing single imide groups and imide pair interactions up to the next-next-nearest neighbor are included.

For continuous domains there are an infinite number of basis functions  $f_{n_i}$  in the complete basis. In practice only the basis functions that should have the most physical relevance are included. For example, for a basis of spherical harmonics one might truncate the basis at a certain angular momentum. For lithium imide a truncated basis from hybridized spherical harmonics has been generated:

$$\begin{aligned}
f_{0_i} &= 1, \\
f_{1_i} &= \frac{1}{2}\sqrt{30}x_i y_i, \\
f_{2_i} &= \frac{1}{2}\sqrt{210}x_i y_i z_i, \\
f_{3_i} &= \frac{1}{4}\left(\sqrt{175}x_i^3 - \sqrt{63}x_i + \sqrt{105}x_i(y_i^2 - z_i^2)\right), \\
f_{4_i} &= \frac{5}{8}\sqrt{7}(x_i^4 + y_i^4 + z_i^4) + \frac{1}{2}\sqrt{\frac{5}{6}}(3x_i^2 - 1) + \frac{1}{2}\sqrt{3}x_i - \frac{3}{8}\sqrt{7},
\end{aligned}
\tag{Equation 15}$$

where  $x_i = \cos\theta_i \sin\phi_i$ ,  $y_i = \sin\theta_i \sin\phi_i$ , and  $z_i = \cos\phi_i$ . The full truncated basis used can be generated by applying the cubic symmetry operations to these functions and keeping the ones that are linearly independent.

To determine the coefficients  $V_n$  of this expansion a library of 98 relaxed structures with different relative N-H orientations was calculated using the projector augmented wavefunction [40] (PAW) method with the Perdew-Burke-Ernzerhof [39] (PBE) generalized gradient approximation (GGA) to DFT as implemented in the Vienna *Ab-initio* Simulation Package (VASP) [71]. The standard hydrogen and nitrogen PAW potentials and *s*-valence Li PAW potentials in VASP were used with a plane-wave cutoff energy of 520 eV. Calculations were considered converged when forces reached less than 80 meV / Å. For a primitive cell calculation, total energy convergence within 1 meV per formula unit was reached with a 7x7x7 Monkhorst-Pack [72]  $k$ -point grid shifted to include the gamma point. For supercell calculations this grid was scaled down proportionately to the size of the supercell.

The coefficients were evaluated using a least-squares regression, where structures were given weights defined by

$$w_i \propto \frac{1}{\Delta E_i^2 + .0009} \tag{Equation 16}$$

where  $\Delta E_i$  is the difference in eV between the calculated formation energy of the *i*th structure and the lowest calculated formation energy. To help prevent overfitting the coefficients were fit to both the energies in eV per formula unit and the forces in eV / Å, with the energies given ten times the weight of the forces. To more accurately represent low-energy interactions, structures with lower energy were assigned higher weights. To ensure the cluster expansion had predictive power a Metropolis algorithm [73] was used to

find a subset of terms in the expansion that had a low leave-one-out cross-validation score [74]. The coefficients for this model generated from the original library of 98 structures and the newfound orthorhombic structure are given in Table 6.

Cluster	Site 1 function	Site 2 function	Terms per formula unit	$V_n$ (kJ / mol f.u.)	$V_n$ (meV / mol f.u.)
Empty	1	1	1	-179.73607	-1862.78463
Point	$f_4(-x)$	1	6	0.91880	9.52246
Nearest Neighbor	$f_1(x,y)$	$f_1(x,y)$	6	0.38115	3.95020
	$f_1(x,y)$	$f_4(x)$	24	0.04595	0.47620
	$f_1(x,y)$	$f_3(x,z,y)$	24	0.01690	0.17514
	$f_1(x,z)$	$f_1(y,z)$	12	0.11320	1.17316
	$f_1(x,z)$	$f_4(-z)$	48	-0.06256	-0.64833
	$f_2(x,y,z)$	$f_2(x,y,z)$	6	0.31453	3.25979
	$f_2(x,y,z)$	$f_4(-z)$	24	-0.08269	-0.85699
	$f_2(x,y,z)$	$f_3(z,y,x)$	24	-0.05846	-0.60589
	$f_4(-x)$	$f_4(-x)$	24	-0.07591	-0.78672
	$f_4(-x)$	$f_4(-y)$	24	0.18900	1.95885
	$f_4(-x)$	$f_3(y,x,z)$	24	0.06545	0.67832
	$f_4(-z)$	$f_4(-z)$	12	0.33858	3.50904
	$f_4(-z)$	$f_4(z)$	12	-0.35140	-3.64189
	$f_4(-z)$	$f_3(x,y,z)$	48	-0.03802	-0.39408
	$f_4(-z)$	$f_3(z,y,x)$	48	-0.11254	-1.16640
	$f_4(x)$	$f_3(y,x,z)$	24	-0.09816	-1.01732
	$f_3(x,z,y)$	$f_3(x,z,y)$	12	0.16971	1.75887
	$f_3(x,z,y)$	$f_3(y,x,z)$	24	-0.05320	-0.55134
	$f_3(x,y,z)$	$f_3(x,y,z)$	12	-0.09459	-0.98028
	Next-Nearest Neighbor	$f_1(x,y)$	$f_1(x,y)$	6	-0.17086
$f_1(x,y)$		$f_4(-y)$	24	-0.04874	-0.50514
$f_1(x,y)$		$f_3(y,x,z)$	12	-0.05271	-0.54629
$f_1(y,z)$		$f_1(y,z)$	3	0.24773	2.56743
$f_1(y,z)$		$f_2(x,y,z)$	6	0.07267	0.75319
$f_4(-y)$		$f_4(-x)$	24	0.21403	2.21818
$f_4(-y)$		$f_4(-y)$	12	0.16783	1.73940
$f_4(-y)$		$f_3(x,y,z)$	24	0.11144	1.15501
	$f_3(x,z,y)$	$f_3(x,z,y)$	6	0.06158	0.63821

Cluster	Site 1 function	Site 2 function	Terms per formula unit	$V_n$ (kJ / mol f.u.)	$V_n$ (meV / mol f.u.)
	$f_3(y,z,x)$	$f_3(y,z,x)$	6	-0.13240	-1.37220
Next- Next- Nearest Neighbors	$f_1(x,y)$	$f_2(x,y,z)$	48	-0.01022	-0.10591
	$f_1(x,y)$	$f_4(-x)$	48	-0.06696	-0.69393
	$f_1(x,y)$	$f_4(-z)$	48	-0.02955	-0.30624
	$f_1(x,y)$	$f_4(z)$	48	-0.02770	-0.28705
	$f_1(x,y)$	$f_3(x,z,y)$	48	-0.03402	-0.35262
	$f_1(x,y)$	$f_3(y,z,x)$	48	-0.03412	-0.35366
	$f_1(x,y)$	$f_3(z,x,y)$	48	0.03978	0.41224
	$f_1(y,z)$	$f_1(y,z)$	12	-0.10254	-1.06277
	$f_1(y,z)$	$f_4(x)$	24	0.02594	0.26880
	$f_2(x,y,z)$	$f_3(y,z,x)$	48	-0.03278	-0.33972
	$f_4(-x)$	$f_4(-y)$	48	-0.05105	-0.52904
	$f_4(-x)$	$f_3(y,z,x)$	48	-0.02563	-0.26559
	$f_4(-y)$	$f_4(-x)$	48	0.06118	0.63411
	$f_4(-y)$	$f_4(y)$	24	-0.23445	-2.42980
	$f_4(-y)$	$f_3(x,y,z)$	48	0.06373	0.66049
	$f_4(-y)$	$f_3(y,x,z)$	48	0.05022	0.52049
	$f_4(x)$	$f_4(-x)$	12	-0.07956	-0.82452
	$f_4(x)$	$f_3(y,x,z)$	48	0.02032	0.21061
	$f_4(x)$	$f_3(y,z,x)$	48	0.04955	0.51353
	$f_4(y)$	$f_4(-z)$	24	0.10778	1.11707
	$f_4(y)$	$f_3(x,z,y)$	48	0.02108	0.21842
	$f_4(y)$	$f_3(z,y,x)$	48	-0.06593	-0.68325
	$f_4(y)$	$f_3(z,x,y)$	48	-0.02153	-0.22311
	$f_3(x,z,y)$	$f_3(x,y,z)$	24	0.03629	0.37610
	$f_3(x,z,y)$	$f_3(y,x,z)$	48	0.01857	0.19244
	$f_3(x,z,y)$	$f_3(y,z,x)$	48	-0.03569	-0.36987
	$f_3(y,x,z)$	$f_3(z,y,x)$	48	-0.03856	-0.39968
	$f_3(y,x,z)$	$f_3(z,x,y)$	24	0.05823	0.60350
$f_3(y,z,x)$	$f_3(z,y,x)$	24	-0.02714	-0.28129	

**Table 6** The parameters for the model Hamiltonian. Each row lists a sample term from an included orbit of functions and the coefficient for functions in that orbit. The imide groups for the sample functions, in conventional anti-fluorite reduced coordinates, are located at (0, 0, 0) for site 1 and (0.5, 0.5, 0) for site 2 for nearest neighbors, (0,0,0) and (1, 0, 0) for next-nearest neighbors, and (0, 0, 0) and (1, 0.5, 0.5) for next-next-nearest neighbors. The

functions are:  $f_1(a, b) = \frac{1}{2}\sqrt{30ab}$ ,  $f_2(a, b, c) = \frac{1}{2}\sqrt{210abc}$ ,  
 $f_3(a, b, c) = \frac{1}{4}\left(\sqrt{175a^3} - \sqrt{63a} + \sqrt{105a(b^2 - c^2)}\right)$ , and  
 $f_4(a) = \frac{5}{8}\sqrt{7}(x^4 + y^4 + z^4) + \frac{1}{2}\sqrt{\frac{5}{6}}(3a^2 - 1) + \frac{1}{2}\sqrt{3a} - \frac{3}{8}\sqrt{7}$ . All functions symmetrically  
equivalent to those listed were included in the Hamiltonian.

### II.3.1.1.2. Enthalpies of formation

The enthalpies of formation of several low-energy structures were calculated. To ensure accurate results the electronic energies were recalculated using hard hydrogen and nitrogen PAW potentials and the cutoff energy was increased to 900 eV. Calculations were considered converged when the forces on the ions were less than 1 meV / Å. Vibrations within the harmonic approximation were evaluated within the linear response approach as implemented in ABINIT [75] with the PBE GGA exchange-correlation functional. The Fritz-Haber Institute pseudopotentials provided with ABINIT were used, and the cutoff energy was 35 Ry. For the electronic supercell calculations a 4x4x4 k-point grid including the gamma point was used, and interatomic forces were calculated on a 2x2x2 grid. For the layered structure a 6x6x6 electronic k-point grid and 3x3x3 interatomic force grid were used. Ideal gas behavior was assumed for the standard state of the reference molecules H<sub>2</sub> and N<sub>2</sub>. The values for H<sub>2</sub> and N<sub>2</sub> reference molecules were calculated in cubic cells with length 15 Å. It was found that increasing the size of the cell to 17 Å changed the calculated energies by less than 0.01 meV.

## II.3.1.2. Results

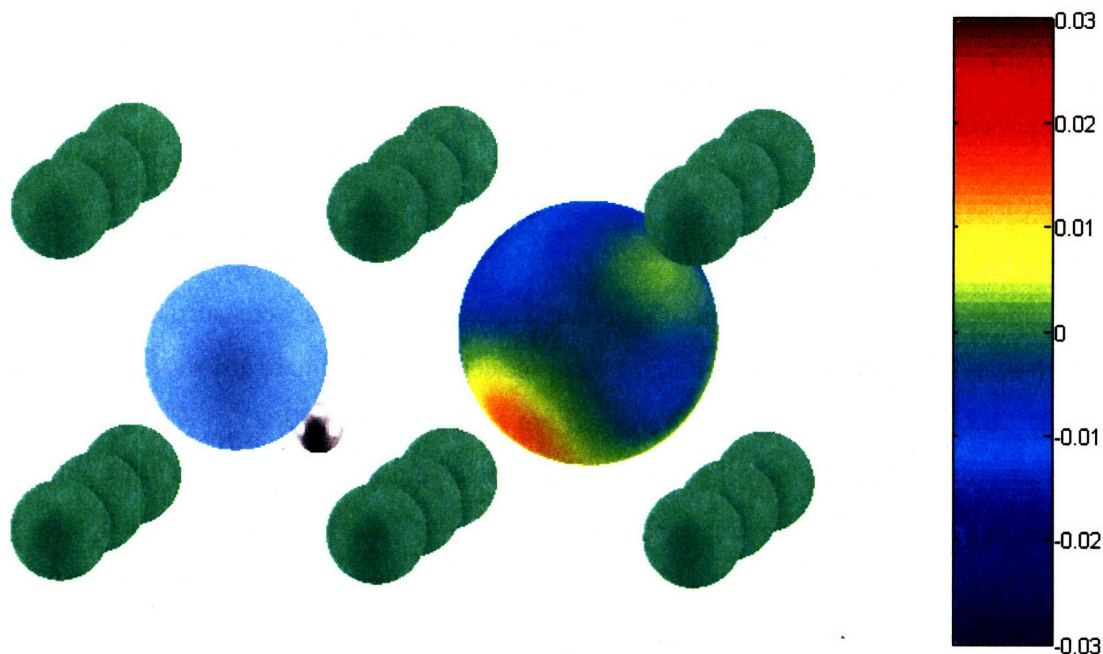
### II.3.1.2.1. Effective interactions

Within the model Hamiltonian, the interaction between a single pair of nearest-neighbor imide groups is represented by

$$\sum_{n_1 \neq 0} V_{n_1} f_{n_1}(\theta_1, \phi_1) + \sum_{n_2 \neq 0} V_{n_2} f_{n_2}(\theta_2, \phi_2) + \sum_{n_1 \neq 0, n_2 \neq 0} V_{n_1, n_2} f_{n_1}(\theta_1, \phi_1) f_{n_2}(\theta_2, \phi_2) \quad \text{Equation 17}$$

where the contributions from single-imide terms have been included as well. A sample nearest-neighbor interaction can be visualized in Figure 11. This figure clearly shows that if

the hydrogen nuclei from two nearest-neighbor imide groups are near the same tetrahedral site, the interaction is unfavorable. On the other hand, an anti-parallel alignment between two nearest-neighbor imide groups is favorable.

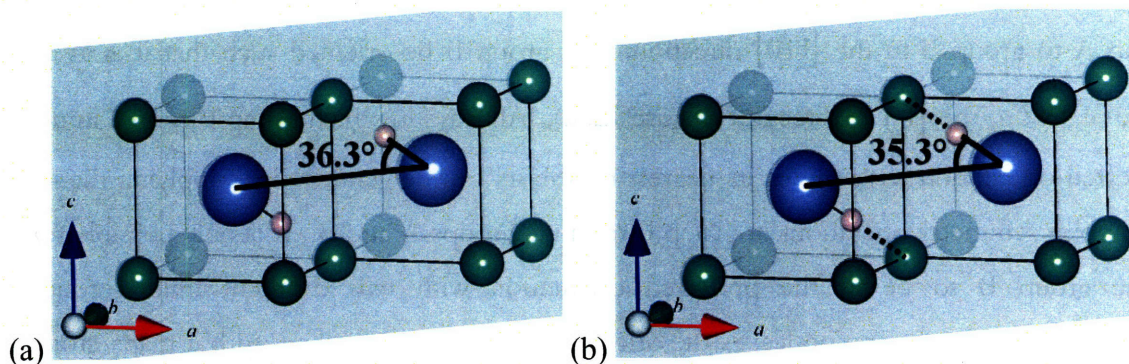


**Figure 11** An example of an effective nearest-neighbor interaction in the model Hamiltonian. The green balls represent lithium nuclei at ideal tetrahedral sites, and they form a cage of four octahedral sites. The blue ball on the left represents a nitrogen nucleus in the front left octahedral site, and the white ball is a hydrogen nucleus that is bound to it to form an imide group. The multi-colored sphere represents a nearest-neighbor imide group in the back right octahedral site. The colors on the sphere represent the effective interaction energy between the two imide groups as a function of the orientation of the N-H bond in the imide group on the right.

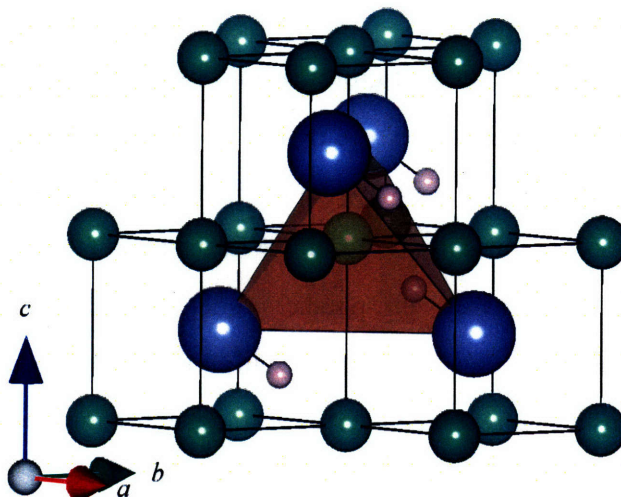
The minimum of Equation 17 represents the most favorable nearest-neighbor imide group orientations and can be visualized in Figure 12(a). This interaction will be referred to as the “preferred” nearest neighbor orientation. Similar to the observation made in Ref. [64], it is found that neighboring imide groups tend to align anti-parallel to one another. The nitrogen and hydrogen nuclei are coplanar with the two lithium nuclei between the imide groups in the  $(1\bar{1}0)$  plane, and a N-N-H bond angle of  $36.3^\circ$ . The hydrogen nucleus on each group tilts towards the nitrogen of the other group, possibly to form a hydrogen bond. In practice the apparent unfavorable proximity between the hydrogen ions and the lithium



ions between groups is resolved as these lithium ions move towards nearby empty octahedral sites. Because this leaves behind a vacancy at a tetrahedral site, it results in relaxation *towards* octahedral sites unlike the full displacement suggested in Ref. [61].



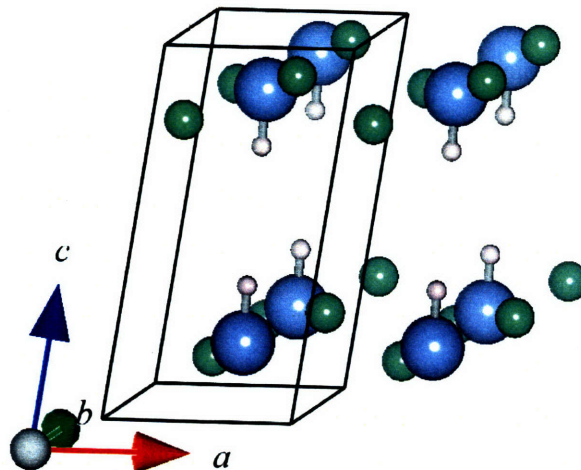
**Figure 12** (a) The “preferred” nearest neighbor orientation. The lithium ions are shown in the ideal anti-fluorite positions for reference. Both N-H bonds are in the  $(1\bar{1}0)$  plane, which is shown. (b) The more high-symmetry “near-preferred” nearest neighbor orientation in which the N-H bonds are aligned along  $[\bar{1}\bar{1}1]$  in the  $(1\bar{1}0)$  plane. Dotted lines indicate the lithium ions to which the N-H bonds point. Large spheres represent nitrogen, medium represent lithium, and small represent hydrogen.



**Figure 13** A tetrahedron of nearest-neighbor imide groups. The lower right group is in the near-preferred nearest neighbor orientation with each of the other three groups, but as a result these groups are not in the near-preferred orientation with each other. Large spheres represent nitrogen, medium represent lithium, and small represent hydrogen.

It is impossible for a single imide group to be in the preferred orientation with more than one of its nearest neighbors. To see this, let one imide group, “group A”, be in the preferred orientation with a nearest neighbor imide group, “group B”, with the vector from group A to group B in the  $[110]$  direction. Let group B be oriented such that it is at an angle of  $1.1^\circ$  to the  $[\bar{1}\bar{1}1]$  direction (Figure 12(a)). All ways in which group B can be in the preferred orientation with any of its nearest neighbors can be generated by applying the 48 FCC point symmetry operations at group B to imide groups A and B. It is only possible for imide group B to be in the preferred orientation with two different imide groups simultaneously if an operation leaves the orientation of group B unchanged but maps group A onto a different nearest neighbor. The only operation that results in group B having the exact same orientation is reflection about the  $(1\bar{1}0)$  plane, but this maps group A onto itself. This frustration can be partially resolved if the N-H bonds rotate slightly so that they are aligned in the high-symmetry  $[\bar{1}\bar{1}1]$  direction (Figure 12(b)). This is still a low-energy orientation because of the ability for the lithium ions to relax towards the empty octahedral sites. This will be referred to as the near-preferred orientation. If groups A and B are in the near-preferred orientation, the symmetry operations corresponding to  $120^\circ$  rotations about the  $[\bar{1}\bar{1}1]$  axis all map the orientation of group B onto itself, but maps group A onto three different nearest neighbors. This allows group B to be in the near-preferred orientation relative to three nearest neighbors at the same time (Figure 13). For every imide group to be in the near-preferred orientation relative to three nearest neighbors, the structure must consist of  $\{111\}$  planes of alternating anti-parallel imide groups where the imide groups in each plane are aligned orthogonally to the plane. This structure, after relaxation with DFT, becomes the one shown in Figure 14. Nuclear coordinates for the relaxed structure, which will be referred to as the “layered” structure, are given in Table 7.





**Figure 14** A 2x2x1 supercell of the layered structure with the unit cell highlighted. Large spheres represent nitrogen, medium represent lithium, and small represent hydrogen.

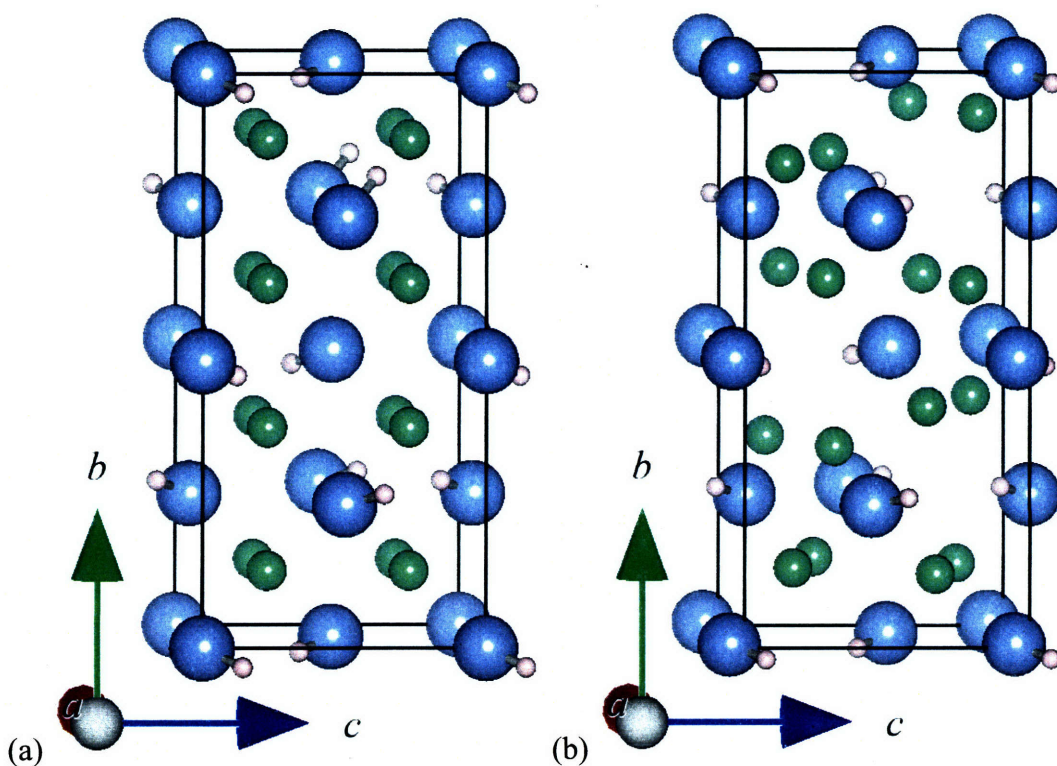
Element	X	Y	Z
Lithium	0.980	0.405	0.258
Lithium	0.674	0.202	0.905
Nitrogen	0.653	0.090	0.203
Hydrogen	0.364	0.956	0.640

**Table 7** The reduced coordinates of the layered structure. All atoms are at Wyckoff position  $2i$  for space group #2 (P-1).  $a=3.57 \text{ \AA}$ ,  $b=3.58 \text{ \AA}$ ,  $c=6.80 \text{ \AA}$ ,  $\alpha = 77.83^\circ$ ,  $\beta=82.23^\circ$ , and  $\gamma=59.91^\circ$ .

### II.3.1.2.2. A new low-energy structure

In addition to providing insight into the effective N-H interactions, the cluster expansion can also be used to very rapidly search for low energy structures of lithium imide. A search for the ground state of this model was performed with Monte Carlo simulation in all supercells up to 8 formula units. The lowest energy configuration is shown in Figure 15(a). A DFT calculation relaxes this structure to the one shown in Figure 15(b), which will be referred to as the “orthorhombic” structure. This structure, although not in the initial library of structures used to fit the coefficients, was added when determining the preferred orientations. The lithium ions in the relaxed structure have moved significantly from their

ideal anti-fluorite positions. The new structure has an orthorhombic unit cell, with coordinates given in Table 8 and calculated lattice parameters of 5.12, 10.51, and 5.27 Angstroms, although these are probably overestimated by a few percent as is typical in the GGA approximation [76-78]. The space group of the structure as determined by ABINIT is  $Pbca$  (#61). Every imide group in this structure is involved in a nearest-neighbor pair interaction resembling the preferred nearest-neighbor orientation, although the imide groups are rotated  $25.4^\circ$  from the preferred nearest-neighbor orientations to be in the low-symmetry  $[0.79, 0.58, 0.20]$  and equivalent directions. The existence of nearest-neighbor orientations similar to the preferred orientation suggests that preference for this type of nearest-neighbor orientation remains high even in an infinite crystal.



**Figure 15** (a) The imide group orientations with the lowest predicted energy after searching structures with up to 8 formula units per unit cell. Lithium ions are shown in the ideal anti-fluorite sites for reference. (b) The structure after relaxation with density functional theory. Large spheres represent nitrogen, medium represent lithium, and small represent hydrogen.

Element	X	Y	Z
Lithium	0.372	0.540	0.651
Lithium	0.480	0.256	0.191
Nitrogen	0.297	0.370	0.451
Hydrogen	0.363	0.610	0.066

**Table 8** The reduced coordinates of the orthorhombic structure. All atoms are at Wyckoff position 8c for space group #61 (Pbca).  $a=5.12 \text{ \AA}$ ,  $b=10.51 \text{ \AA}$ , and  $c=5.27 \text{ \AA}$ .

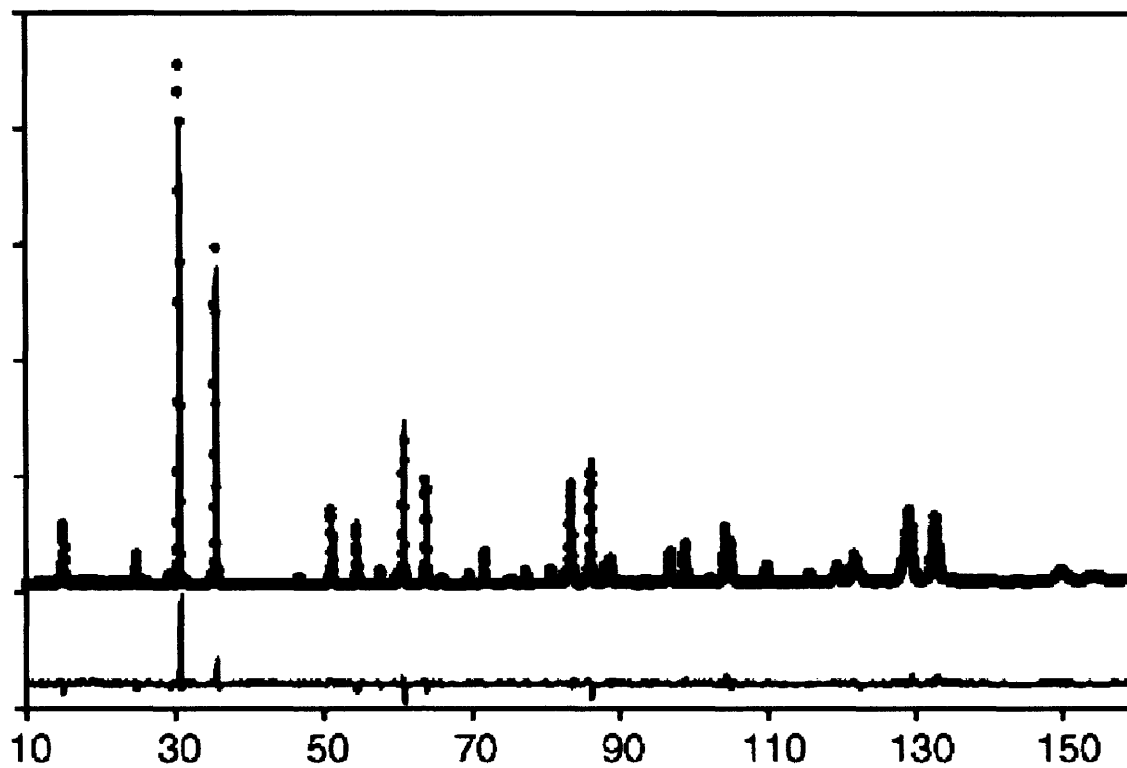
The formation energies for the layered and orthorhombic structures and the lowest-energy structures proposed in Refs. [61, 64] have been calculated using DFT. The results of these calculations are in Table 9. The calculated formation energy for the new orthorhombic structure is significantly lower than that of the any other known low-energy structure.

Structure	$\Delta E_{\text{electronic}}$	$E_{\text{vibration}}$		$\Delta E_{\text{total}}$	
		0K	298.2K	0K	298.2K
Ref. [61]	-183.5 (-1.902)	46.7 (0.484)	56.6 (0.587)	-166.3 (-1.724)	-173.6 (-1.799)
Ref. [64]	-186.6 (-1.934)	47.5 (0.492)	56.9 (0.590)	-168.6 (-1.748)	-176.4 (-1.829)
Layered Structure	-186.3 (-1.931)	47.0 (0.487)	56.3 (0.584)	-168.8 (-1.750)	-176.7 (-1.831)
Orthorhombic Structure	-188.3 (-1.952)	47.2 (0.489)	56.6 (0.587)	-170.7 (-1.769)	-178.4 (-1.849)

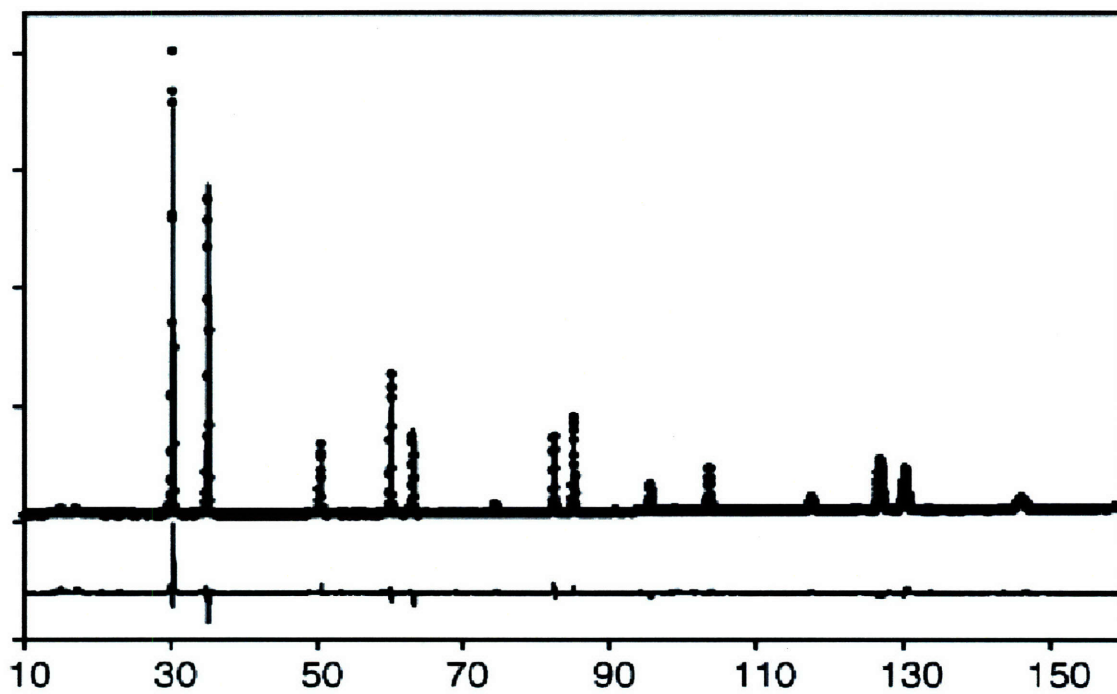
**Table 9** The calculated electronic and total formation energies and vibrational energies for the structures proposed in Refs. [61, 64] and the layered and orthorhombic structures presented in this thesis. The values are given in kJ / mol f.u. (eV / mol f.u.).

The density of the orthorhombic structure, 0.028 formula units per  $\text{\AA}^3$ , is approximately 8 to 10 percent lower than the experimentally derived density for lithium imide [59, 60, 63]. A portion of this density difference is likely due to the use of GGA, which typically overestimates the lattice parameter of materials by 1 to 2 percent [76-78]. More significant is the fact that this structure has different symmetry than that indicated by diffraction studies [59, 60, 63]. A comparison between published experimental diffraction patterns and

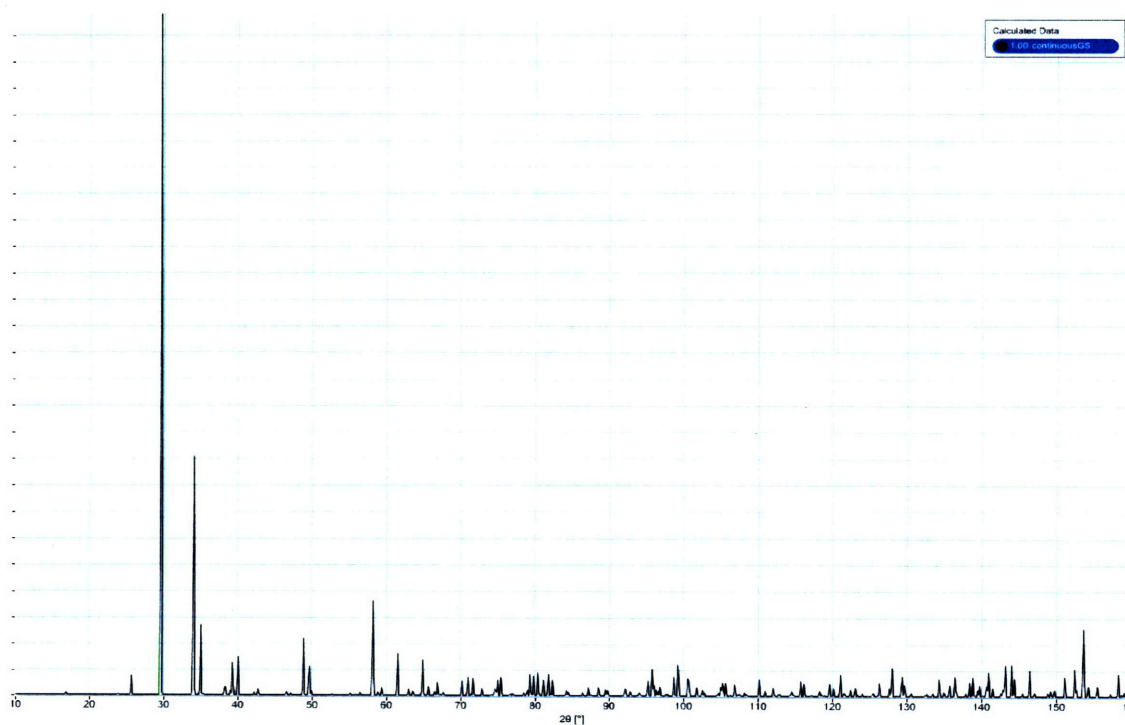
calculated diffraction patterns for the orthorhombic structure can be seen on the following pages in Figure 16 - Figure 24. Based on these comparisons, the orthorhombic structure is most likely not the experimentally observed structure.



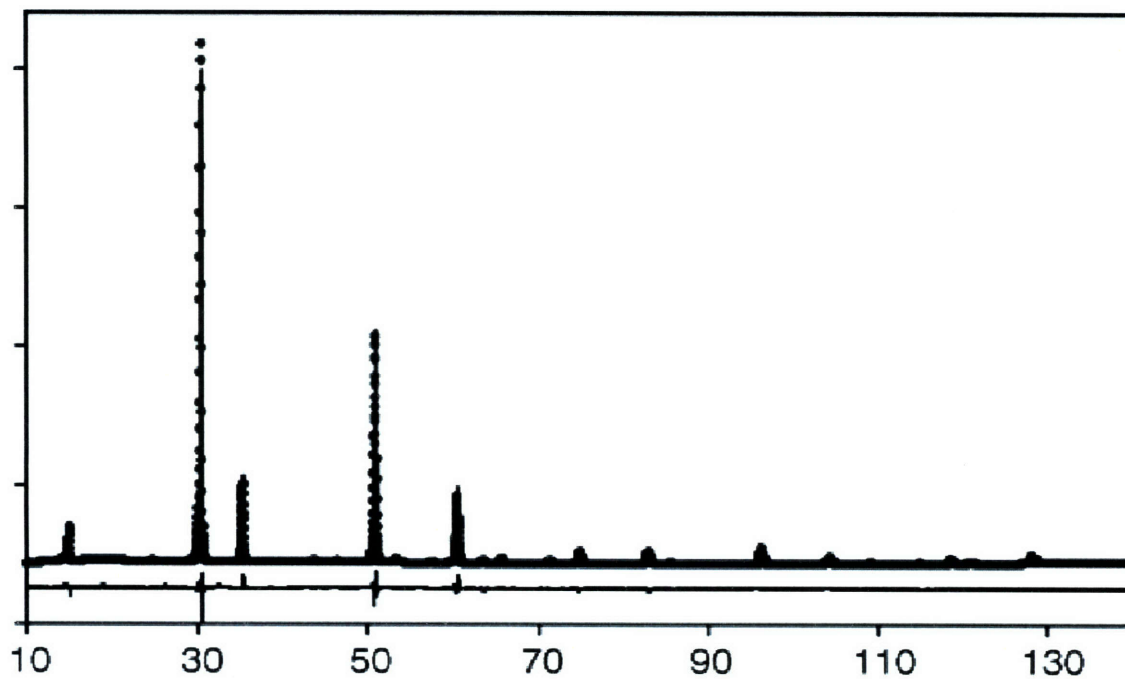
**Figure 16** The neutron powder diffraction pattern for deuterated lithium imide ( $\text{Li}_2\text{ND}$ ) at 100K from Ref. [63].



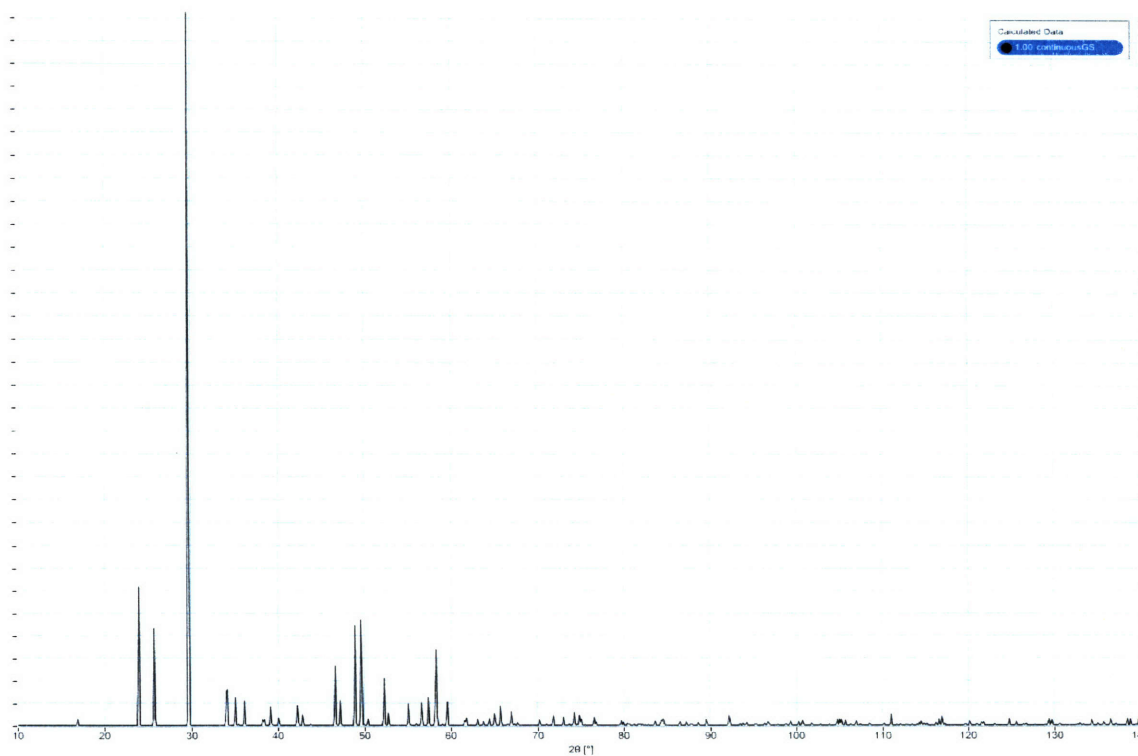
**Figure 17** The neutron powder diffraction pattern for deuterated lithium imide ( $\text{Li}_2\text{ND}$ ) at 400K from Ref. [63].



**Figure 18** The neutron powder diffraction pattern for  $\text{Li}_2\text{ND}$  in the orthorhombic structure, as calculated by CrystalDiffraction. The wavelength is the same as that used in Ref. [63] (Figure 16, Figure 17).

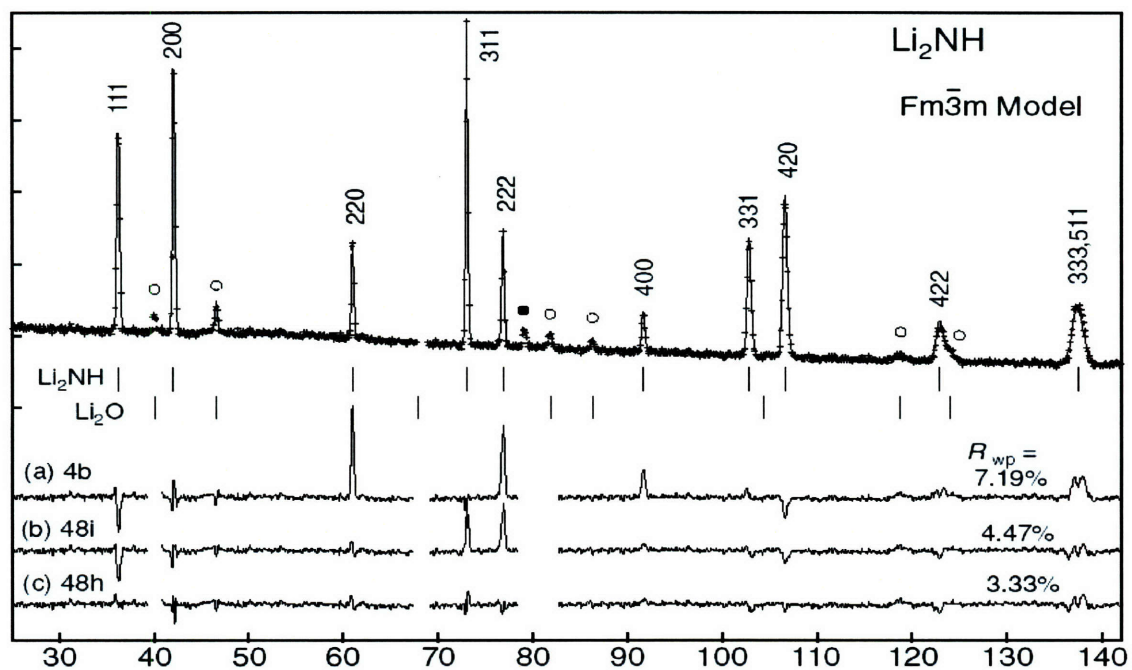


**Figure 19** The x-ray powder diffraction pattern for deuterated lithium imide ( $\text{Li}_2\text{ND}$ ) at room temperature from Ref. [63].

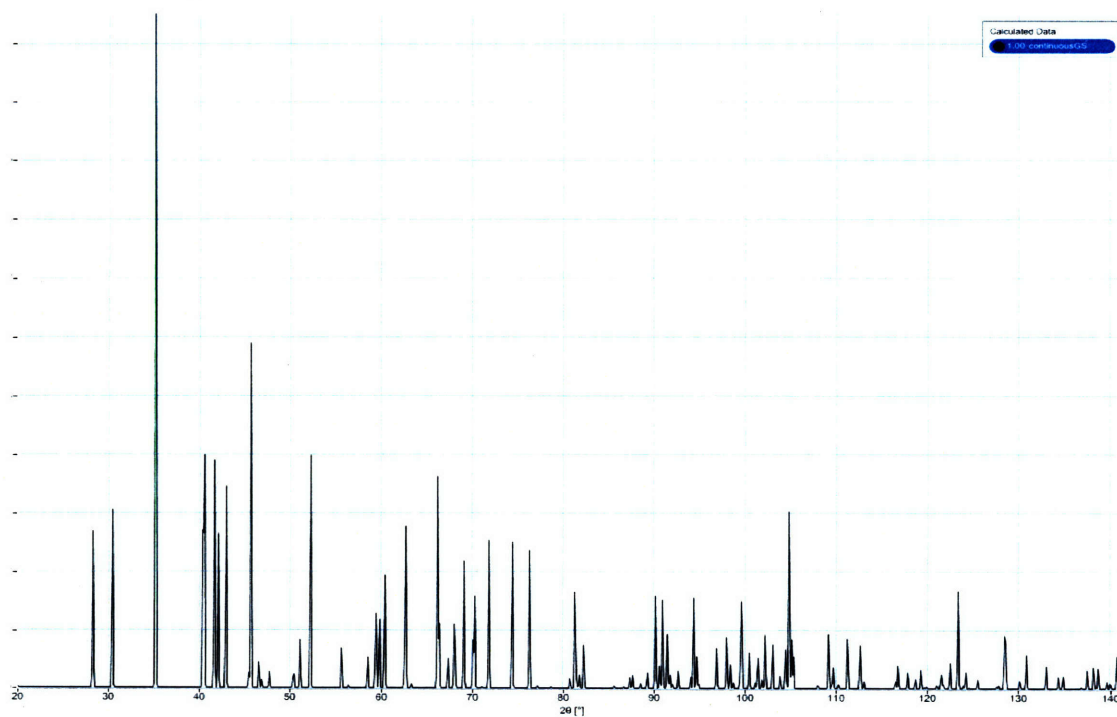


**Figure 20** The x-ray powder diffraction pattern for  $\text{Li}_2\text{ND}$  in the orthorhombic structure, as calculated by CrystalDiffract. The wavelength is the same as that used in Ref. [63] (Figure 19).

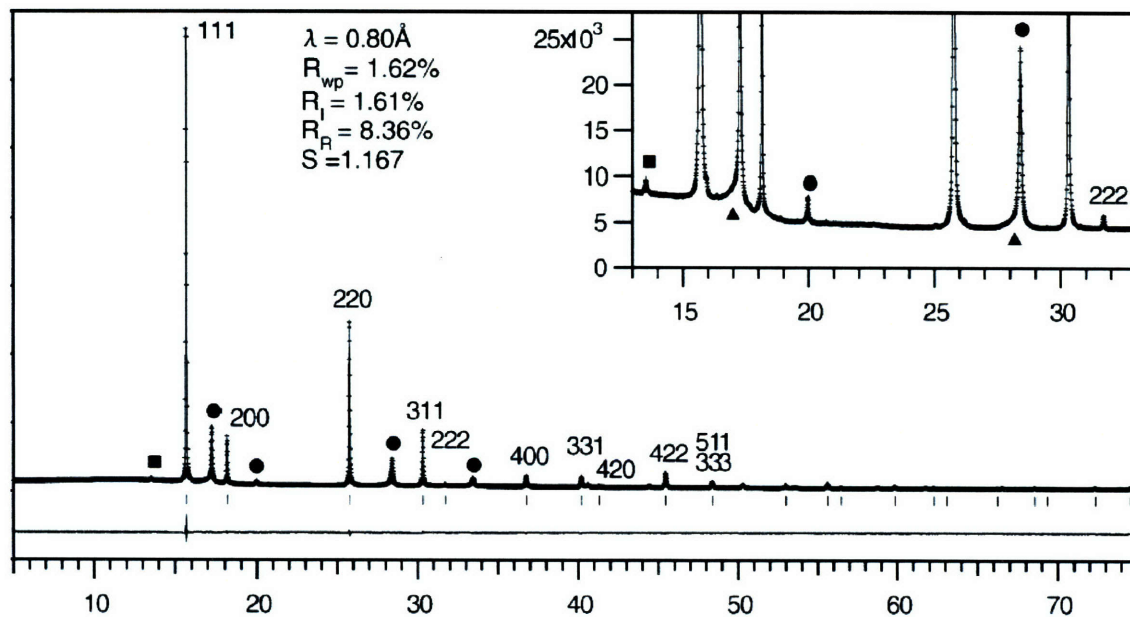




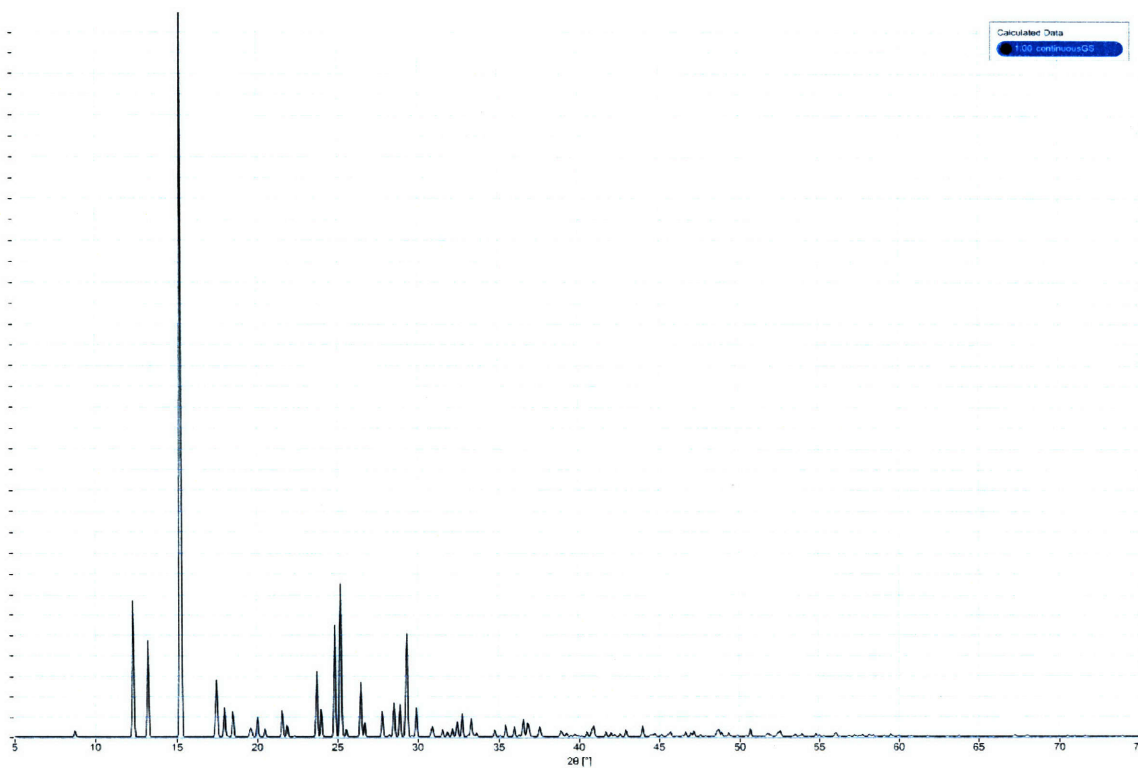
**Figure 21** The neutron powder diffraction pattern for lithium imide at room temperature from Ref. [59]. The points marked with open circles are attributed to Li<sub>2</sub>O, and the closed circle is unexplained.



**Figure 22** The neutron powder diffraction pattern for lithium imide in the orthorhombic structure, as calculated by CrystalDiffract. The wavelength is the same as that used in Ref. [59] (Figure 21).



**Figure 23** The synchrotron x-ray powder diffraction pattern for lithium imide at room temperature from Ref. [60]. The points marked with circles and triangles are attributed to  $\text{Li}_2\text{O}$ , and the square is attributed to  $\text{Li}_2\text{NCN}$ .



**Figure 24** The x-ray powder diffraction pattern for lithium imide in the orthorhombic structure, as calculated by CrystalDiffract. The wavelength is the same as that used in Ref. [60] (Figure 23).



There are several potential reasons why the data for the experimentally observed structure does not correspond to the orthorhombic structure. While the structure search was extensive it is not exhaustive, and it is possible that structures with even lower energy exist which are in better agreement with the experimentally measured diffraction pattern. It is possible that these were not found because of the limits on the search space or the limits of the model Hamiltonian. Because the nearest-neighbor interactions and their topology in the fluorite structure lead to some frustration, it is possible that more complex ordering patterns with larger unit cells have even lower energy as often is the case for frustrated systems.

It is also possible that the experimentally observed structure is not the energetic ground state but is stabilized by equilibrium finite-temperature entropy effects that are not captured by the harmonic approximation. Similarly, it may be instead a higher-symmetry metastable state with some level of frozen-in disorder. This state may be a partially disordered variation of the structure proposed in this thesis, or it may be an entirely different structure.

It is also worth considering that GGA is not accurate enough for this material. However the energy difference between the orthorhombic structure and the next-best known is about 5 meV per atom. For a comparison between similar structures, this would be an unusually large error for GGA.

Better agreement with experiment might be achieved by determining a more complete solution of the hydrogen nuclear wavefunction. Even at low temperatures there may be rapid tunneling of the hydrogen atoms between low-energy sites, similar to that proposed by Zhang et al. [62] The potential energy surface of one hydrogen nucleus is affected by location of nearby nuclei, so it is likely that any motion of hydrogen ions is locally correlated and dependent on lithium and nitrogen positions at low temperatures.

In the next section, the discrepancy between the experimental data and the calculated ground state will be examined in more detail.

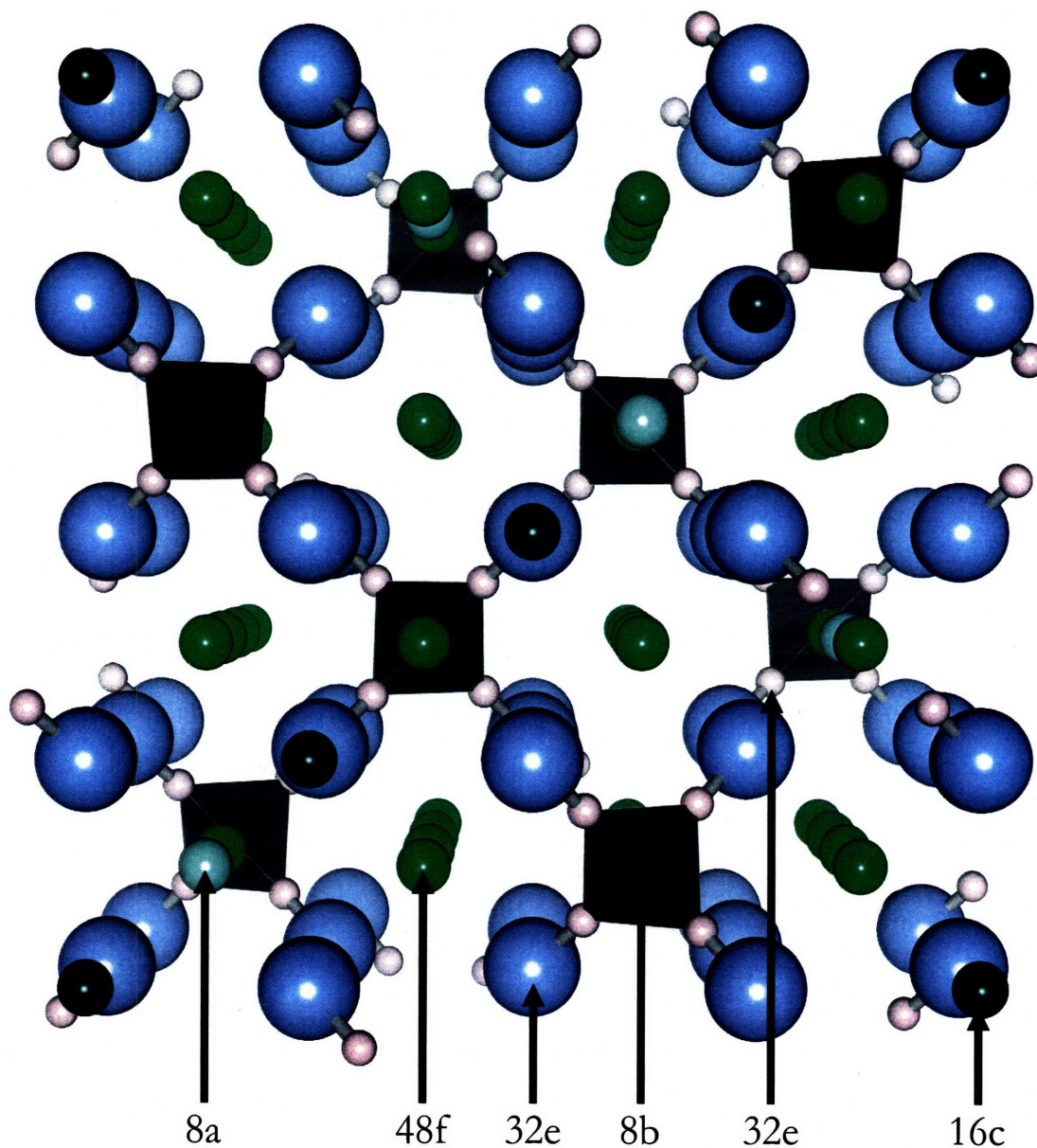
### II.3.2. The low temperature phase

---

Based on their own neutron and x-ray diffraction data, Balogh et al. conclude that the low-temperature phase has an FCC lattice, with  $Fd\bar{3}m$  symmetry and a lattice parameter of approximately 10.1 Å. The size of this unit cell corresponds to a 2x2x2 supercell of

conventional anti-fluorite unit cells. They note that this indexing also explains the previously unexplained peaks in Ref. [59] (Figure 21). However, they note that there were not any unexplained peaks in Ref. [60] (Figure 23), suggesting that the material in Ref. [60] was in a different phase, possibly due to contamination [63]. In this section, it will be assumed that the neutron and x-ray diffraction data in Refs. [59, 63] are correct, and that the low-temperature phase has a structure that closely resembles an FCC structure with  $Fd\bar{3}m$  symmetry and a lattice parameter of approximately 10.1 Å.

Based on the diffraction work by Balogh et al. [63], Herbst and Hector proposed an ordered structure for the low-temperature phase [61]. The structure proposed by Herbst and Hector differs from the antifluorite model of lithium imide in that not all tetrahedral sites are populated with lithium ions [61]. It belongs to a class of structures in which one in eight tetrahedral sites is vacant, and the N-H bonds are oriented so that four hydrogen nuclei surround each vacant tetrahedral site. With the eight lithium vacancies, structures in this class have the experimentally observed  $Fd\bar{3}m$  symmetry, as shown in Figure 25. Because this is the same space group as spinel, structures in this class will be referred to as “spinel-like” structures. In general the  $Fd\bar{3}m$  symmetry is broken when the stoichiometry is restored by placing the eight displaced lithium nuclei in vacant sites.



**Figure 25** The important sites for spinel-like lithium imide structures. The blue balls are nitrogen, the white are hydrogen, and the green three different lithium sites. The pale green (8a) are lithium sites that are as far as possible from the 8b vacancy. The medium green (48f) are lithium sites corresponding to the tetrahedral antiferro sites. The dark green (16c) are the octahedral lithium sites. In each unit cell, eight of the indicated lithium sites must be vacant.

In the structure proposed by Herbst and Hector, the eight displaced lithium nuclei are placed in alternating layers of 16c sites. Although this structure does not have the  $Fd\bar{3}m$  symmetry predicted experimentally, it agrees very well with the observed diffraction patterns

[61, 63]. The formation energy of the structure proposed by Herbst and Hector at room temperature is calculated to be about 50 meV per formula unit higher than that of the orthorhombic structure (Table 9). Despite this relatively high formation energy, it is the lowest-energy proposed structure that agrees well with the observed diffraction patterns.

The refinement of the diffraction data by Balogh et al. suggests that the lithium ions in the octahedral sites are located not at the ideal 16c sites, but at nearby 32e sites that are between the 16c and 8a sites [63]. There are two 32e sites near each 16c site, and the refined proximity of these sites to each other (approximately 1.10 Å – 1.37 Å) suggests that both cannot be populated at the same time. For this reason, the lithium in the 32e sites can be considered to be effectively associated with a single octahedral 16c site.

The distance between the octahedral lithium nuclei and nearby tetrahedral lithium nuclei in the structure proposed by Herbst and Hector, calculated as described in section II.3.1.1.2, is between 2.2 Å and 2.3 Å. In solid lithium metal, the distance between nearest-neighbor lithium ions is calculated to be 2.98 Å, which is close to the experimental value of 3.04 Å [79]. In contrast, Balogh et al. refine the distance between the octahedral lithium site and the 8a site to be between 1.50 Å and 1.64 Å [63]. The proximity between these sites makes it unlikely that both are simultaneously occupied. Because it is necessary for some of the octahedral sites to be occupied with lithium ions displaced from the 8b sites, this suggests that the nearby 8a sites are vacant.

Based on the above analysis of the experimental data, a general model for the spinel-like structure of lithium imide can be developed. The 8b sites are vacant and surrounded by hydrogen nuclei, as in the structure proposed by Herbst and Hector. However, unlike this structure, the 8a sites may be vacant as well. The lithium ions displaced from the 8a and 8b sites are near the octahedral 16c sites, most likely relaxed towards a nearby vacant 8a site. In the following sections, a cluster expansion based on this model will be used to search for new ground state structures and calculate the configurational contribution to the free energy.

### II.3.2.1. Methodology

Some of the techniques used to generate the cluster expansion for the spinel-like structures are developed and explained in the second part of this thesis. It will be assumed that the reader is familiar with these techniques.

Per conventional unit cell, there are 24 possible sites for 16 lithium nuclei, which means that the composition of the 8a and 16c sites should be 2/3 lithium. The analytical method for cluster expansions at a fixed composition was used to calculate the domain matrix for this cluster expansion, and a Monte Carlo algorithm was used to find fifteen structures that complemented the set of twenty-four manually generated spinel-like structures. The energies for all structures were calculated using DFT as described in section II.3.1.1.1.

Candidate clusters for the expansion were the eleven smallest pair clusters and three-site, four-site, five-site, and six-site clusters up to the third nearest neighbor. To fit the ECI, energies were expressed relative to the energy of the structure in the data set with the lowest known energy. Gaussian prior distributions defined on the ECI with a hierarchical width generating function were used, and a grid search was used to find the parameters that minimized the leave-one-out cross-validation score. After the ECI were fit once, the clusters with ECI below 0.1 meV per formula unit were removed from the set of candidate clusters, and the procedure was repeated until no clusters with ECI below 0.1 meV remained.

A Metropolis algorithm [73] was used to search for the structure with the lowest predicted energy. To calculate the configurational free energy, the following thermodynamic relationship was used:

$$F(T)\beta = F(T_0)\beta_0 + \int_{\beta_0}^{\beta_T} \langle E(\beta) \rangle d\beta \quad \text{Equation 18}$$

where  $F(T)$  is the Helmholtz free energy at temperature  $T$  and  $\beta = \frac{1}{k_B T}$ , where  $k_B$  is Boltzmann's constant.  $\langle E(\beta) \rangle$  is the average energy as calculated using a Metropolis algorithm run on a 6x6x6 supercell of the 128-atom conventional unit cell.  $T_0$  is a reference temperature, and  $\beta_0$  is defined accordingly. For the purposes of the integration in this section the high-temperature reference limit was used, in which  $\beta_0 = 0$  and  $F\beta_0 = -\frac{S_\infty}{k_B}$ , where  $S_\infty$  is the entropy at infinite temperature.

It was observed that in the calculated structures some lithium nuclei relaxed so that they ended up closer to sites other than their initial sites. Because the cluster expansion is a model of the local minima of the potential energy surface, the sites closest to the relaxed

nuclei were used when determining the occupation variables. To prevent the calculation of excess configurational entropy, decorations representing structures in which the nuclei relax to other sites should be excluded from configuration space. To approximately identify such decorations, the local configurations of nuclei in the candidate clusters were examined in the initial and relaxed structures. A local configuration was considered to be unstable if it was present no more than one third as often in relaxed structures as in initial structures. Structures with unstable local configurations were excluded from thermodynamic sampling. Because of the existence of the unstable structures, the high-temperature entropy limit per primitive unit cell can be estimated by:

$$\frac{S_{\infty}}{k_B} = 6 \left[ -\frac{2}{3} \ln\left(\frac{2}{3}\right) - \frac{1}{3} \ln\left(\frac{1}{3}\right) + \lim_{N_{sites} \rightarrow \infty} \left( \frac{\ln(x)}{N_{sites}} \right) \right] \quad \text{Equation 19}$$

where  $x$  is the fraction of stable configurations in structures with  $N_{sites}$  independent 8a or 16c sites. The factor of six represents the fact that there are six possible sites per primitive cell.

## II.3.2.2. Results

### II.3.2.2.1. Local interactions

The fitting procedure resulted in eleven clusters with non-zero ECI (Table 1). The leave-one-out cross-validation score for the fit is 4.6 meV per formula unit. As expected, the pair interaction between nearest-neighbor 8a and 16c sites favors different nuclei on each site. The most favorable interaction is a vacancy on the 8a site next to a lithium ion on the 16c site.

Reduced coordinates of sites	Wyckoff positions	Maximum distance between sites (Å)	Multiplicity per primitive cell	ECl (meV / formula unit)
{0, 0, 0}	16c	N/A	4	9.50102
{0.125, 0.125, 0.125}	8a	N/A	2	14.87512
{0, 0, 0}	16c	2.20143	8	5.46169
{0.125, 0.125, 0.125}	8a			
{0, 0, 0}	16c	3.59492	12	0.59523
{0, 0.25, 0.25}	16c			
{0.125, 0.125, 0.125}	8a	4.40286	4	0.22825
{-0.125, -0.125, -0.125}	8a			
{0, 0, 0}	16c	6.22658	24	0.87240
{-0.25, -0.5, 0.25}	16c			
{0, 0, 0}	16c	7.18984	12	1.05321
{-0.5, -0.5, 0}	16c			
{0, 0, 0}	16c	7.18984	12	0.11260
{0.5, 0.5, 0}	16c			
{0, 0, 0}	16c	8.03848	24	0.42894
{-0.75, 0.25, 0}	16c			
{0, 0, 0}	16c	3.59492	12	1.15799
{0.25, 0.25, 0}	16c			
{0.125, 0.125, 0.125}	8a			
{0, 0, 0}	16c	3.59492	8	0.831504
{0.25, 0.25, 0}	16c			
{0, 0.25, 0.25}	16c			

**Table 10** The clusters with non-zero ECl. The coordinates are given in terms of reduced coordinates of the conventional unit cell for the  $Fd\bar{3}m$  space group with origin 2. In the cluster expansion, occupation with lithium was given a value of +1, and vacancies were given a value of -1.

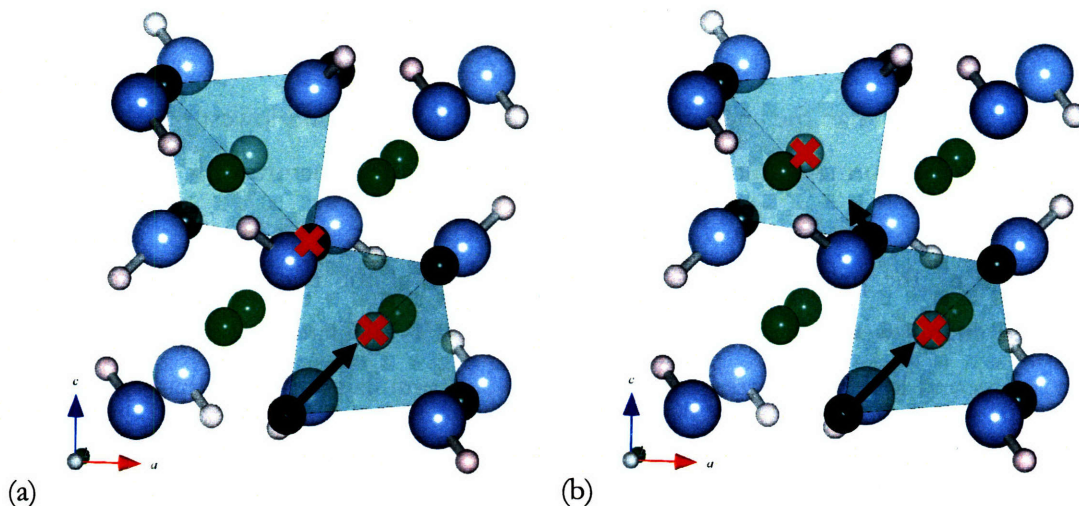
Using the method described in section II.3.2.1, it was determined that two local configurations are frequently unstable. The first configuration, shown in Figure 26(a), is a vacancy at an 8a site adjacent to a vacancy at a nearest-neighbor 16c site. Upon inspection of the calculated structures, the instability can be explained by considering the following situations:

- 1) Another nearest-neighbor 16c site adjacent to the 8a site is occupied, in which case the nucleus at that 16c site will frequently relax to the 8a site. If all four 16c sites surrounding an 8a site are occupied, this does not happen as the lithium nuclei at the 16c sites repel each other from relaxing to the center of the tetrahedron.
- 2) There are no other lithium ions in the four 16c sites next to the 8a site. In this situation, there is a cluster of five vacancies, creating a lithium void in the structure. The void tends to attract lithium from other nearby sites.

The consequence of this instability is that stable configurations will typically have either an occupied 8a site or a vacant 8a site that is surrounded by a complete tetrahedron of occupied 16c sites. By inspecting the ECI, it is apparent that local configurations in which a vacant 8a site is surrounded by four occupied 16c sites should have low energy.

The second unstable local configuration (Figure 26(b)) is closely related to the first. In the second configuration, there is a vacancy at an 8a site and a vacancy at one of its nearest neighbor 8a sites. In between the two 8a sites is a single 16c site. If the 16c site is vacant, the configuration is equivalent to the first unstable configuration. If the 16c site is occupied, the lithium nucleus on the 16c site tends to relax to the 32e site in between the 16c site and one of the 8a sites. The other 8a site sees a vacancy on the 16c site as the lithium ion relaxes away, creating a situation similar to the first unstable configuration.





**Figure 26** (a) The first unstable local configuration. (b) The second unstable local configuration. The colors are the same as in Figure 25. The green tetrahedra have at their centers 8a sites, and on each vertex a nearest-neighbor 16c site. The red crosses represent vacancies, and the dark arrows show how nearby nuclei might relax.

#### II.3.2.2.2. Ground state search

The cluster expansion was used to search for new low-energy structures. As might be predicted by inspecting the largest ECI, the lowest-energy structures were all ones in which every lithium ion is on a 16c site adjacent to a single vacant 8a site. Every 8a site is in turn surrounded by a complete tetrahedron of occupied 16c sites. To narrow down the search for low-energy structures, the configuration space was reduced further to structures which fit this description. The only remaining variable is the ordering of the vacant 8a sites. Because each vacant 8a site is coupled to four occupied 16c sites, exactly 1/3 of the 8a sites must be vacant to preserve stoichiometry. To avoid the second unstable local configuration (Figure 26(b)), none of the vacant 8a sites can be nearest-neighbors with each other.

A second cluster expansion was developed to more accurately model the ordering of the vacant 8a sites under the constraints described above. Candidate clusters included 2-site, 3-site, 4-site, and 5-site 8a clusters with sites up to 9 Å apart. The leave-one-out cross-validation score for the final cluster expansion was 3.5 meV.

Because of the numerous constraints on the system, there are only 323 possible structures with fewer than 384 atoms per unit cell. The energies of all 323 structures were

predicted using the second cluster expansion, and one structure had lower predicted formation energy than all others. The predicted energy of this structure was 3.2 meV per formula unit lower than the energy of the structure with the second-lowest predicted energy, and 4.9 meV per formula unit lower than the energy of the structure with the third-lowest predicted energy.

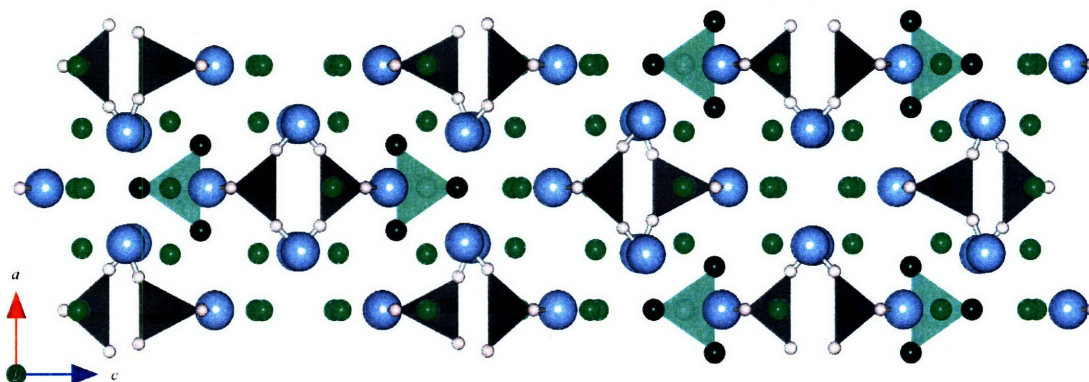
To check the predictions, the electronic energies of the two structures with lowest predicted energy were calculated using DFT. The DFT-calculated energy of the structure with the second-lowest predicted energy is 1 meV per formula unit higher than the predicted value. The DFT-calculated energy of the structure with the lowest predicted energy is 0.6 meV per formula unit lower than the predicted value, and is the lowest of all spinel-like structures. These results, along with the cross-validation score, suggest that the structure with the lowest predicted energy also has the lowest DFT-calculated energy of all 323 structures. It is proposed that this structure is the ground state of the spinel-like phase.

The space group of the proposed spinel-like ground state, as determined by ABINIT, is  $I 4_1/amd$  (#141). The lattice parameters are  $a=b=7.16 \text{ \AA}$  and  $c=30.38 \text{ \AA}$ , and the nuclear coordinates are given in Table 11. The electronic formation energy of the proposed spinel-like ground state, calculated as described in section II.3.1.1.2, is  $-1.932 \text{ eV / formula unit}$ , or  $-186.4 \text{ kJ / mol}$ . This value is about  $30 \text{ meV / formula unit}$ , or  $2.9 \text{ kJ / mol}$ , lower than the calculated energy of the structure proposed by Herbst and Hector. However it is about  $20 \text{ meV / formula unit}$ , or  $1.9 \text{ kJ / mol}$ , higher than the calculated energy of the antiferroite-like orthorhombic structure.

Element	Wyckoff Position	X	Y	Z
Lithium	8e	0	0.25	0.1246
Lithium	8e	0	0.75	0.0333
Lithium	8e	0.5	0.25	0.0349
Lithium	8e	0.5	0.75	0.0414
Lithium	16g	0.2715	0.0215	0.125
Lithium	16h	0.1746	0.75	0.0954
Lithium	32i	0.2685	0.5042	0.0390
Nitrogen	16h	0	0.9754	0.0882
Nitrogen	16h	0.5	0.9800	0.0817
Nitrogen	16h	0.7392	0.75	0.0022
Hydrogen	16h	0	0.0837	0.0657
Hydrogen	16h	0.5	0.3851	0.0943
Hydrogen	16h	0.8444	0.25	0.0213

**Table 11** The reduced coordinates of the spinel-like ground state structure, in terms of space group  $I 4_1/a m d$  (#141) orientation 2.  $a=b=7.16 \text{ \AA}$  and  $c=30.38 \text{ \AA}$ .

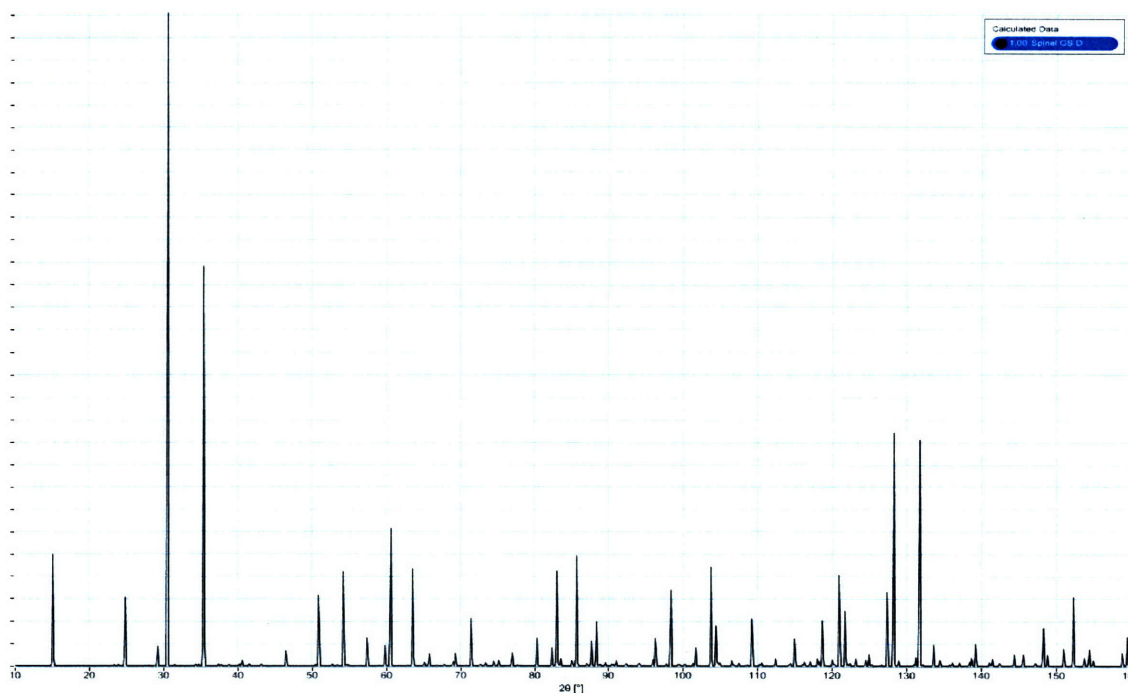
A conventional unit cell of the relaxed nuclear positions of the proposed spinel-like ground state is shown in Figure 27. The structure is simpler than it might initially appear. The structure can be created by starting with the ideal spinel-like structure shown in Figure 25, and placing vacancies in the 8a sites of every third row of lithium ions. Lithium nuclei are then placed in each of the four 16c sites around the 8a vacancies.



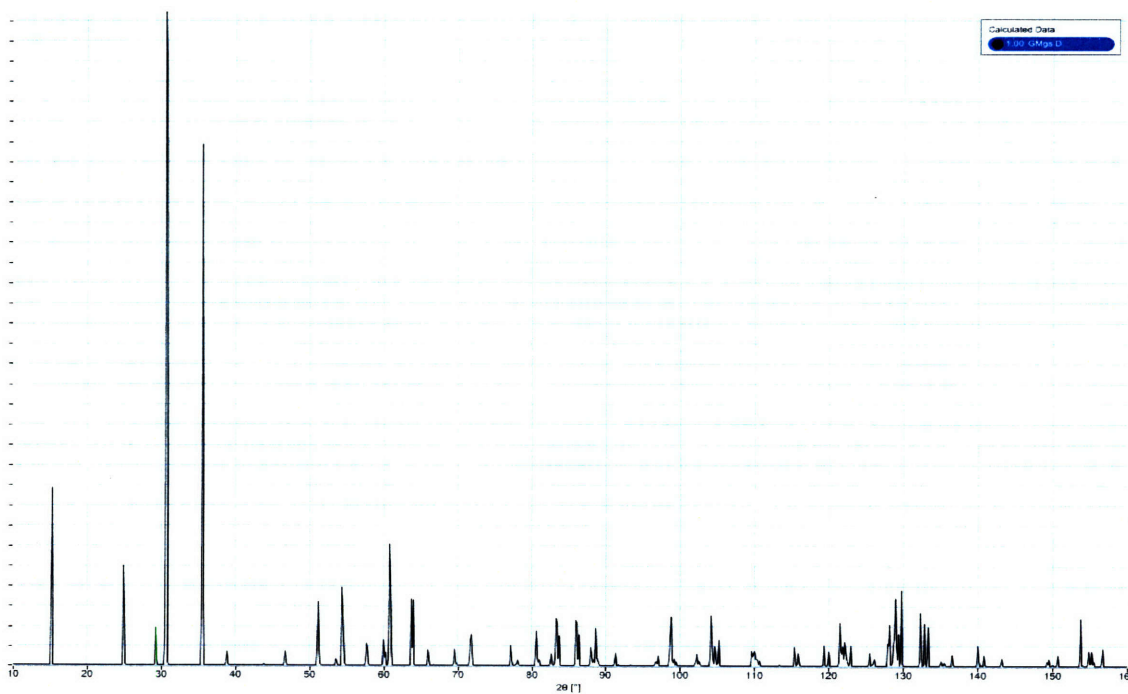
**Figure 27** A conventional unit cell of the spinel-like ground state structure. The colors are the same as in Figure 25. The grey tetrahedral are vacant 8b sites that are surrounded by hydrogen nuclei, and the green tetrahedral are vacant 8a sites surrounded by lithium nuclei in the 16c sites.

The calculated distance between vacant 8a sites and the octahedral lithium is  $1.54 \text{ \AA}$ , in excellent agreement with the 100K value  $1.52 \text{ \AA}$  resolved by Balogh et al. [63], considering that GGA typically underbinds. This distance is much closer than the calculated distance of  $2.2 \text{ \AA} - 2.3 \text{ \AA}$  in the structure proposed by Herbst and Hector. The calculated lattice parameter in the c direction for 12 layers of lithium nuclei is  $30.38 \text{ \AA}$ , meaning that the lattice parameter per four layers of lithium nuclei in the c direction is  $10.1261 \text{ \AA}$ . The calculated lattice parameter per four layers of lithium nuclei in the directions orthogonal to the c direction is  $10.1288 \text{ \AA}$ . These values are in good agreement with the isotropic 100 K lattice parameter of  $10.0873 \text{ \AA}$  resolved by Balogh et al. [63]

The predicted x-ray and neutron diffraction spectra for the proposed spinel-like ground state, and the structure proposed by Herbst and Hector, are given in Figure 28 to Figure 35. These spectra can be compared directly to the experimental data in Figure 16, Figure 17, Figure 19, Figure 21, and Figure 23. The patterns for the proposed spinel-like ground state and the structure proposed by Herbst and Hector are similar, with the proposed spinel-like ground state generally having slightly sharper peaks. This is positive for the proposed spinel-like ground state, as the structure proposed by Herbst and Hector is known to agree well with experimental data [63]. The proposed spinel-like ground state appears to agree slightly better with some experimental data, such as the peaks near  $2\theta=130^\circ$  in Figure 28.

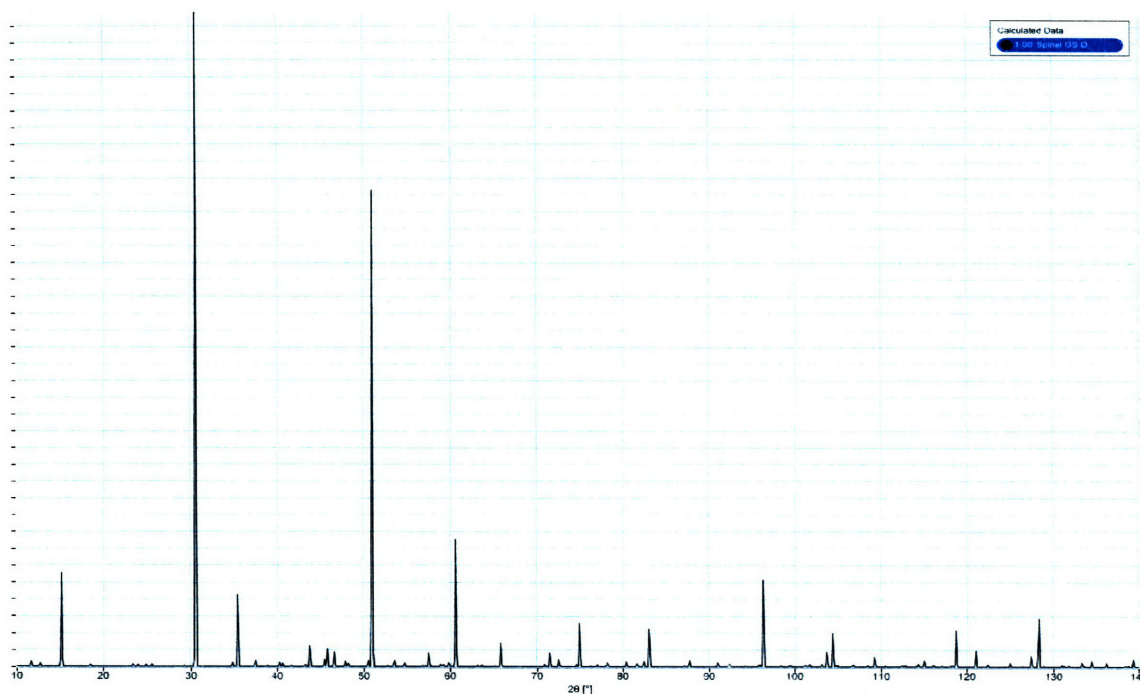


**Figure 28** The neutron powder diffraction pattern for  $\text{Li}_2\text{ND}$  in the proposed spinel ground state, as calculated by CrystalDiffract. The wavelength is the same as that used in Ref. [63] (Figure 16, Figure 17).

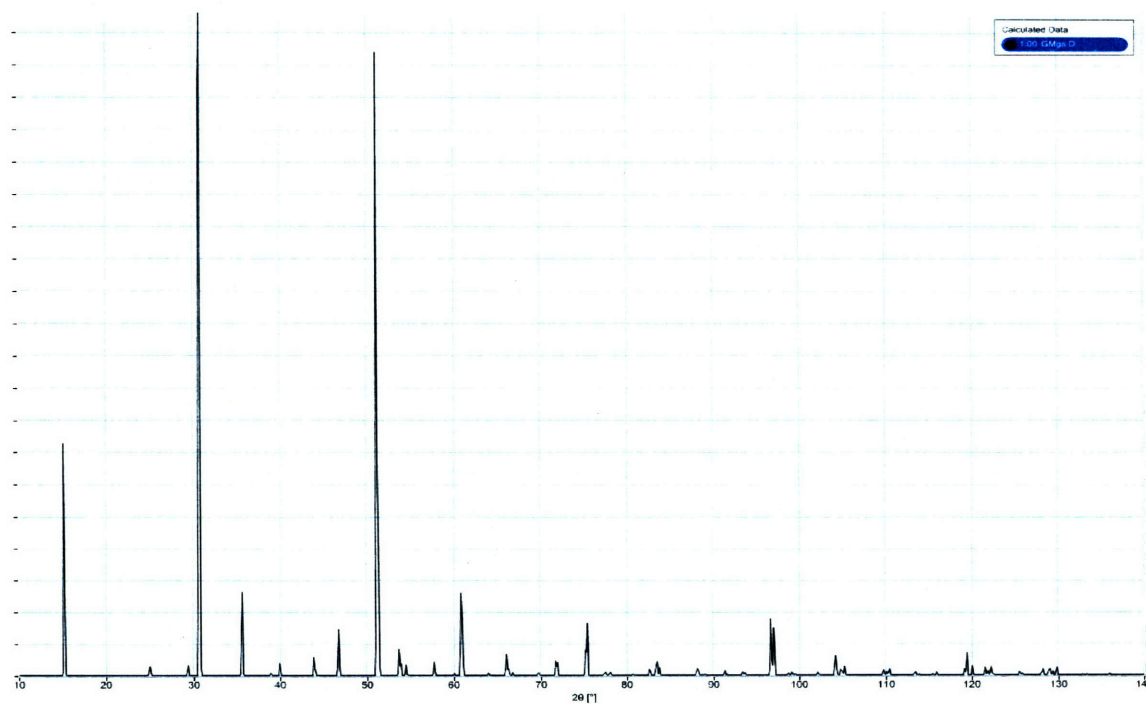


**Figure 29** The neutron powder diffraction pattern for  $\text{Li}_2\text{ND}$  in the structure proposed by Herbst and Hector [61], as calculated by CrystalDiffract. The wavelength is the same as that used in Ref. [63] (Figure 16, Figure 17).

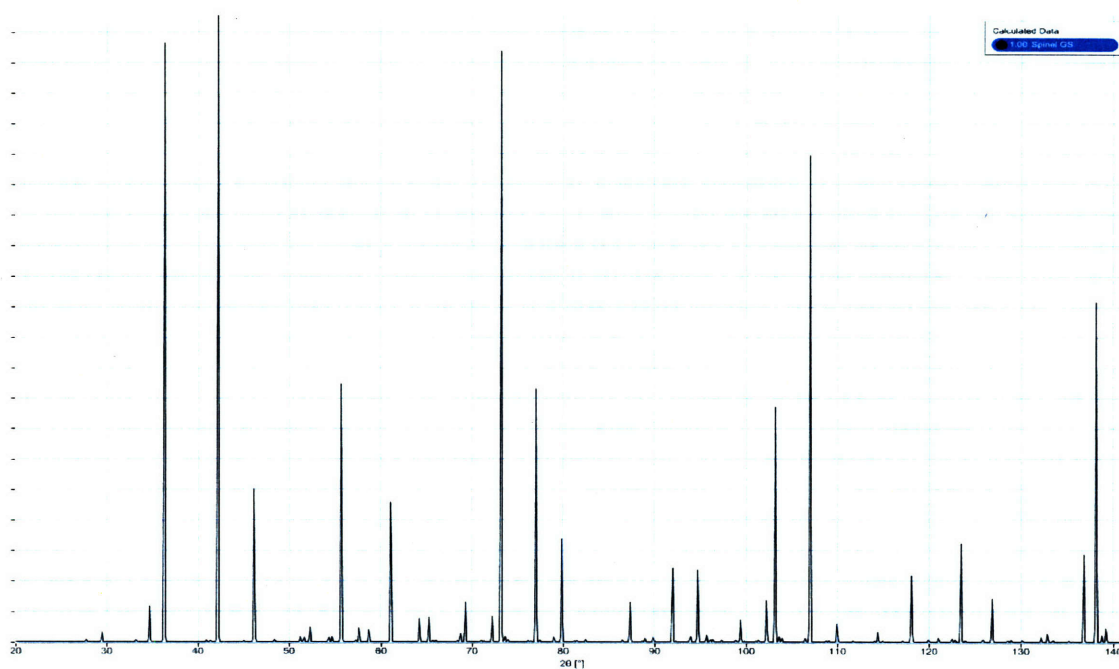




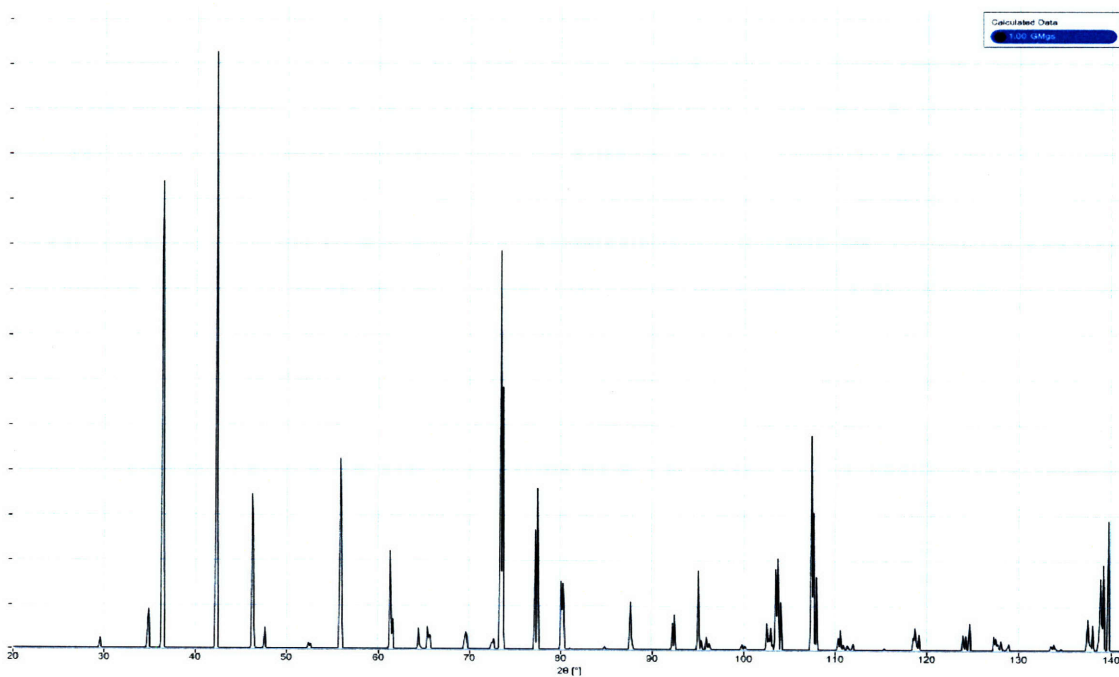
**Figure 30** The x-ray powder diffraction pattern for  $\text{Li}_2\text{ND}$  in the proposed spinel ground state, as calculated by CrystalDiffract. The wavelength is the same as that used in Ref. [63] (Figure 19).



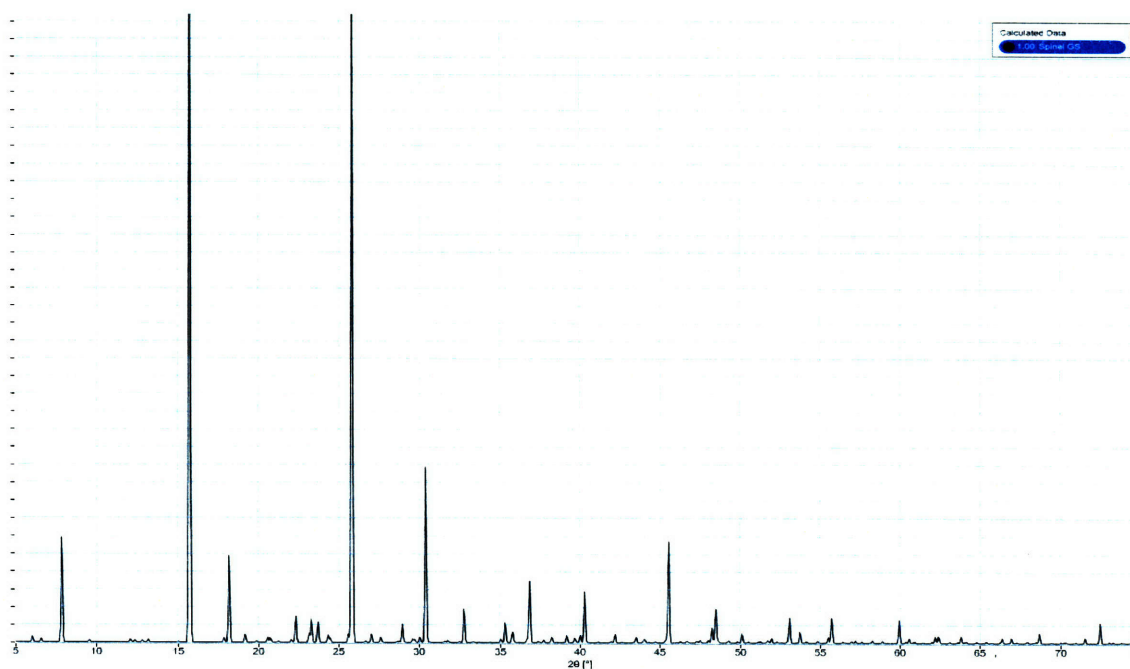
**Figure 31** The x-ray powder diffraction pattern for  $\text{Li}_2\text{ND}$  in the structure proposed by Herbst and Hector [61], as calculated by CrystalDiffract. The wavelength is the same as that used in Ref. [63] (Figure 19).



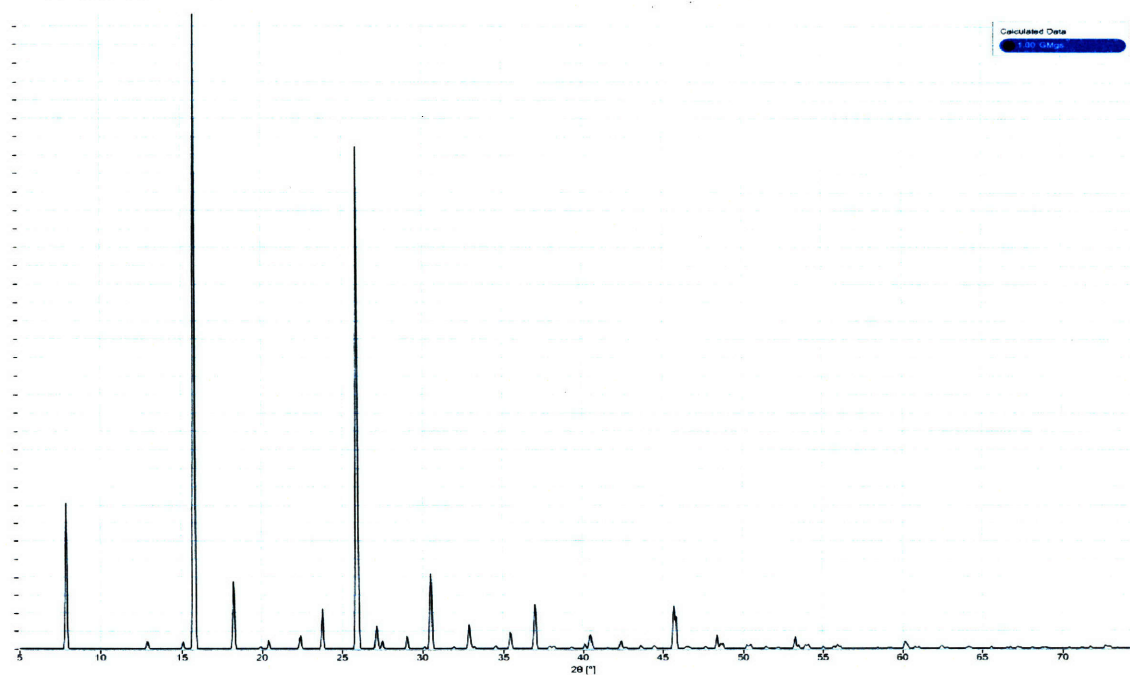
**Figure 32** The neutron powder diffraction pattern for lithium imide in the proposed spinel ground state, as calculated by CrystalDiffract. The wavelength is the same as that used in Ref. [59] (Figure 21).



**Figure 33** The neutron powder diffraction pattern for lithium imide in the structure proposed by Herbst and Hector [61], as calculated by CrystalDiffract. The wavelength is the same as that used in Ref. [59] (Figure 21).



**Figure 34** The x-ray powder diffraction pattern for lithium imide in the proposed spinel-like ground state, as calculated by CrystalDiffract. The wavelength is the same as that used in Ref. [60] (Figure 23).



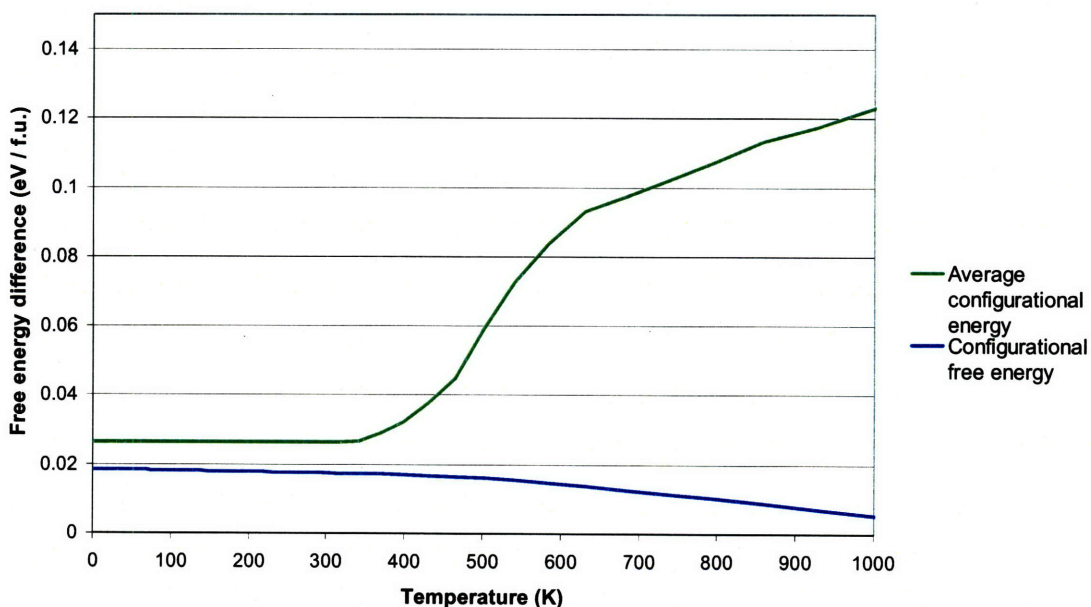
**Figure 35** The x-ray powder diffraction pattern for lithium imide in the structure proposed by Herbst and Hector [61], as calculated by CrystalDiffract. The wavelength is the same as that used in Ref. [60] (Figure 23).



### II.3.2.2.3. Finite-temperature behavior

The proposed spinel-like ground state has a significantly higher calculated energy than the antiferroite-based orthorhombic structure. However, the spinel-like structure agrees much better with experimental data. It is possible the orthorhombic structure is preferred at very low temperatures, but the spinel-like structures are entropically stabilized at slightly higher temperatures. This entropic stabilization could be caused by vibrational and/or configurational entropy.

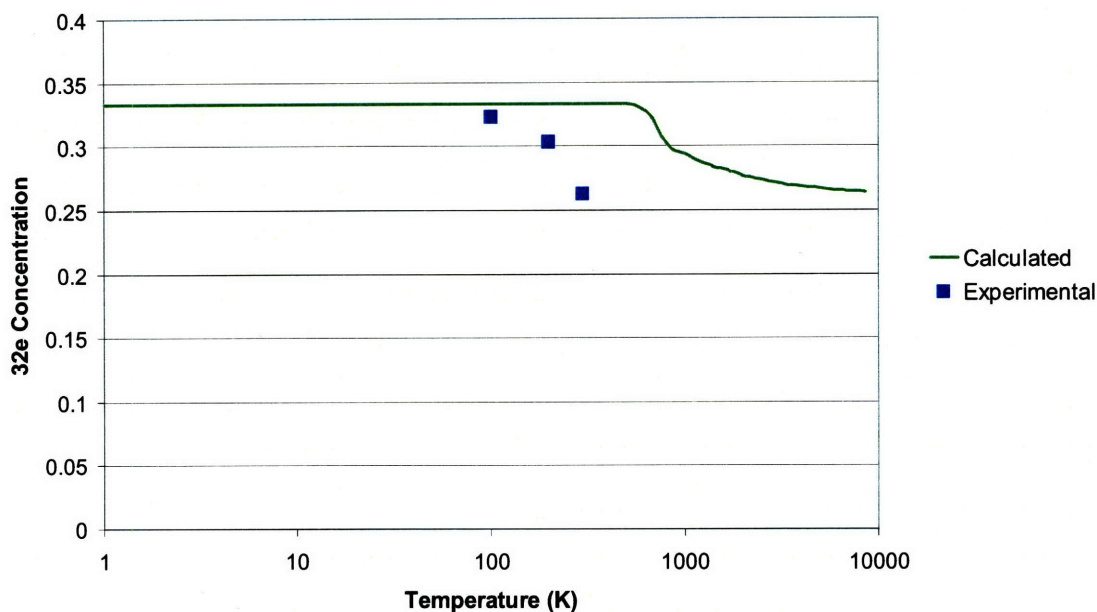
The configurational free energy of the spinel-like phase was calculated by thermodynamic integration as described in section II.3.2.1. The high-temperature entropy limit was approximated using Equation 19 with  $x$  calculated by sampling  $1.5 \times 10^{12}$  decorations from a  $3 \times 3 \times 3$  supercell of the 32-atom primitive unit cell. The high-temperature configurational entropy per formula unit, in units of  $k_B$  (Boltzmann's constant), was estimated to be  $2.87 \pm 0.03$ . The configurational energy and free energy are plotted in Figure 36. The difference between the two is the entropic contribution to the configurational free energy. There is a calculated phase transition at around 450K, in which the proposed spinel-like ground state disorders.



**Figure 36** The values of the average and free energies of the proposed spinel-like ground state relative to the energy of the orthorhombic structure. The difference between the two lines is the contribution of configurational entropy to the stabilization of the spinel-like phase.

As the temperature is increased from 100K to 300K, the experimentally observed occupation of lithium in the 32e (octahedral) site decreases from 32% to 26% [63]. (Because there are two 32e sites per octahedral site, and they cannot both be occupied at the same time, this corresponds to octahedral occupations of 64% to 52%.) In the structure proposed by Herbst and Hector, the occupation of the 32e sites is 25%. Based on this structure, Balogh et. al suggested that the higher-than-expected octahedral occupation might be due to excess lithium [63]. However this explanation should be reconsidered in the context of the proposed spinel-like ground state. In the proposed spinel-like ground state, 33% of the 32e sites are occupied, which is more consistent with the low-temperature 100K measurement of 32%. Because the average occupation of the 8a and 16c (octahedral) sites is  $2/3$ , it should be expected that the high-temperature limit of occupation of 32e sites should also be 33% lithium. However, at 300K, it is experimentally observed that the occupation decreases to 26%.

A possible reason for this decrease is that experimentally observed decorations are metastable. Decorations that are unstable are likely very short-lived, and the lithium nuclei spend most of their time close to the local minima in the potential energy surface. If decorations that are expected to be unstable are disallowed, the calculated occupation of 32e sites decreases with increasing temperature to a limit of about 26.1%. Intuitively, this occurs because vacancies at 8a sites are generally only stable if surrounded by a complete tetrahedron of occupied 32e sites. As the temperature increases, these tetrahedra begin to break up, which causes lithium nuclei to collapse from 32e sites to nearby 8a sites (Figure 26). The more tetrahedra are broken up, the more nuclei move to 8a sites, reducing the percentage of lithium on the 32e sites. A comparison of the calculated occupation with experimental observations is shown in Figure 37.

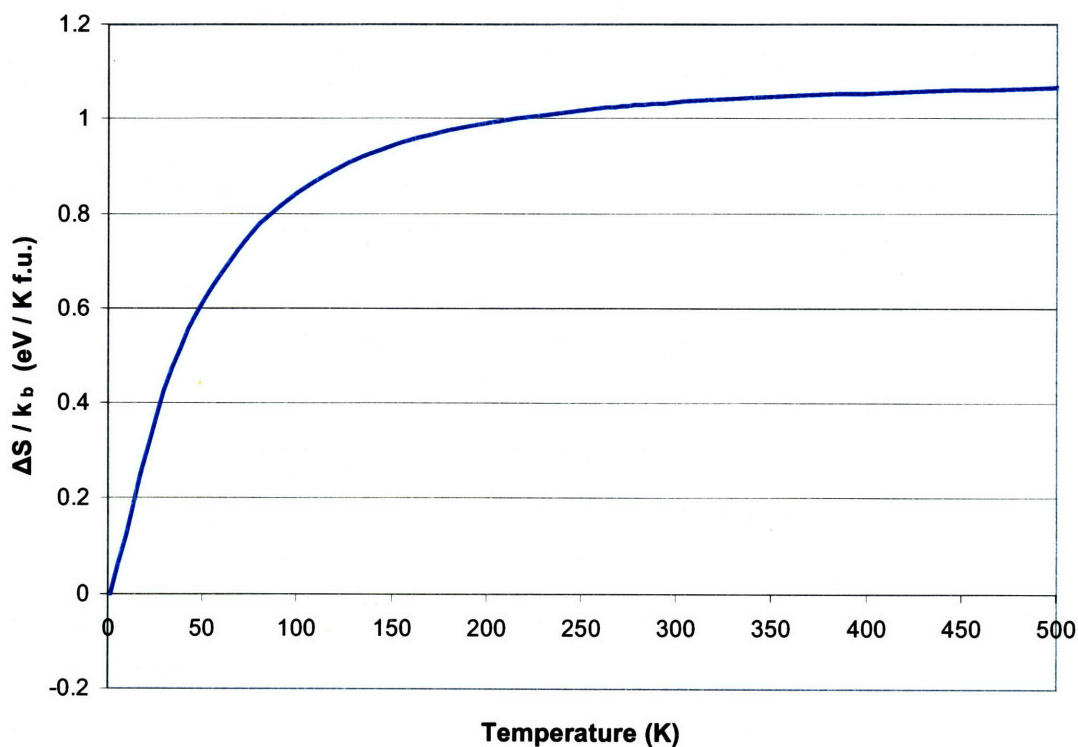


**Figure 37** A comparison of the experimental and calculated lithium occupation of the 32e sites. The occupation decreases to a limit of about 26.1% due to the instability of vacant 8a sites adjacent to fewer than four occupied 32e sites.

The reduction in 32e occupation is calculated to occur at a higher temperature than experimentally observed. A better fit to the experimental data might be achieved if a more complex model were used to predict which decorations represent unstable structures. In addition, the 8a sites might become more favorable with increasing temperature due effects not considered here, such as volume expansion or greater vibrational entropy for structures with more highly occupied 8a sites. Comparison with experimental data suggests that the phase transition to a disordered spinel-like structure might occur at a lower temperature than calculated, which could significantly reduce the configurational free energy below what is shown in Figure 37, especially for temperatures between 100K and 1000K.

To estimate the vibrational entropy of spinel-like structures, the vibrational entropy of the structure proposed by Herbst and Hector is used as a prototype. The calculation was done as described in section II.3.1.1.2. The vibrational entropy difference between the antifluorite-based orthorhombic structure and the spinel-like structure is shown in Figure 38. The vibrational entropy is higher in the spinel-like structure, likely because the ability for lithium ions to move between 8a and 16c sites without a potential barrier favors soft modes.

There is a loss in configurational entropy due to disallowing unstable decorations, but this is partially offset by the increase in the vibrational entropy. Intuitively, there are fewer local minima, which reduces the configurational entropy, but the local minima are wider, which increases the vibrational entropy.

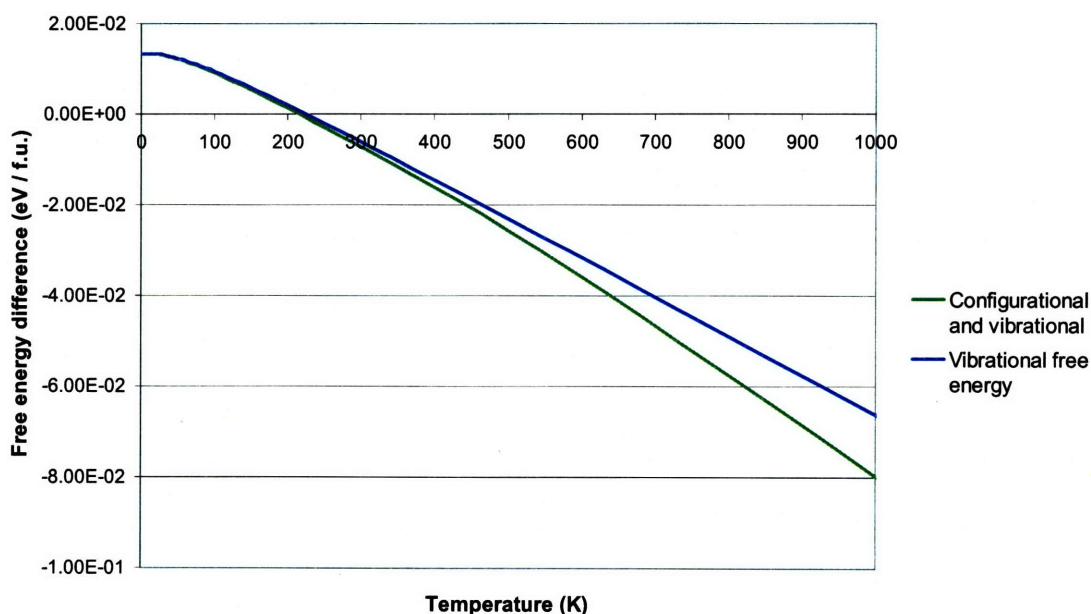


**Figure 38** The vibrational entropy of the structure proposed by Herbst and Hector relative to the vibrational entropy of the orthorhombic structure.

The extra vibrational entropy in the spinel-like phase, combined with the relatively low energy of the proposed spinel-like ground state, suggests that the spinel-like phase may be entropically stabilized. A plot of the estimated free energy, including vibrational effects, of the spinel-like phase relative to that of the orthorhombic structure is shown in Figure 39. It can be seen that the spinel-like phase is estimated to become more stable at temperatures above about 225K. If configurational entropy is included, the transition is predicted to be slightly lower. By comparing the observed and calculated occupations of the 32e sites as a function of temperature (Figure 37), it can be inferred that the calculated configurational entropy might be too low in this temperature range. In this case, it can be expected that



spinel-like phase becomes more stable at an even lower temperature. This could explain why the antiferroite-like orthorhombic structure is not seen experimentally. At the temperature at which the transition to the orthorhombic structure should take place, the kinetics might be very slow, effectively trapping the material in a metastable spinel-like state.



**Figure 39** The estimated vibrational free energy of the spinel-like phase relative to the vibrational free energy of the orthorhombic structure. The spinel-like phase becomes more stable above approximately 225K. The combined configurational and vibrational free energy difference is also shown.

### II.3.2.3. Discussion

An explanation for the discrepancy between the experimental data and the calculated ground state structure has been proposed. The orthorhombic structure introduced in section II.3.1.2.2 remains the most likely ground state structure, but at low temperatures a spinel-like phase becomes entropically stabilized. The existence of this phase agrees very well with published diffraction data. The orthorhombic structure most likely has not been seen experimentally because lithium imide is typically synthesized above the temperature at which it becomes thermodynamically unstable. It is possible that even at temperatures below the transition temperature, slow reaction kinetics have prevented the observation of the ground state structure.

Experiments indicate a second-order phase transition in lithium imide near 360K [63, 66]. Above this transition temperature, diffraction data indicate that lithium imide transitions to more of an antiferroite structure. It is likely that the high temperature phase involves disorder among the imide group orientations as well as continuing disorder in the positions of the lithium nuclei. The free energy of lithium imide above this temperature must be even lower than the free energy curve for the spinel-like phase. Lithium, nitrogen, and hydrogen are key elements in a variety of potential hydrogen storage systems. These results suggest that the configurational and vibrational contributions to the free energy should be considered when determining reaction products and equilibrium temperatures in such systems.

### III. A Bayesian approach to building cluster expansions





---

## III.1. Background and motivation

---

Approximate values for many material properties can be calculated using the Born-Oppenheimer approximation [80], in which atomic nuclei are treated as classical particles, if the potential energy surface of the material is known. The potential energy surface is an expression of the potential energy of the material as a function of the positions of the atomic nuclei:

$$V(\{\bar{r}_i, s_i\}) \tag{Equation 20}$$

where  $\bar{r}_i$  is the position of the *i*th nucleus and  $s_i$  is a discrete variable representing the number of protons and neutrons in the *i*th nucleus. Among the important material properties that can be derived from the potential energy surface are:

- The ground-state structure. This is the atomic arrangement the material would achieve if it were left at a very low temperature for an infinite period of time. It corresponds to the global minimum on the potential energy surface.
- Free energy of the material. This can be calculated statistically by sampling the potential energy surface with a sample weight that is a function of the potential energy.

The calculation of these and other material properties requires sampling or searching the potential energy surface. Because the domain of  $\{\bar{r}_i, s_i\}$  grows exponentially with the number of particles in the system, effective searching or sampling of the potential energy surface for systems with more than a few unique atoms requires the ability to rapidly evaluate  $V(\{\bar{r}_i, s_i\})$ .

Evaluating  $V(\{\bar{r}_i, s_i\})$  exactly within the Born-Oppenheimer approximation would require solving the multi-electron Schrödinger equation. Because no analytical solution to the Schrödinger equation is known for any system with more than one electron, a variety of methods for calculating approximate solutions to the Schrödinger equation using numerical methods have been developed. In general, the more accurately a method approximates a solution to the Schrödinger equation, the worse the method scales with the number of electrons in the system. For example, using one of the most popular methods, density

functional theory [31], calculating a reasonable approximation for  $V(\{\bar{r}_i, s_i\})$  on a modern x86 processor might take on the order of CPU-hours for a system with twenty unique nuclei. Effective sampling or searching of the potential energy surface of such a system using density functional theory is computationally infeasible. However, frequently it *is* computationally feasible to use a small sample of accurate, but expensive, calculations to parameterize an effective potential function that can be rapidly evaluated. Such an approach often strikes a good balance between the needs for speed and accuracy.

Many of the important properties of materials can be well approximated by considering only the *local minima* of the potential energy surface. It is therefore often sufficient to develop an effective potential that only models these local minima. One form of such effective potentials that has been successfully used to model a variety of systems is the cluster expansion [69, 81]. Like all potential functions, the success of the cluster expansion depends on its ability to accurately predict the values of the potential energy surface. The focus of this chapter will be on the development of methods to efficiently develop cluster expansions that accurately predict these values.

### III.1.1. The cluster expansion

---

In this section a brief introduction to the cluster expansion formalism and some of its important properties will be presented. A more thorough review of the cluster expansion and its applications can be found in references [69, 81].

In the cluster expansion, it is assumed that structures which correspond to local minima on the potential energy surface are ones in which all nuclei are at or near a fixed set of low-energy sites. It is assumed that there is nearly a one-to-one mapping between the different ways in which these sites can be occupied by the atomic nuclei and the low-energy local minima on the potential energy surface. For example, the low-energy local minima might all correspond to ways in which atomic nuclei can be arranged on a cubic lattice.

If the low-energy local minima are assumed to correspond to different nuclei occupying a fixed set of sites, the problem of evaluating  $V(\{\bar{r}_i, s_i\})$  at the local minima can be replaced with the problem of evaluating  $V(\{c_k\})$ , where  $c_k$  is a discrete “occupation variable” that represents the type of nucleus present on the *k*th high-symmetry site (one value could

represent the existence of nothing on the site). For simplicity,  $\{c_k\}$  will be represented with the vector symbol  $\vec{c}$  :

$$V(\vec{c}) = V(c_1, c_2 \dots) \quad \text{Equation 21}$$

The multi-variable function  $V(c_1, c_2 \dots)$  can be expanded in a set of basis functions [69]. To build this basis, single-variable bases for the occupation variables at each of the sites are first created. Let the  $j$ th basis function for the  $k$ th occupation variable be denoted  $\phi_{j,k}(\cdot)$ . The complete multi-variable basis can be constructed by taking the tensor product of the single-variable bases. The function  $V(c_1, c_2 \dots)$  can be expanded in terms of these tensor product basis functions:

$$V(c_1, c_2 \dots) = \sum_{\vec{j}} V_{\vec{j}} \prod_k \phi_{j_k, k}(\sigma_k) \quad \text{Equation 22}$$

where  $\vec{j}$  represents a particular product of single-variable basis functions and  $j_k$  is the  $k$ th element of  $\vec{j}$ . The sum is over all possible  $\vec{j}$ .

It is convenient use single-variable bases in which  $\phi_{0,k} = 1$  for all  $k$ . This allows the expansion of Equation 22 in terms of products of a finite number of basis functions:

$$V(c_1, c_2 \dots) = V_{\vec{j}_0} + \sum_{\vec{j} \in J_1} V_{\vec{j}} \prod_k \phi_{j_k, k}(c_k) + \sum_{\vec{j} \in J_2} V_{\vec{j}} \prod_k \phi_{j_k, k}(c_k) + \dots \quad \text{Equation 23}$$

where  $\vec{j}_0$  is the vector of all 0's and  $J_n$  is the set of all  $\vec{j}$  with exactly  $n$  non-zero elements. The non-zero elements in  $\vec{j}$  represent the sites for which the occupation affects the value of the cluster function, and set of such sites can be thought of as the cluster represented by  $\vec{j}$ . The coefficients  $V_{\vec{j}}$  are known as the effective cluster interactions (ECI). The first term represents the average potential energy, the second term represents the effect the occupation of single sites has on the potential energy, the third term represents the effect of pairs of sites, etc.

The expansion in Equation 23 is known as the ‘‘cluster expansion’’, because it expands the potential energy into contributions from clusters of sites. The product  $\prod_k \phi_{j_k, k}(c_k)$  will be referred to as a ‘‘cluster function’’. Each cluster function is defined by a given vector  $\vec{j}$

and represents an interaction between all sites for which  $j_k \neq 0$ . The variable  $\varphi_j$ , defined by

$$\varphi_{\vec{j}}(\vec{c}) \equiv \prod_k \phi_{j_k, k}(c_k) \quad \text{Equation 24}$$

will be used to represent a cluster function.

Combining Equation 23 and Equation 24 it can be seen that the cluster expansion is an expansion of  $V(c_1, c_2 \dots)$  into a linear combination of cluster functions:

$$V(c_1, c_2 \dots) = V_{j_0} + \sum_{\vec{j} \in J_1} V_{\vec{j}} \varphi_{\vec{j}}(\vec{c}) + \sum_{\vec{j} \in J_2} V_{\vec{j}} \varphi_{\vec{j}}(\vec{c}) + \dots \quad \text{Equation 25}$$

Often the arrangement of sites in a cluster expansion is symmetric. To take advantage of this symmetry, the same single-variable basis functions and occupation variables must be used at symmetrically equivalent sites. By doing this, orbits of symmetrically equivalent cluster functions can be created. Because all functions in a given orbit are symmetrically equivalent, they must all have the same coefficient  $V_{\vec{j}}$ . Let  $\alpha$  represent a set of vectors  $\vec{j}$  corresponding to symmetrically equivalent cluster functions, and  $V_\alpha$  be the common coefficients for these functions. Equation 25 can be re-written in terms of these common coefficients:

$$V(c_1, c_2 \dots) = V_{j_0} + \sum_{\alpha \in A_1} V_\alpha \sum_{\vec{j} \in \alpha} \varphi_{\vec{j}}(\vec{c}) + \sum_{\alpha \in A_2} V_\alpha \sum_{\vec{j} \in \alpha} \varphi_{\vec{j}}(\vec{c}) + \dots \quad \text{Equation 26}$$

where  $A_n$  is the set of all different orbits  $\alpha$  for which the vectors  $\vec{j} \in \alpha$  have  $n$  non-zero elements.

Let  $N_\alpha$  be defined the number of elements in  $\alpha$ , and the ‘‘correlation’’  $\langle \varphi \rangle_\alpha$  be defined by

$$\langle \varphi(\vec{c}) \rangle_\alpha \equiv \frac{\sum_{\vec{j} \in \alpha} \varphi_{\vec{j}}(\vec{c})}{N_\alpha} \quad \text{Equation 27}$$

Substituting Equation 27 into Equation 26 results in

$$V(c_1, c_2 \dots) = V_{j_0} + \sum_{\alpha \in A_1} V_\alpha N_\alpha \langle \varphi(\vec{c}) \rangle_\alpha + \sum_{\alpha \in A_2} V_\alpha N_\alpha \langle \varphi(\vec{c}) \rangle_\alpha + \dots \quad \text{Equation 28}$$

Often the value of interest is not the potential energy of the entire structure, but the average potential energy per some symmetrically repeated group of sites. This is especially

true for cluster expansions defined on infinite crystal lattices, where only the averaged potential energy is necessarily finite. Let the average value of a quantity  $x$  per symmetrically repeated group of sites be denoted by  $\langle x \rangle_{Unit}$ . Using this notation, the multiplicity of a cluster orbit  $m_\alpha$  will be defined as

$$m_\alpha \equiv \langle N_\alpha \rangle_{Unit} \quad \text{Equation 29}$$

From Equation 29 and Equation 28, the average potential energy can be expressed as

$$\langle V(c_1, c_2, \dots) \rangle_{Unit} = \langle V_{j_0} \rangle_{Unit} + \sum_{\alpha \in A_1} V_\alpha m_\alpha \langle \phi(\bar{c}) \rangle_\alpha + \sum_{\alpha \in A_2} V_\alpha m_\alpha \langle \phi(\bar{c}) \rangle_\alpha + \dots \quad \text{Equation 30}$$

For simplicity, the first term in the expansion as is defined as  $V_0$ . This term is referred to as the coefficient for the empty cluster function. This substitution results in

$$\langle V(c_1, c_2, \dots) \rangle_{Unit} = V_0 + \sum_{\alpha \in A_1} V_\alpha m_\alpha \langle \phi(\bar{c}) \rangle_\alpha + \sum_{\alpha \in A_2} V_\alpha m_\alpha \langle \phi(\bar{c}) \rangle_\alpha + \dots \quad \text{Equation 31}$$

Equation 31 can be expressed in a more familiar form, by defining the following variables:

$$\begin{aligned} x_\alpha &= \langle \phi(\bar{c}) \rangle_\alpha \\ \beta_\alpha &= V_\alpha m_\alpha \\ x_0 &= 1 \\ \beta_0 &= V_0 \end{aligned} \quad \text{Equation 32}$$

By defining  $\vec{\beta}$  as the vector of all  $\beta_\alpha$  and  $\vec{x}$  as the matrix of all  $x_\alpha$ , Equation 31 can be written in vector form:

$$\langle V(c_1, c_2, \dots) \rangle_{Unit} = \vec{\beta} \cdot \vec{x} \quad \text{Equation 33}$$

Up to this point the expansion is exact. The correlations  $x_\alpha$  are properties of a particular decoration  $\bar{c}$ , and the coefficients  $\beta_\alpha$  are unknown. Approximate values for  $\beta_\alpha$  are typically derived from calculations performed using a relatively accurate method, such as density functional theory.

### III.1.1.1. Effective cluster interactions (ECI)

The coefficients  $V_\alpha$  are known as the effective cluster interactions, or ECI. The ECI for a given cluster function orbit indicates the degree to which interactions between the nuclei in the clusters affect the potential energy. To analyze the ECI, it is convenient to use orthonormal single-variable bases, where orthonormality is defined by:

$$\frac{\sum_{i=1}^{N_k} \phi_{a,k}(c_{k,i}) \phi_{b,k}(c_{k,i})}{N_k} = \delta_{ab} \quad \text{Equation 34}$$

where the sum is over all  $N_k$  possible values of the occupation variable  $c_k$ .

If the single-variable bases are orthonormal, it can be shown that the cluster functions must also be orthonormal:

$$\frac{\sum_{i=1}^{N_{\bar{c}}} \varphi_{\bar{j}}(\bar{c}_i) \varphi_{\bar{j}'}(\bar{c}_i)}{N_{\bar{c}}} = \frac{\prod_k \sum_{i=1}^{N_k} \phi_{j_k,k}(c_{k,i}) \phi_{j'_k,k}(c_{k,i})}{N_{\bar{c}}} = \frac{\prod_k N_k \delta_{j_k,j'_k}}{N_{\bar{c}}} = \delta_{\bar{j},\bar{j}'} \quad \text{Equation 35}$$

Equation 35 can be used to derive an exact expression for the ECI by projecting the known potential energies onto the cluster functions. For example, to calculate the value of  $V_{\bar{j}'}$ , where  $\bar{j}'$  represents a particular cluster function, first multiply both sides of Equation 25 by  $\varphi_{\bar{j}'}(\bar{c})$ :

$$V(\bar{c}) \varphi_{\bar{j}'}(\bar{c}) = V_{\bar{j}_0} \varphi_{\bar{j}_0}(\bar{c}) \varphi_{\bar{j}'}(\bar{c}) + \sum_{\bar{j} \in J_1} V_{\bar{j}} \varphi_{\bar{j}}(\bar{c}) \varphi_{\bar{j}'}(\bar{c}) + \sum_{\bar{j} \in J_2} V_{\bar{j}} \varphi_{\bar{j}}(\bar{c}) \varphi_{\bar{j}'}(\bar{c}) + \dots \quad \text{Equation 36}$$

Next sum over all possible decorations

$$\sum_{i=1}^{N_{\bar{c}}} V(\bar{c}_i) \varphi_{\bar{j}'}(\bar{c}_i) = V_{\bar{j}_0} \sum_{i=1}^{N_{\bar{c}}} \varphi_{\bar{j}_0}(\bar{c}_i) \varphi_{\bar{j}'}(\bar{c}_i) + \sum_{\bar{j} \in J_1} V_{\bar{j}} \sum_{i=1}^{N_{\bar{c}}} \varphi_{\bar{j}}(\bar{c}_i) \varphi_{\bar{j}'}(\bar{c}_i) + \sum_{\bar{j} \in J_2} V_{\bar{j}} \sum_{i=1}^{N_{\bar{c}}} \varphi_{\bar{j}}(\bar{c}_i) \varphi_{\bar{j}'}(\bar{c}_i) + \dots \quad \text{Equation 37}$$

The orthonormality condition (Equation 35) allows this expression to be simplified:

$$\sum_{i=1}^{N_{\bar{c}}} V(\bar{c}_i) \varphi_{\bar{j}'}(\bar{c}_i) = V_{\bar{j}_0} N_{\bar{c}} \delta_{\bar{j}_0,\bar{j}'} + \sum_{\bar{j} \in J_1} V_{\bar{j}} N_{\bar{c}} \delta_{\bar{j},\bar{j}'} + \sum_{\bar{j} \in J_2} V_{\bar{j}} N_{\bar{c}} \delta_{\bar{j},\bar{j}'} + \dots = V_{\bar{j}'} N_{\bar{c}} \quad \text{Equation 38}$$

Re-arranging the terms produces the desired result:

$$V_{\bar{j}'} = \frac{\sum_{i=1}^{N_{\bar{c}}} V(\bar{c}_i) \varphi_{\bar{j}'}(\bar{c}_i)}{N_{\bar{c}}} \quad \text{Equation 39}$$

Equation 39 is known as direct configuration averaging. Because it involves a sum over all possible site decorations, it generally cannot be used. However it does provide a better understanding of the ECI. Because  $\varphi_{j_0}(\bar{c})=1$ ,  $V_{j_0}$  is just the average potential energy over all decorations, and the remaining ECI represent the covariance between the potential energy and the cluster functions. The single-site terms (also known as “point” terms) are the main effects, and the multi-body cluster functions are interactions.

### III.1.2. A new approach to developing cluster expansions

The key to the successful application of the cluster expansion formalism is the determination of values for the ECI that accurately reproduce the values of the local minima in the potential energy surface. Physically, it is expected that the ECI will become smaller as the cluster functions include more sites. In addition, it is physically expected that the closer sites are to each other, the more the occupations of those sites will affect the potential energy. For these reasons, the ECI should be largest for cluster functions containing a small number of physically proximate sites.

In this chapter, this well-known insight will be used in conjunction with Bayes’ theorem to develop new methods for generating cluster expansions. The first half will focus on the theory and algorithms behind these methods, and in the second half these methods will be applied to different binary systems and evaluated relative to more traditional methods.

---

## III.2. Bayesian Regression

---

In this section Bayesian regression will be derived in the context of a cluster expansion, and the relationship to Tikhonov regularization [82] will be briefly discussed. More details on Bayesian regression can be found in refs. [83] and [84].

### III.2.1. Derivation of Bayesian regression

---

Let there be a set of input values and output values, where the  $i^{\text{th}}$  set of input values will be labeled  $\bar{x}_i$  and the  $i^{\text{th}}$  output value will be labeled  $y_i$ . Assume that the output value is an unknown function  $f_{opt}(\cdot)$  of the input values, plus some random error  $\tilde{e}_i$  (the tilde indicates that  $\tilde{e}_i$  is a random variable). The error  $\tilde{e}_i$  will be referred to as the noise.

$$y_i = f_{opt}(\bar{x}_i) + \tilde{e}_i \quad \text{Equation 40}$$

The noise is a random value drawn from a distribution  $n_i$  with mean 0 and variance  $\sigma_i^2$ .

Assuming the samples are independent, the probability of the training output can be calculated, given the training input and a trial function  $f(\cdot)$ .

$$P(f(X) + \tilde{e} = \bar{y} \mid X, \bar{y}, \bar{n}) = \prod_i P(y_i - f(\bar{x}_i) = \tilde{e}_i \mid \bar{x}_i, y_i, n_i) \quad \text{Equation 41}$$

where  $\bar{y}$  is a column vector which has  $y_i$  as the  $i^{\text{th}}$  element,  $\bar{n}$  and  $\tilde{e}$  are similarly defined, and  $X$  is a matrix with  $\bar{x}_i$  as the  $i^{\text{th}}$  row. The product is taken over all training samples. For the remainder of this section, it will be assumed that the training data ( $X$  and  $\bar{y}$ ) and the error distributions ( $\bar{n}$ ) are given, and explicit statement of this assumption will be dropped from the equations for clarity.

An intuitive way to estimate the hidden function  $f_{opt}(\cdot)$  might be to search for the function that maximizes  $P(f(X) + \tilde{e} = \bar{y})$ . Because the natural log is a monotonically increasing function, this is the same as maximizing  $\ln[P(f(X) + \tilde{e} = \bar{y})]$ . Combining this with Equation 41 produces the following expression:



$$\hat{f}_{erm}(\cdot) = \arg \max_{f(\cdot)} \sum_i \ln[P(y_i - f(\bar{x}_i) = \tilde{e}_i)] \quad \text{Equation 42}$$

where the carat symbol ( $\hat{\cdot}$ ) is used to indicate an estimated quantity.

Equation 42 is known as empirical risk minimization, with the empirical loss function defined by

$$I_S[f(\cdot)] = -\sum_i \ln[P(y_i - f(\bar{x}_i) = \tilde{e}_i)] \quad \text{Equation 43}$$

For example, suppose the distributions  $n_i$  are Gaussian, i.e.:

$$P(y_i - f(\bar{x}_i) = \tilde{e}_i | f(\cdot), \bar{x}_i, y_i, n_i) = \frac{1}{\sigma_i \sqrt{2\pi}} e^{-\frac{(y_i - f(\bar{x}_i))^2}{2\sigma_i^2}} \quad \text{Equation 44}$$

The loss function becomes

$$I_S[f(\cdot)] = \ln(\sigma_i \sqrt{2\pi}) + \frac{1}{2} \sum_i \frac{1}{\sigma_i^2} (y_i - f(\bar{x}_i))^2 \quad \text{Equation 45}$$

where C is some constant. This is just the squared loss function, where each sample is given a weight of  $\sigma_i^{-2}$ .

Depending on the function space from which functions are selected, empirical risk minimization may be ill-posed [82]. In other words, it may not satisfy the following conditions:

- 1) For every set of input and output values, there is a unique minimum to the loss function.
- 2) Small changes in the input and output values result in predictably small changes in the function that minimizes the loss function.

An alternative approach for selecting a function is to try to find the function that is most likely to be  $f_{opt}(\cdot)$  given the training data. Using Bayes' rule [85], the probability of a function  $f(\cdot)$  being  $f_{opt}(\cdot)$  given the training data is:

$$P(f(\cdot) = f_{opt}(\cdot) | f_{opt}(X) + \tilde{e} = \bar{y}) = \frac{P(f_{opt}(X) + \tilde{e} = \bar{y} | f(\cdot) = f_{opt}(\cdot)) P(f(\cdot) = f_{opt}(\cdot))}{P(f_{opt}(X) + \tilde{e} = \bar{y})} \quad \text{Equation 46}$$

The optimization problem now becomes

$$\hat{f}_{Bayes}(\cdot) = \arg \max_{f(\cdot)} \frac{P(f_{opt}(X) + \tilde{\epsilon} = \bar{y} | f(\cdot) = f_{opt}(\cdot)) P(f(\cdot) = f_{opt}(\cdot))}{P(f_{opt}(X) + \tilde{\epsilon} = \bar{y})} \quad \text{Equation 47}$$

Because the logarithm is monotonically increasing, Equation 47 can be re-written

$$\hat{f}_{Bayes}(\cdot) = \arg \max_{f(\cdot)} \left( \begin{array}{l} \ln[P(f_{opt}(X) + \tilde{\epsilon} = \bar{y} | f(\cdot) = f_{opt}(\cdot))] \\ + \ln[P(f(\cdot) = f_{opt}(\cdot))] \\ - \ln[P(f_{opt}(X) + \tilde{\epsilon} = \bar{y})] \end{array} \right) \quad \text{Equation 48}$$

The last term on the right is not dependent on  $\hat{f}(\cdot)$ , so it can be dropped.

$$\hat{f}_{Bayes}(\cdot) = \arg \max_{f(\cdot)} \left( \begin{array}{l} \ln[P(f_{opt}(X) + \tilde{\epsilon} = \bar{y} | f(\cdot) = f_{opt}(\cdot))] \\ + \ln[P(f(\cdot) = f_{opt}(\cdot))] \end{array} \right) \quad \text{Equation 49}$$

Combining Equation 49 with Equation 41 and Equation 43 produces:

$$\hat{f}_{Bayes}(\cdot) = \arg \min_{f(\cdot)} (I_S[f(\cdot)] - \ln[P(f(\cdot) = f_{opt}(\cdot))]) \quad \text{Equation 50}$$

If  $-\ln[P(f(\cdot) = f_{opt}(\cdot))]$  is appropriately defined (e.g. a norm of the function space containing  $f(\cdot)$ ), Equation 50 is equivalent to Tikhonov regularization. Such a problem is guaranteed to be well-posed [82].

The advantage to deriving regularization this way is that it gives some meaning to the regularization term, in the form of the prior distribution  $P(f(\cdot) = f_{opt}(\cdot))$ . The prior distribution is a probability distribution over functions that assigns to every function  $f(\cdot)$  a probability that it is  $f(\cdot) = f_{opt}(\cdot)$ , before the training data is taken into consideration. The distribution  $P(f(\cdot) = f_{opt}(\cdot) | f_{opt}(X) + \tilde{\epsilon} = \bar{y})$  on the left-hand side of Equation 46 is known as the posterior distribution. Like the prior distribution, this is a distribution over functions that assigns to every function  $f(\cdot)$  a probability that  $f(\cdot) = f_{opt}(\cdot)$ . The difference is that using Bayes' rule, the posterior distribution takes into account the training data. The posterior distribution provides a useful way to evaluate candidate functions, but it is dependent on the prior guess.<sup>4</sup> In the following sections prior distributions will be examined in more detail.

<sup>4</sup> In this text the terms "prior distribution" and "posterior distribution" will occasionally be shortened to "prior" and "posterior" respectively.

### III.2.2. Prior distributions

---

The choice of prior distribution is an important factor in determining how quickly the predicted function,  $\hat{f}_{Bayes}(\cdot)$ , converges to  $f_{opt}(\cdot)$ . The important attributes to consider when choosing a prior distribution are the mean of the prior, the shape of the prior, and the width of the prior. In the following sections each of these attributes will be investigated both generally and in the context of cluster expansions.

#### III.2.2.1. The mean of the prior distribution

Physical insight might provide a sense of what values the ECI are expected to have. Perhaps typical ECI values for similar systems are known, or approximations for the ECI have been analytically calculated. These expected ECI values will be used as the mean for the prior distribution. Specifically, this is written as

$$\tilde{f}(\cdot) = f_p(\cdot) + \tilde{f}_\Delta(\cdot) \quad \text{Equation 51}$$

where  $\tilde{f}_p(\cdot)$  is a function drawn from the prior distribution of functions,  $f_p(\cdot)$  is the mean of the prior distribution of functions, and  $\tilde{f}_\Delta(\cdot)$  represents the variation from the mean.

The mean of the prior distribution can be treated in a straightforward way if it is assumed that the loss functional  $I_S[f(\cdot)]$  can be written as  $I_S[f(X) - \bar{y}]$ . Equation 50 becomes:

$$\hat{f}_{Bayes}(\cdot) = \arg \min_{f(\cdot)} \left( I_S[f(X) - \bar{y}] - \ln[P(f(\cdot) = f_{opt}(\cdot))] \right) \quad \text{Equation 52}$$

If Equation 51 is substituted into Equation 52, the fact that  $f_p(\cdot)$  is a known function can be used to write the minimization in terms of  $\tilde{f}_\Delta(\cdot)$ :

$$\hat{f}_{Bayes}(\cdot) = f_p(\cdot) + \arg \min_{\tilde{f}_\Delta(\cdot)} \left( I_S[f_p(X) + \tilde{f}_\Delta(X) - \bar{y}] - \ln[P(f_p(\cdot) + \tilde{f}_\Delta(\cdot) = f_{opt}(\cdot))] \right) \quad \text{Equation 53}$$

Making the following variable substitution:

$$\bar{y}_\Delta = \bar{y} - f_p(X) \quad \text{Equation 54}$$

and re-arranging Equation 53 produces

$$\hat{f}_{Bayes}(\cdot) = f_p(\cdot) + \arg \min_{f_\Delta(\cdot)} \left( I_S[f_\Delta(X) - \bar{y}_\Delta] - \ln[P(f_\Delta(\cdot) = f_{opt}(\cdot) - f_p(\cdot))] \right) \quad \text{Equation 55}$$

The minimization problem in Equation 55 is equivalent to the one in Equation 52 with variable substitutions. The training output  $\bar{y}$  has been replaced with  $\bar{y}_\Delta$ , and the prior distribution with mean  $f_p(\cdot)$  has been replaced by a distribution with a mean of zero. In the following sections the shape and the width of the prior distribution will be investigated, under the assumption that the variables have been transformed so that the mean of the distribution is zero.

### III.2.2.2. The shape of the prior distribution

Given that the problem has been transformed so that the prior distribution has zero mean, it should be expected that the constant zero function has the highest prior probability. The prior should also ideally be chosen so that it is computationally convenient. In the following two sections, two shapes for prior distributions for cluster expansions that meet these criteria will be investigated.

#### III.2.2.2.1. The Gaussian prior

A common form for the prior distribution might be:

$$\|\tilde{f}_\Delta(\cdot)\| \sim n(\mathbf{0}, \sigma_p^2) \quad \text{Equation 56}$$

where  $\|\cdot\|$  is a function space norm and  $n(\mathbf{0}, \sigma_p^2)$  is a normal distribution with mean  $\mathbf{0}$  and variance  $\sigma_p^2$ . The prior distribution can be written as

$$P(f_\Delta(\cdot) = f_{opt}(\cdot) - f_p(\cdot)) = \frac{1}{\sigma_p \sqrt{2\pi}} e^{-\frac{\|f_\Delta(\cdot)\|^2}{2\sigma_p^2}} \quad \text{Equation 57}$$

Plugging Equation 57 into Equation 55 produces:

$$\hat{f}_{Bayes}(\cdot) = f_p(\cdot) + \arg \min_{f_\Delta(\cdot)} \left( I_S[f_\Delta(X) - \bar{y}_\Delta] + \frac{\|f_\Delta(\cdot)\|^2}{2\sigma_p^2} + \ln(\sigma_p \sqrt{2\pi}) \right) \quad \text{Equation 58}$$

Because the last term on the right does not depend on  $\hat{f}(\cdot)$ , it can be dropped.

Defining  $\lambda_p = \frac{1}{\sigma_p^2}$ , Equation 58 can be written as

$$\hat{f}_{Bayes}(\cdot) = f_p(\cdot) + \arg \min_{f_\Delta(\cdot)} \left( I_S[f_\Delta(X) - \bar{y}_\Delta] + \frac{1}{2} \lambda_p \|f_\Delta(\cdot)\|^2 \right) \quad \text{Equation 59}$$

Equation 59 can be applied to the cluster expansion. For example, consider using the  $L_2$  norm for a cluster expansion. The  $L_2$  norm for a function defined over the discrete space of occupation variables is defined by

$$\|V(\cdot)\|_{L_2} = \sqrt{\frac{\sum_{i=1}^{N_\varepsilon} V(\bar{c}_i)^2}{N_\varepsilon}} \quad \text{Equation 60}$$

Substituting Equation 25 into Equation 60 produces

$$\|V(\cdot)\|_{L_2} = \sqrt{\frac{\sum_{i=1}^{N_\varepsilon} \sum_j V_j V_{j'} \varphi_j(\bar{c}_i) \varphi_{j'}(\bar{c}_i)}{N_\varepsilon}} \quad \text{Equation 61}$$

Equation 61 can be re-arranged as:

$$\|V(\cdot)\|_{L_2} = \sqrt{\sum_j \sum_{j'} V_j V_{j'} \frac{\sum_{i=1}^{N_\varepsilon} \varphi_j(\bar{c}_i) \varphi_{j'}(\bar{c}_i)}{N_\varepsilon}} \quad \text{Equation 62}$$

If an orthonormal basis is used, Equation 35 can be combined with Equation 62 to arrive at the final result:

$$\|V(\cdot)\|_{L_2} = \sqrt{\sum_j V_j^2} \quad \text{Equation 63}$$

Equation 25 and Equation 63 can be combined to express the regularization term as a function of the cluster expansion ECI:

$$\hat{V}_{Bayes}(\cdot) = V_p(\cdot) + \arg \min_{V_\Delta(\cdot)} \left( I_S[V_\Delta(X) - \bar{y}_\Delta] + \frac{1}{2} \lambda_p \sum_j V_{j,\Delta}^2 \right) \quad \text{Equation 64}$$

where  $V_{j,\Delta} = V_j - V_{j,p}$  and  $V_{j,p}$  is the mean of the prior distribution for  $V_j$ .

If Gaussian error distributions are assumed, Equation 45 can be combined with Equation 64 to produce

$$\hat{V}_{Bayes}(\cdot) = V_p(\cdot) + \arg \min_{V_\Delta(\cdot)} \left( \frac{1}{2} \sum_i \frac{1}{\sigma_i^2} (V_\Delta(\bar{c}_i) - y_{\Delta,i})^2 + \frac{1}{2} \lambda_p \sum_{\bar{j}} V_{\bar{j},\Delta}^2 \right) \quad \text{Equation 65}$$

One problem with this approach is that the regularization term penalizes errors in all ECI equivalently. Physically, it is expected that ECI representing interactions between sites that are far apart, or interactions between a large number of sites, should be very small. The root of this problem can be seen by combining Equation 63 with Equation 57 to produce an expression for the prior distribution of functions, assuming a Gaussian distribution of the L<sub>2</sub> norm:

$$P(V_\Delta(\cdot) = V_{opt}(\cdot) - V_p(\cdot)) = \frac{1}{\sigma_p \sqrt{2\pi}} e^{-\frac{\sum_{\bar{j}} V_{\bar{j},\Delta}^2}{2\sigma_p^2}} \quad \text{Equation 66}$$

The sum in the exponent can be expanded into a product:

$$P(V_\Delta(\cdot) = V_{opt}(\cdot) - V_p(\cdot)) = \frac{1}{\sigma_p \sqrt{2\pi}} \prod_{\bar{j}} e^{-\frac{V_{\bar{j},\Delta}^2}{2\sigma_p^2}} \quad \text{Equation 67}$$

Within a normalization constant, this can be written as:

$$P(V_\Delta(\cdot) = V_{opt}(\cdot) - V_p(\cdot)) \propto \prod_{\bar{j}} P(V_{\bar{j},\Delta}) \quad \text{Equation 68}$$

where

$$P(V_{\bar{j},\Delta}) = \frac{1}{\sigma_p \sqrt{2\pi}} e^{-\frac{V_{\bar{j},\Delta}^2}{2\sigma_p^2}} = \frac{1}{\sigma_p \sqrt{2\pi}} e^{-\frac{(V_{\bar{j}} - V_{\bar{j},p})^2}{2\sigma_p^2}} \quad \text{Equation 69}$$

Thus using the L<sub>2</sub> norm is equivalent to expressing the prior probability of a particular function as the product of independent prior probabilities on the individual ECI. The prior distribution for a given ECI  $V_{\bar{j}}$  is Gaussian, with mean  $V_{\bar{j},p}$  and variance  $\sigma_p^2$ . It is because the width of the prior distribution for each cluster function  $\bar{j}$  is identical that the regularization is physically unrealistic. A more realistic prior would have narrower distributions for the variables for which the difference  $V_{\bar{j}} - V_{\bar{j},p}$  is expected to be smaller. It is straightforward to create such a prior by replacing the universal  $\sigma_p$  with  $\sigma_{\bar{j}}$ , which is dependent on the cluster function represented by  $\bar{j}$ :

$$P(V_{\Delta}(\cdot) = V_{opt}(\cdot) - V_p(\cdot)) \propto \prod_j e^{\frac{-V_{j,\Delta}^2}{2\sigma_j^2}} \quad \text{Equation 70}$$

The optimization problem can now be expressed as

$$\hat{V}_{Bayes}(\cdot) = V_p(\cdot) + \arg \min_{V_{\Delta}(\cdot)} \left( \frac{1}{2} \sum_i \frac{1}{\sigma_i^2} (V_{\Delta}(\bar{c}_i) - y_{\Delta,i})^2 + \frac{1}{2} \sum_j \lambda_j V_{j,\Delta}^2 \right) \quad \text{Equation 71}$$

where

$$\lambda_j = \frac{1}{\sigma_j^2} \quad \text{Equation 72}$$

If the average potential energy is of interest, prior distributions over the correlation coefficients  $\beta_{\alpha}$  can be similarly defined:

$$\hat{V}_{Bayes}(\cdot) = V_p(\cdot) + \arg \min_{V_{\Delta}(\cdot)} \left( \frac{1}{2} \sum_i \frac{1}{\sigma_i^2} (V_{\Delta}(\bar{c}_i) - y_{\Delta,i})^2 + \frac{1}{2} \sum_{\alpha} \lambda_{\alpha} \beta_{\alpha,\Delta}^2 \right) \quad \text{Equation 73}$$

where the sum is now over cluster function orbits  $\alpha$  instead of the individual cluster functions. This can be expressed in matrix-vector notation:

$$\hat{V}_{Bayes}(\bar{x}) = V_p(\bar{x}) + \bar{x} \cdot \arg \min_{\bar{\beta}_{\Delta}} \left( \frac{1}{2} (\bar{y}_{\Delta} - X\bar{\beta}_{\Delta})^T W_v (\bar{y}_{\Delta} - X\bar{\beta}_{\Delta}) + \bar{\beta}_{\Delta}^T \Lambda_v \bar{\beta}_{\Delta} \right) \quad \text{Equation 74}$$

where the matrix  $W_v$  is diagonal, with elements defined by:

$$W_{v,ii} = \frac{1}{\sigma_i^2} \quad \text{Equation 75}$$

and  $\Lambda_v$  is a diagonal matrix with elements defined by

$$\Lambda_{v,\alpha\alpha} = \lambda_{\alpha} = \frac{1}{\sigma_{\alpha}^2} \quad \text{Equation 76}$$

To know the values for the noise variances,  $\sigma_i^2$ , would require knowing how well the model can reproduce the actual data. In general, this is not known, which makes the weight matrix  $W_v$  impossible to specify. However, the desired magnitude of the noise variance for a given sample relative to other samples can be specified. This is accomplished by defining a matrix  $W$  by

$$W_{ii} = \frac{\sigma_w^2}{\sigma_i^2} \quad \text{Equation 77}$$

for some unknown variance  $\sigma_w^2$ . This has the desirable property that

$$\frac{W_{ii}}{W_{jj}} = \frac{W_{v,ii}}{W_{v,jj}} = \frac{\sigma_j^2}{\sigma_i^2} \quad \text{Equation 78}$$

but no assumptions about the magnitude of the noise variances are necessary. The term  $\sigma_w^2$  can be thought of as the uniform variance of the noise for a regression performed with input values  $XW^{1/2}$  and output values  $W^{1/2}\bar{y}_\Delta$ .

If a matrix  $\Lambda$  is defined by

$$\Lambda_{\alpha\alpha} = \frac{\sigma_w^2}{\sigma_\alpha^2} \quad \text{Equation 79}$$

Equation 74 can be re-written as

$$\hat{V}_{Bayes}(\bar{x}) = V_p(\bar{x}) + \bar{x} \cdot \arg \min_{\bar{\beta}_\Delta} \left( \frac{1}{2} (\bar{y}_\Delta - X\bar{\beta}_\Delta)^T W (\bar{y}_\Delta - X\bar{\beta}_\Delta) + \bar{\beta}_\Delta^T \Lambda \bar{\beta}_\Delta \right) \quad \text{Equation 80}$$

Equation 80 is a weighted regularized least-squares regression on the output vector  $\bar{y}_\Delta$  with regularization matrix  $\Lambda$ . The solution for this problem is known, and is given by:

$$\hat{V}_{Bayes}(\bar{x}) = V_p(\bar{x}) + \bar{x} (X^T W X + \Lambda)^{-1} X^T W \bar{y}_\Delta \quad \text{Equation 81}$$

Through Equation 77 and Equation 79, all of the unknowns from the probability distributions have been transferred to the matrix  $\Lambda$ . Later ways to find good values for  $\Lambda$  will be discussed.

### III.2.2.2.2. The Laplace prior

An alternative to Gaussian priors on the coefficients for cluster functions would be to use a Laplace prior. The prior distribution is expressed as:

$$P(\beta_{\alpha,\Delta}) = \frac{1}{\sigma_\alpha \sqrt{2}} e^{-\frac{|\beta_{\alpha,\Delta}| \sqrt{2}}{\sigma_\alpha}} = \frac{1}{\sigma_\alpha \sqrt{2}} e^{-\frac{|\beta_\alpha - \beta_{\alpha,p}| \sqrt{2}}{\sigma_\alpha}} \quad \text{Equation 82}$$

The optimization problem becomes



$$\hat{V}_{Bayes}(\cdot) = V_p(\cdot) + \arg \min_{V_\Delta(\cdot)} \left( \frac{1}{2} \sum_i \frac{1}{\sigma_i^2} (V_\Delta(\bar{c}_i) - y_{\Delta,i})^2 + \sum_\alpha \lambda_\alpha |\beta_{\alpha,\Delta}| \right) \quad \text{Equation 83}$$

where

$$\lambda_\alpha = \frac{\sqrt{2}}{\sigma_\alpha} \quad \text{Equation 84}$$

Because of the discontinuity in the first derivative of the absolute value, it is more difficult to efficiently solve this problem. This problem can be written in matrix-vector notation as follows:

$$\hat{V}_{Bayes}(\bar{x}) = V_p(\bar{x}) + \bar{x} \cdot \arg \min_{\bar{\beta}_\Delta} \left( \frac{1}{2} (\bar{y}_\Delta - X\bar{\beta}_\Delta)^T W (\bar{y}_\Delta - X\bar{\beta}_\Delta) + \bar{s}_{\beta_\Delta}^T \Lambda \bar{\beta}_\Delta \right) \quad \text{Equation 85}$$

where  $\bar{s}_{\beta_\Delta}$  is a vector whose *ith* has an absolute value of 1 and the same sign as the *ith* element of  $\bar{\beta}_\Delta$ . Under the assumption that all diagonal elements of  $\Lambda$  are non-zero, the following variable transformations can be made:

$$\begin{aligned} \bar{\beta}_{\Delta,T} &= \Lambda \bar{\beta}_\Delta \\ X_T &= W^{1/2} X \Lambda^{-1} \\ \bar{y}_T &= W^{1/2} \bar{y}_\Delta \end{aligned} \quad \text{Equation 86}$$

These transformations produce the following minimization problem:

$$\hat{V}_{Bayes}(\bar{x}) = V_p(\bar{x}) + \bar{x} \cdot \arg \min_{\bar{\beta}_{\Delta,T}} \left( \frac{1}{2} \|\bar{y}_T - X_T \bar{\beta}_{\Delta,T}\|^2 + (\bar{s}_{\beta_{\Delta,T}} \cdot \bar{\beta}_{\Delta,T}) \right) \quad \text{Equation 87}$$

using the fact that  $\bar{s}_{\beta_{\Delta,T}} = \bar{s}_{\beta_\Delta}$  as long as all diagonal elements of  $\Lambda$  are positive. The regression method that is expressed by Equation 87 is commonly known as LASSO (least absolute shrinkage and selection operator) [86]. One of the advantages to this method is that in general, the minimizing function will have only a finite set of non-zero ECI. Thus the optimal set of cluster function orbits to be included in the fit are automatically selected. Efron et al. recently produced an algorithm that finds the minimum in Equation 87 with the same computational efficiency as a least-squares fit [87]. In the discussion of their algorithm, they assume that the input data has been transformed to fit the following assumptions:

$$\sum_i y_i = 0 \quad \text{Equation 88}$$

$$\begin{aligned}\forall j \quad \sum_i x_{i,j} &= 0 \\ \forall j \quad \sum_i x_{i,j}^2 &= 1\end{aligned}$$

These assumptions are restrictive, as they prevent regularization of the constant term and the last assumption is incompatible with the second transformation in Equation 86. However, none of these assumptions are actually required for the derivation of the algorithm; they are only required for analysis and comparisons to similar algorithms.

The algorithm as presented by Efron et al., applied to the transformed variables, can be summarized as follows:

- 1) Let  $\bar{\beta}_{\Delta,T} = \bar{\mathbf{0}}$ , a vector of all zeroes.
- 2) Let  $\bar{c} = X_T^T (\bar{y}_T - X_T \bar{\beta}_{\Delta,T})$ .
- 3) Let  $C = \max(|c_j|)$ .
- 4) Let  $A = \{j : |c_j| = C\}$ .
- 5) Let  $s_i = \begin{cases} 1: \hat{c}_{A_i} > 0 \\ -1: \hat{c}_{A_i} < 0 \end{cases}$  where  $A_i$  is the  $i$ th element of  $A$ .
- 6) Let  $X_{T,A}$  be the matrix composed of columns  $s_i X_{T,A_i}$  where  $X_{T,j}$  is the  $j$ th column of  $X_T$  and  $A_i$  is the  $i$ th element of  $A$ .
- 7) Let  $A_A = \frac{1}{\sqrt{\bar{\mathbf{1}}_A^T (X_{T,A}^T X_{T,A})^{-1} \bar{\mathbf{1}}_A}}$  where  $\bar{\mathbf{1}}_A$  is a vector of all 1's with length  $|A|$
- 8) Let  $\bar{w}_A = A_A (X_{T,A}^T X_{T,A})^{-1} \bar{\mathbf{1}}_A$ .
- 9) Let  $\bar{a} = X_T^T X_{T,A} \bar{w}_A$ .
- 10) Let  $\hat{\gamma} = \min_{j \in A}^+ \left( \frac{C - c_j}{A_A - a_j}, \frac{C + c_j}{A_A + a_j} \right)$ , where  $\min^+$  means the minimum is over only positive elements.
- 11) Let  $d_j = \begin{cases} s_i w_i : j = A_i \\ 0 : j \notin A \end{cases}$ , where  $A_i$  is the  $i$ th element of  $A$ .

- 12) Let  $\gamma_j = \frac{-(\beta_{\Delta,T})_j}{d_j}$ .
- 13) If there are some  $j \in A$  for which  $\gamma_j > 0$ , then
- a. Let  $\tilde{j} = \arg \min_{j \in A: \gamma_j > 0} (\gamma_j)$ .
  - b. If  $\gamma_{\tilde{j}} < \hat{\gamma}$ , then
    - i. Remove  $\tilde{j}$  from  $A$ .
    - ii. Set  $\hat{\gamma} = \gamma_{\tilde{j}}$ .
- 14) Let  $\bar{\beta}_{\Delta,T} = \bar{\beta}_{\Delta,T} + \hat{\gamma} \bar{d}$ .
- 15) Goto 2).

There is one thing missing from this algorithm: a break condition. To derive this condition, start with Equation 5.22 from reference [87]:

$$-S_A X_{T,A}^T (\bar{y}_T - X_{T,A} S_A \bar{\beta}_{\min,\Delta,T}) + S_A \bar{1}_A = 0 \quad \text{Equation 89}$$

where  $S_A$  is a diagonal matrix whose  $i$ th diagonal element is  $s_i$ ,  $\bar{\beta}_{\min,\Delta,T}$  is the minimizing set of coefficients as defined in Equation 87, and the fact that  $\lambda = 1$  in Equation 87 has been used. Using the definitions given in the above algorithm, this formula can be written as

$$-\bar{c}_{\min} + S_A \bar{1}_A = 0 \quad \text{Equation 90}$$

where  $\bar{c}_{\min}$  is the vector  $\bar{c}$  corresponding to the minimizing solution. Re-arranging and pre-multiplying by  $S_A$  yields

$$\bar{1}_A = \bar{C}_{\min} \quad \text{Equation 91}$$

where  $\bar{C}_{\min}$  is a vector of length  $|A|$  whose elements are given by  $C_{\min,j} = |c_{\min,j}|$ . This shows that the algorithm reaches the minimizing solution when  $C = 1$ . Equation 2.16 in reference [87] says that for each step,  $C$  can be expressed as the following function of  $\gamma$ :

$$C(\gamma) = C_0 - \gamma A_A \quad \text{Equation 92}$$

where  $C_0$  is the value of  $C$  at the beginning of the step. The minimizing value of  $\gamma$  for a given step can therefore be expressed as:

$$\gamma_{\min} = (C_0 - 1)A_A^{-1} \quad \text{Equation 93}$$

Equation 93 provides the break condition for the algorithm . Steps 14) and 15) are replaced with:

- 14) Let  $\gamma_{\min} = (C - 1)A_A^{-1}$
- 15) If  $\gamma_{\min} < \hat{\gamma}$  then
  - a.  $\bar{\beta}_{\Delta,T} = \bar{\beta}_{\Delta,T} + \gamma_{\min} \bar{d}$
  - b. End.
- 16) Let  $\bar{\beta}_{\Delta,T} = \bar{\beta}_{\Delta,T} + \hat{\gamma} \bar{d}$ .
- 17) Goto 2).

A limitation of this algorithm is that all diagonal elements of  $\Lambda$  must be non-zero. Otherwise, the columns of  $X_T$  corresponding to the elements with value of zero would be infinite. It is possible to get around this restriction by re-writing the algorithm in a way that never involves infinity. Start with the observation that any variables for which  $\Lambda_{kk} = 0$  should be automatically included in the “active set”  $A$  because there is absolutely no penalty for doing so. This is equivalent to noting that their correlations,  $c_j$ , will be infinite as long as they are non-zero. Therefore the LASSO solution must always be one that forces the correlations for these variables to be zero. This is done by including these variables in  $A$  and ensuring that every “update vector”  $u$  is orthogonal to the columns representing variables for which  $\Lambda_{kk} = 0$ .

Once these variables are included in  $A$ , the rest of the algorithm can be written in terms of terms of  $X_W$ , defined as

$$X_W = W^{1/2} X \quad \text{Equation 94}$$

Including some optimizations, the new algorithm becomes:

- 1) Define the set of positive-regularized variables as  $PR = \{j : \Lambda_{jj} > 0\}$ . Let the its complement, the set of zero-regularized variables, be  $ZR$ . For  $j \in ZR$ , let  $\beta_{\Delta_j}$  be the coefficient determined using a least-squares fit to the variables in  $ZR$ . For  $j \in PR$ , let  $\beta_{\Delta_j} = 0$ .
- 2) Let  $\bar{c} = X_w^T (\bar{y}_T - X_w \bar{\beta}_\Delta)$ .
- 3) Let  $C = \max_{j \in PR} \left( \frac{|c_j|}{\Lambda_{jj}} \right)$
- 4) Let  $A = \left\{ j : \frac{|c_j|}{\Lambda_{jj}} = C \right\} \cup ZR$ .
- 5) Let  $s_i = \begin{cases} 1: \hat{c}_{A_i} \geq 0 \\ -1: \hat{c}_{A_i} < 0 \end{cases}$  where  $A_i$  is the  $i$ th element of  $A$ .
- 6) Let  $X_{w,A}$  be the matrix composed of columns  $X_{w,A_i}$  where  $X_{w,j}$  is the  $j$ th column of  $X_w$  and  $A_i$  is the  $i$ th element of  $A$ .
- 7) Let  $A_A = \frac{1}{\sqrt{\bar{\lambda}_A^T (X_{w,A}^T X_{w,A})^{-1} \bar{\lambda}_A}}$  where  $\bar{\lambda}_A$  is a vector of length  $|A|$  and the  $i$ th element of  $\bar{\lambda}_A$  is  $s_i \Lambda_{A_i, A_i}$ .
- 8) Let  $\bar{w}_A = A_A (X_{w,A}^T X_{w,A})^{-1} \bar{\lambda}_A$ .
- 9) Let  $\bar{a} = X^T X_{w,A} \bar{w}_A$ .
- 10) Let  $\hat{\gamma} = \min_{j \in A}^+ \left( \frac{C - \frac{c_j}{\Lambda_{jj}}}{A_A - \frac{a_j}{\Lambda_{jj}}}, \frac{C + \frac{c_j}{\Lambda_{jj}}}{A_A + \frac{a_j}{\Lambda_{jj}}} \right)$ , where  $\min^+$  means the minimum is over only positive elements.
- 11) Let  $d_j = \begin{cases} \Lambda_{jj} s_i w_i : j = A_i \\ 0 : j \notin A \end{cases}$ , where  $A_i$  is the  $i$ th element of  $A$ .

- 12) Let  $\gamma_j = \frac{-(\beta_{\Delta,T})_j}{d_j}$  for  $j \in A$ .
- 13) If there are some  $j \in A$  for which  $\gamma_j > 0$ , then
- a. Let  $\tilde{j} = \underset{j \in A: \gamma_j > 0}{\operatorname{argmin}}(\gamma_j)$ .
  - b. If  $\gamma_{\tilde{j}} < \hat{\gamma}$ , then
    - i. Remove  $\tilde{j}$  from  $A$ .
    - ii. Set  $\hat{\gamma} = \gamma_{\tilde{j}}$ .
- 14) Let  $\gamma_{\min} = (C-1)A_A^{-1}$
- 15) If  $\gamma_{\min} < \hat{\gamma}$  then
- a.  $\bar{\beta}_{\Delta,T} = \beta_{\Delta,T} + \gamma_{\min} \bar{d}$
  - b. End.
- 16) Let  $\bar{\beta}_{\Delta,T} = \beta_{\Delta,T} + \hat{\gamma} \bar{d}$ .
- 17) Goto 2).

This algorithm can be made more efficient by noting that each update of the matrix  $(X_{W,A}^T X_{W,A})^{-1}$  can be accomplished in  $O(|A|^2)$  time using one of a variety of incremental update methods for matrix inverses.

### III.2.2.3. The width of the prior distribution

In previous sections, solutions for Equation 50 were examined for two different shapes for the prior distribution: a Gaussian shape and a Laplace shape. With the transformed variables, the mean of the prior distributions will be zero. If only Laplace and Gaussian shapes are considered, this leaves only the problem of determining  $\sigma_\alpha$ , the width of the prior distribution for each variable. In this section different generating functions for  $\sigma_\alpha$  will be discussed. In the next section ways parameterize these generating functions will be investigated.

### III.2.2.3.1. Cluster selection

In theory, for a periodic system there are an infinite number of cluster orbits that can be included in a cluster expansion. Determining the ECI for all of these cluster orbits is a computationally intractable problem, but this is usually not necessary. One of the advantages to the cluster expansion is that for clusters of atoms that are far apart, or for clusters of many atoms, the ECI are expected to be small. In a Bayesian sense, this means that the width of the prior distributions for the ECI for these clusters should be small. At some point, the prior becomes so small that  $\sigma_\alpha = 0$  is a reasonable approximation for cluster orbits containing clusters larger than a given size. This approximation allows the exclusion of large clusters from consideration in the fit, reducing the problem to one of a finite number of variables. The cluster orbits for which  $\sigma_\alpha$  is not set to zero will be referred to as “candidate cluster orbits.”

From the set of candidate cluster orbits, a common strategy for determining the ECI is to choose a subset of the candidate orbits and perform a least-squares fit on this subset. This is equivalent to choosing a prior distribution with the constraint that  $\sigma_\alpha \in \{0, \infty\}$ . Although this is of course an unrealistic constraint, it is computationally and conceptually simple and is widely used to develop cluster expansions. The parameters for this prior distribution are binary variables that determine whether each cluster orbit is included in the fit.

In the following sections forms for prior distributions for which the prior widths can take on a continuous range of values will be investigated. In all forms  $\sigma_\alpha = \infty$  is used for the empty cluster function.

### III.2.2.3.2. Constant width

A common strategy for determining the prior distribution is to assume the same prior distribution on the ECI for each cluster, so that  $\Lambda = \lambda I$  for some  $\lambda$ . It is a simple and popular form for the prior, because it only requires the determination of one parameter.

This prior is equivalent to saying that the expected magnitude of the ECIs for all candidate clusters is the same. In the case of a cluster expansion, this is an unrealistic statement. It suggests that the ECI for larger clusters are expected to get smaller not in a

continuous manner, but as a step function. For candidate clusters the width of the prior is a constant  $\sigma_\lambda$ , and beyond some cutoff point the width of the prior distributions is set to 0. Physically, the width of the prior distributions should shrink more continuously.

### III.2.2.3.3. Independent widths

An alternative to using a constant width is to set the width of the prior distribution for each variable independently. This is equivalent to trying to find the best set of non-negative diagonal elements for  $\Lambda$ . This form of a prior, applied to cluster expansions, was recently suggested by Drautz and Díaz-Ortiz [88].

One of the challenges with using independent widths is that determining the “best”  $\Lambda$  can be computationally expensive. It is an optimization problem on a continuous  $m$ -dimensional search space, where  $m$  is the number of candidate clusters. It is possible to use a local minimization algorithm to find a good value, but this makes the result dependent on the initial conditions.

Another concern with this approach is that with so many degrees of freedom available to choose the prior distribution, there is a risk of over-fitting the data. The parameters of the prior distribution are usually chosen based on a given set of training data, and there is a risk that a given set of parameters work well for the training data but not in general. This risk increases as the number of degrees of freedom in the prior parameter space is increased. Although this might be a reasonable approach for a system for which there is no insight into the expected magnitude of the ECI, it is perhaps excessive for the cluster expansion. There is a general physical expectation for what a cluster expansion should look like: smaller clusters are more likely to have large ECI than larger clusters. In the next two sections distributions that take advantage of this insight will be investigated.

### III.2.2.3.4. The exponential width function

For the cluster expansion, it makes sense to incorporate physical insight into the width of the prior distribution:

- 1) As the number of sites in a cluster increases, the width of the prior distribution on the ECI for that cluster should decrease.



2) As the distance between sites in a cluster increases, the width of the prior distribution on the ECI for that cluster should decrease.

3) The cluster expansion should converge with respect to the size of the clusters in the expansion. To define convergence, let  $V_{\Delta, cut}(\cdot)$  be the cluster expansion with the “true” ECI (as defined in Equation 39) for all clusters with fewer than  $n_{cut}$  sites and no two sites further apart than  $r_{cut}$ , and ECI of zero for all other clusters. Convergence is then defined as follows:

$$\forall \varepsilon > 0, \exists \{n_{cut}, r_{cut}\} \text{ s.t. } E\left(\left(V_{\Delta}(\bar{c}) - V_{\Delta, cut}(\bar{c})\right)^2\right) < \varepsilon \quad \text{Equation 95}$$

By this definition, every cluster expansion for a material with a finite number of sites converges because the cluster expansion that includes all possible clusters can perfectly reproduce  $V_{\Delta}(\bar{c})$ . For this reason, it will be assumed in the rest of this discussion that systems with an infinite number of sites are being considered.

Equation 25 can be substituted into Equation 95 to get

$$\forall \varepsilon > 0, \exists \{n_{cut}, r_{cut}\} \text{ s.t. } E\left(\left(\sum_{\bar{j} \in J_{cut}} V_{\bar{j}, \Delta} \phi_{\bar{j}}(\bar{c})\right)^2\right) < \varepsilon \quad \text{Equation 96}$$

where  $J_{cut}$  is the set of clusters with non-zero ECI in  $V_{\Delta, cut}(\cdot)$ , and  $V_{\bar{j}, \Delta}$  is used instead of  $V_{\bar{j}}$  as a reminder that the mean of the prior expectations of the ECI is zero. The structure-dependent part of the expectation value can be factored out to get

$$\forall \varepsilon > 0, \exists \{n_{cut}, r_{cut}\} \text{ s.t. } \frac{\sum_i E\left(\left(\sum_{\bar{j} \in J_{cut}} V_{\bar{j}, \Delta} \phi_{\bar{j}}(\bar{c}_i)\right)\right)^2}{N_{\bar{c}}} < \varepsilon \quad \text{Equation 97}$$

where the sum is over structures.

The square can be expanded to get

$$\forall \varepsilon > 0, \exists \{n_{cut}, r_{cut}\} \text{ s.t. } \frac{\sum_i \sum_{\bar{j} \in J_{cut}} \sum_{\bar{j}' \in J_{cut}} V_{\bar{j}, \Delta} V_{\bar{j}', \Delta} \phi_{\bar{j}}(\bar{c}) \phi_{\bar{j}'}(\bar{c})}{N_{\bar{c}}} < \varepsilon \quad \text{Equation 98}$$

Re-arranging Equation 98 produces the following expression:

$$\forall \varepsilon > 0, \exists \{n_{cut}, r_{cut}\} \text{ s.t. } \sum_{\bar{j} \in J_{cut}} \sum_{\bar{j}' \in J_{cut}} E \left( V_{\bar{j}, \Delta} V_{\bar{j}', \Delta} \frac{\sum_i \varphi_{\bar{j}}(\bar{c}_i) \varphi_{\bar{j}'}(\bar{c}_i)}{N_{\bar{c}}} \right) < \varepsilon \quad \text{Equation 99}$$

The orthonormality condition (Equation 35) yields the following expression:

$$\forall \varepsilon > 0, \exists \{n_{cut}, r_{cut}\} \text{ s.t. } \sum_{\bar{j} \in J_{cut}} E(V_{\bar{j}, \Delta}^2) < \varepsilon \quad \text{Equation 100}$$

Because the expected values of the  $V_{\bar{j}, \Delta}$  are zero, Equation 100 can be written in terms of the variances of the prior distributions on the ECI:

$$\forall \varepsilon > 0, \exists \{n_{cut}, r_{cut}\} \text{ s.t. } \sum_{\bar{j} \in J_{cut}} \sigma_{\bar{j}}^2 < \varepsilon \quad \text{Equation 101}$$

The term  $\sum_{\bar{j} \in J_{cut}} \sigma_{\bar{j}}^2$  will in general be very large unless  $\sigma_{\bar{j}}^2$  gets very small. In particular, using the constant prior implies that the prior expectation is that the cluster expansion does not converge. How small  $\sigma_{\bar{j}}^2$  must become as a function of cluster size can be approximated by considering the density of clusters of a given size. By expressing the variance of the prior distributions be a function of cluster size, the following continuous approximation can be used:

$$\sum_{\bar{j} \in J_{cut}} \sigma_{\bar{j}}^2 \approx \sum_{n=n_{cut}}^{n_{sites}} \int_{r_c}^{r_{max}} \rho(n, r) \sigma^2(n, r) dr \quad \text{Equation 102}$$

where  $\rho(n, r)$  is the approximate density of clusters with  $n$  sites and a maximum radius of  $r$ , and  $\sigma^2(n, r)$  is the variance of the prior distributions expressed as a function of  $n$  and  $r$ . Under this approximation, the convergence criterion becomes

$$\forall \varepsilon > 0, \exists \{n_{cut}, r_{cut}\} \text{ s.t. } \sum_{n=n_{cut}}^{n_{sites}} \int_{r_c}^{r_{max}} \rho(n, r) \sigma^2(n, r) dr < \varepsilon \quad \text{Equation 103}$$

A sufficient condition to satisfy Equation 103 is

$$\int_{r_c}^{r_{max}} \rho(n, r) \sigma^2(n, r) dr \leq a^{-n} \quad \text{Equation 104}$$

for  $a > 1$ .

A function  $\sigma^2(n, r)$  that satisfies Equation 104 can be found by calculating an upper bound for  $\rho(n, r)$  in the limit of large  $r$  and  $n$ . The maximum number of sites that can be selected such that the distance between any two sites is no greater than  $r$  grows as  $r^d$ , where  $d$  is the dimensionality of the structure. The number of clusters with  $n$  sites and a distance of no more than  $r$  between sites grows at a rate less than  $r^{nd}$ . Therefore,  $r^{nd}$  is an upper bound on  $\rho(n, r)$  in the limit of large  $r$  and  $n$ , yielding

$$\int_{r_c}^{r_{\max}} \rho(n, r) \sigma^2(n, r) dr < \int_{r_c}^{r_{\max}} r^{nd} \sigma^2(n, r) dr \quad \text{Equation 105}$$

Consider the following form for  $\sigma^2(n, r)$ :

$$\sigma^2(n, r) = \lambda r^{-nd - \lambda n - 1} \quad \text{Equation 106}$$

where  $\lambda > 0$ . Combined with Equation 105, this produces

$$\int_{r_c}^{r_{\max}} \rho(n, r) \sigma^2(n, r) dr < \lambda \int_{r_c}^{r_{\max}} r^{-n\lambda - 1} dr \leq \left( r_c^{-\lambda} \right)^{-n} \quad \text{Equation 107}$$

With the appropriate choice of  $\lambda$ , the approximate convergence criterion in Equation 104 can be satisfied.

Although this analysis is based on rough approximations, it suggests a useful form for  $\sigma^2(n, r)$ . The prior distribution for the ECI for a cluster with  $n$  sites and a maximum distance of  $r$  between sites has a width given by:

$$\sigma^2(n, r) = \lambda_1 (\lambda_2 r + \lambda_3 + 1)^{-\lambda_4 n - \lambda_5} \quad \text{Equation 108}$$

where the  $\lambda_i$  are non-negative parameters to be determined (for example, by cross-validation). For a cluster of only one site, set  $r = 0$ . This form allows for prior distributions that will converge, although it doesn't force such distributions. It also guarantees that as  $n$  and  $r$  get larger the widths of the prior distributions on the ECI get smaller (or stay the same). Because it depends on only five parameters, it reduces the risk of over fitting the data relative to the individual prior. Equation 108 will be referred to as the exponential width function. A similar approach was introduced by Laks et al. to treat problems of long-range lattice strain, although it was only applied to pair functions [89].

### III.2.2.3.5. The hierarchical width function

The exponential width function is just one of many width-generating functions that allow the prior expectation to be that the cluster expansion converges. The exponential width function assigns a width to the prior distribution of the ECI for a given cluster based on two parameters: the number of sites in a cluster and the maximum distance between those sites. Width-generating functions can use a richer set of data. For example, it might make sense to consider the distances between all sites in a cluster when determining the width of the prior distribution on the ECI. It also might make sense to consider what types of sites are included in a cluster. The following form takes these factors into consideration:

$$\sigma^2(\alpha) = \left( \sum_{s \in \alpha} \left( \lambda_s \prod_{s' \in \alpha, s' \neq s} (1 + \lambda_1 r_{s,s'}) \right)^{1/\lambda_2} \right)^{-2\lambda_2} \quad \text{Equation 109}$$

where  $s$  and  $s'$  label sites in cluster  $\alpha$  and  $r_{s,s'}$  is the distance between sites  $s$  and  $s'$ . The non-negative parameters  $\lambda_s$  are site-dependent, and all symmetrically equivalent sites share the same  $\lambda_s$ . These, and the non-negative parameters  $\lambda_1$  and  $\lambda_2$ , are to be determined.

The form given in Equation 109 ensures that as sites are added to clusters, and as the distance between the sites grows, the width of the prior distribution decreases. It also takes into account, through the  $\lambda_s$  parameters, the possibility that the types of sites contained in a cluster may affect the ECI. The distances between every pair of sites in the clusters are accounted for through the  $r_{s,s'}$  terms. For point terms, the expression reduces to  $\sigma_s^2 = \lambda_s^{-2}$ . Because there are so few parameters in this form, and the point terms are special in that they involve no interactions between sites, it might make sense to introduce a new parameter  $\lambda_3$  to explicitly deal with the point clusters. This is the approach used in this thesis. Because this form determines the width of the prior distribution in a way that considers the properties of the components of a cluster, and it ensures that all clusters have narrower priors than their subclusters, it will be referred to as the hierarchical width function.

### III.2.2.4. Parameterizing the width-generating functions

All of the width-generating functions discussed have parameters that need to be determined. This is known as model selection, and there are a variety of known ways to approach this problem. In this section two common approaches, leave-one-out cross-validation (LOO-CV) and generalized cross-validation (GCV), will be investigated.

#### III.2.2.4.1. Leave-one-out cross-validation

In leave-one-out cross validation, one sample is removed from the training set and a model is built from the remaining training samples. This model is used to predict the value for the sample that was left out. This is repeated for each of the samples in the training set. The leave-out-out cross-validation score is the root mean squared error for the predictions. This technique was introduced for cluster expansions by Van de Walle and Ceder in the context of cluster selection [74].

Leave-one-out cross-validation can also be used to evaluate prior distributions. The lower the leave-one-out cross-validation score is for a given prior, the better the models generated using the prior are at predicting the values of the left-out samples. A reasonable way to choose parameters might therefore be to find the set of parameters that gives the lowest leave-one-out cross-validation score.

Calculating the leave-one-out cross-validation can be expensive, as it requires fitting up to as many models as there are samples in the training set. In the case of the Gaussian prior, a more computationally efficient solution can be found. Start by transforming the input variables:

$$\begin{aligned} X_w &= W^{1/2} X \\ \bar{y}_T &= W^{1/2} (\bar{y} - X\bar{\beta}_p) \\ \tilde{y}_T &= W^{1/2} (\hat{V}_{Bayes}(X) - V_p(X)) \end{aligned} \quad \text{Equation 110}$$

Next substitute these variables into Equation 81:

$$\tilde{y}_T = X_w (X_w^T X_w + \Lambda)^{-1} X_w^T \bar{y}_T \quad \text{Equation 111}$$

Let  $X_{w,out}$  be a matrix in which each row is a sample being left out of the training set. Let  $X_{w,in}$  be the matrix consisting of the remaining samples. Similarly, let  $\bar{y}_{T,in}$  be the set

of transformed output values corresponding to the input values in  $X_{W,in}$ . Let  $\bar{y}_{T,out}$  and  $\tilde{y}_{T,out}$  be defined similarly. Let  $\tilde{y}_{T,out,CV}$  be the transformed output values for samples left out of the training set, predicted by a model trained only on the samples left in the training set. Equation 111 then becomes

$$\tilde{y}_{T,out,CV} = X_{W,out} \left( X_{W,in}^T X_{W,in} + \Lambda \right)^{-1} X_{W,in}^T \tilde{y}_{T,in} \quad \text{Equation 112}$$

Equation 112 can be re-written in terms of  $X_W$ ,  $X_{W,out}$ ,  $\bar{y}_T$ , and  $\bar{y}_{T,out}$ .

$$\tilde{y}_{T,out,CV} = X_{W,out} \left( X_W^T X_W + \Lambda - X_{W,out}^T X_{W,out} \right)^{-1} \left( X_W^T \bar{y}_T - X_{W,out}^T \bar{y}_{T,out} \right) \quad \text{Equation 113}$$

For readability, it is convenient to define a new matrix  $\Omega$ :

$$\Omega = X_W^T X_W + \Lambda \quad \text{Equation 114}$$

After applying the Sherman-Morrison-Woodbury formula [90] and doing a bit of algebra, the following expression is derived:

$$\tilde{y}_{T,out,CV} - \bar{y}_{T,out} = \left( I - X_{W,out}^T \Omega^{-1} X_{W,out} \right)^{-1} \left( \tilde{y}_{T,out} - \bar{y}_{T,out} \right) \quad \text{Equation 115}$$

The cross-validation score is then given by:<sup>5</sup>

$$CV = \sqrt{\text{Avg} \left\| \tilde{y}_{T,out,CV} - \bar{y}_{T,out} \right\|^2} \quad \text{Equation 116}$$

where the average is taken over different sets of samples left out of the fit. For leave-one-out cross-validation, usually the average is taken by leaving every sample out of the fit, one at a time. For leave-k-out cross-validation, the number of possible terms in the average grows roughly exponentially with k as long as k is significantly less than the size of the training set. For this reason, it may be preferable not to leave out every possible combination of samples for leave-k-out cross-validation.

Equation 115 is similar to a well-known method for calculating the leave-one-out cross-validation score for least-squares fits (see e.g. [74]). The advantage to this approach is that it only involves solving for  $\hat{y}_{T,out}$  once, using the full set of training data. The incremental cost

---

<sup>5</sup> Often the cross-validation and generalized cross-validation scores are given as the squares of the scores presented in this thesis. The form in this thesis was chosen to make units of the score equivalent to the units of the output data.

of using Equation 115 for each set of  $k$  samples left out is dominated by the cost of inverting the  $k \times k$  matrix  $(I - X_{w,out}^T \Omega^{-1} X_{w,out})$ . For leave-one-out cross-validation, this is very fast.

For the Laplace prior, the straightforward method of fitting  $n$  models for a training set of size  $n$  will be used to calculate the leave-one-out cross-validation score.

#### III.2.2.4.2. Generalized cross-validation

It can be shown that the leave-one-out cross-validation score is dependent on the basis used to represent the input data  $X$ . A widely used alternative that does not have this dependency is known as generalized cross-validation (GCV) [91]. It is equivalent to the leave-one-out cross-validation score for a system in which the input data have been rotated to a standard form. For systems in which the prior is Gaussian, the generalized cross-validation score is given by:

$$GCV = \frac{\sqrt{N \|\tilde{y}_T - \bar{y}_T\|^2}}{\text{Tr}(I - X_w^T \Omega^{-1} X_w)} \quad \text{Equation 117}$$

where  $N$  is the number of training samples and  $\Omega$  is given by Equation 114. In addition to being basis-independent, the GCV score has the advantage of being faster to compute than the leave-k-out cross-validation score.

For the Laplace prior, Tibshirani has suggested using a form similar to Equation 117, in which  $\Omega$  is given by

$$\Omega = X_w^T X_w + B^{-1} \Lambda \quad \text{Equation 118}$$

where  $B$  is a diagonal matrix with elements given by  $B_{ii} = |\beta_{\Delta,i}|$  [86]. Although an approximation, Equation 118 is much faster than calculating the leave-one-out cross-validation score.

For the Gaussian prior, an interesting interpretation of the generalized cross-validation score can be derived. This derivation uses some definitions were made previously, but are repeated here for clarity:

$$X_w = W^{1/2} X \quad \text{Equation 119}$$

$$\begin{aligned}\Omega &= X_w^T X_w + \Lambda \\ \bar{y}_T &= W^{1/2}(\bar{y} - X\bar{\beta}_p) \\ \hat{y}_T &= W^{1/2}(\hat{V}_{Bayes}(X) - V_p(X))\end{aligned}$$

Combining Equation 119 with Equation 81 and re-arranging yields:

$$\hat{y}_T - \bar{y}_T = (X_w \Omega^{-1} X_w^T - I) \bar{y}_T \quad \text{Equation 120}$$

If the training data have not been generated, the values for  $\bar{y}_T$  are unknown. Therefore  $\bar{y}_T$  can be treated as a random vector, denoted  $\tilde{y}_T$ . By making use of the prior distribution for  $\bar{y}_T$ , an expression for the prior distribution for  $(\hat{y}_T - \tilde{y}_T)$  can be derived:

$$\hat{y}_T - \tilde{y}_T = (X_w \Omega^{-1} X_w - I) W^{1/2} (X \tilde{\beta}_\Delta + \tilde{e}_n) \quad \text{Equation 121}$$

where  $\tilde{\beta}_\Delta$  is a vector of variables drawn from the prior distributions of coefficients,  $\tilde{e}_n$  is a vector of variables drawn from the prior distributions for the noise. The fact that the expectation of every element of  $\bar{y}_T$  is zero has been used to derive Equation 121. Taking the squared magnitude of each side produces:

$$\|\hat{y}_T - \tilde{y}_T\|^2 \sim Tr \left[ (X_w \Omega^{-1} X_w^T - I) W^{1/2} (X \tilde{\beta}_\Delta + \tilde{e}_n) (X \tilde{\beta}_\Delta + \tilde{e}_n)^T W^{1/2} (X_w \Omega^{-1} X_w^T - I) \right] \quad \text{Equation 122}$$

The expectation value of each side can be calculated, using Equation 77 and Equation 79 and the fact that  $\tilde{\beta}_\Delta$  and  $\tilde{e}_n$  are uncorrelated:

$$E \left( \|\hat{y}_T - \tilde{y}_T\|^2 \right) = \sigma_w^2 Tr \left[ (X_w \Omega^{-1} X_w^T - I) (X_w \Lambda^{-1} X_w^T + I) (X_w \Omega^{-1} X_w^T - I) \right] \quad \text{Equation 123}$$

Simplifying Equation 123 produces the following expression:

$$E \left( \|\hat{y}_T - \tilde{y}_T\|^2 \right) = \sigma_w^2 Tr \left[ (I - X_w \Omega^{-1} X_w^T) \right] \quad \text{Equation 124}$$

Therefore, for a given positive diagonal matrix  $\Lambda$ ,

$$E \left( \frac{\|\hat{y}_T - \tilde{y}_T\|^2}{Tr \left[ (I - X_w \Omega^{-1} X_w^T) \right]} \right) = \sigma_w^2 \quad \text{Equation 125}$$



Comparing this to the GCV score produces

$$E(GCV^2) = \frac{\sigma_w^2}{\frac{1}{N} \text{Tr} \left[ \left( I - X_w \Omega^{-1} X_w^T \right) \right]} \quad \text{Equation 126}$$

The numerator reflects the noise of the model; it is the error that cannot be recovered no matter how well the model parameters are fit. In the cluster expansion, this is the error introduced by the fact that not all possible clusters are included in the expansion. The denominator is a measure of the degrees of freedom of the model [92]. If no regularizer is used, this is simply the number of samples minus the number of variables. The regularizer serves to decrease the effective number of variables in the fit. This suggests that the GCV score is a way of quantifying the intuition that the fewer degrees of freedom required to build a model with little noise, the better the predictive power is expected to be.

---

### III.3. Sample generation

---

Although the generation of sample data for a cluster expansion can be expensive, the expected computational cost can be reduced by an intelligent choice of which samples are generated. In this sense, fitting a cluster expansion is a problem in active learning. To reduce the computational expense of building a cluster expansion, it is useful to consider how to best choose the structures that should be included in the training set.

Van de Walle and Ceder argued that when the correlation coefficients are determined using a least-squares fit, the predictive power of the fit over the universe of all structures should be approximately proportional to the trace of  $(X^T X)^{-1}$  [74]. Here this idea will be expanded upon to develop a method for structure selection that works for a Bayesian cluster expansion and requires fewer approximations.

The derivation begins by calculating the covariance matrix for the correlation coefficients in the case of regularized regression based on Gaussian prior distributions. The correlation coefficients are given by:

$$\tilde{\beta}_\Delta = \Omega^{-1} X_w^T \bar{y}_T \quad \text{Equation 127}$$

where  $\tilde{\beta}_\Delta$  is a vector of the estimated correlation coefficients, and  $\Omega$ ,  $X_w$ , and  $\bar{y}_T$  are defined in Equation 119. The output variables can be written in terms of the “true” coefficients  $\tilde{\tilde{\beta}}_\Delta$ , and a vector of random noise,  $\tilde{e}_n$ . The tilde over  $\tilde{\beta}_\Delta$  is used to represent the fact that the coefficients are drawn from a prior distribution. Equation 125 then becomes

$$\tilde{\beta}_\Delta = \Omega^{-1} X_w^T W^{1/2} (X \tilde{\tilde{\beta}}_\Delta + \tilde{e}_n) \quad \text{Equation 128}$$

The difference between the predicted and “true” coefficients is

$$\tilde{\beta}_\Delta - \tilde{\tilde{\beta}}_\Delta = (\Omega^{-1} X_w^T X_w - I) \tilde{\tilde{\beta}}_\Delta + \Omega^{-1} X_w^T W^{1/2} \tilde{e}_n \quad \text{Equation 129}$$

This is the sum of two random, uncorrelated variables. The covariance matrix is given by

$$\text{Cov}\left(\tilde{\beta}_\Delta - \tilde{\beta}_\Delta\right) = \sigma_w^2 \left[ \left( \Omega^{-1} X_w^T X_w - I \right) \Lambda^{-1} \left( X_w^T X_w \Omega^{-1} - I \right) + \Omega^{-1} X_w^T X_w \Omega^{-1} \right] \quad \text{Equation 130}$$

where the facts that  $\text{Cov}\left(\tilde{\beta}_\Delta\right) = \sigma_w^2 \Lambda^{-1}$  and  $\text{Cov}\left(\tilde{e}_n\right) = \sigma_w^2 W^{-1}$  have been used. This expression can be simplified to

$$\text{Cov}\left(\tilde{\beta}_\Delta - \tilde{\beta}_\Delta\right) = \sigma_w^2 \Omega^{-1} \quad \text{Equation 131}$$

For a given input vector  $\bar{x}$ , the difference between the predicted output and the actual output is given by

$$\hat{y}_{T,i} - y_{T,i} = w_i^{1/2} \bar{x}_i^T \tilde{\beta}_\Delta - \left( w_i^{1/2} \bar{x}_i^T \tilde{\beta}_\Delta + w_i^{1/2} \tilde{e}_i \right) \quad \text{Equation 132}$$

Rearranging produces

$$\hat{y}_{T,i} - y_{T,i} = w_i^{1/2} \bar{x}_i^T \left( \tilde{\beta}_\Delta - \tilde{\beta}_\Delta \right) - w_i^{1/2} \tilde{e}_i \quad \text{Equation 133}$$

The first term on the right is the variance of the fit, and the second is the noise. Once again, the right side is the sum of two uncorrelated random variables. Taking the covariance of each side produces

$$E\left(\hat{y}_{T,i} - y_{T,i}\right)^2 = \sigma_w^2 \left( w_i^{1/2} \bar{x}_i^T \Omega^{-1} \bar{x}_i w_i^{1/2} + 1 \right) \quad \text{Equation 134}$$

Recognizing that the term on the right is the trace of a 1x1 matrix, and using the cyclic property of the trace, the following expression can be derived:

$$E\left(\hat{y}_{T,i} - y_{T,i}\right)^2 = \sigma_w^2 \text{Tr}\left(\Omega^{-1} \bar{x}_i w_i \bar{x}_i^T\right) + \sigma_w^2 \quad \text{Equation 135}$$

It is desirable to minimize the expected squared error over the entire domain of structures, which will be referred to as the “test set” or “test domain”. Equation 135 can be generalized for any set of test data.

$$E\left(\left\| \tilde{y}_{T,test} - \bar{y}_{T,test} \right\|^2\right) = \sigma_w^2 \left[ \Omega^{-1} \cdot X_{W,test}^T X_{W,test} + N_{test} \right] \quad \text{Equation 136}$$

where  $N_{test}$  is the number of structures in the test set, and  $\Omega^{-1} \cdot X_{W,test}^T X_{W,test}$  is the Frobenius inner product (the sum over element-wise multiplication) of the matrices  $\Omega^{-1}$  and  $X_{W,test}^T X_{W,test}$ . This is a product between  $p \times p$  matrices, where  $p$  is the number of distinct correlation coefficients to be determined.

The expected squared prediction error per structure can be written as

$$E\left(\frac{\|\hat{\bar{y}}_{T,test} - \bar{y}_{T,test}\|^2}{N_{test}}\right) = \sigma_w^2 [\Omega^{-1} \cdot D + 1] \quad \text{Equation 137}$$

where the matrix  $D$  is defined by

$$D = \frac{X_{W,test}^T X_{W,test}}{N_{Test}} \quad \text{Equation 138}$$

This provides a useful method for trying to select structures that minimize the prediction error over the entire function domain. The term  $\sigma_w^2$  is a function of the model selection, which has been discussed in the previous section. It is independent of the structures that are included in the training set. Therefore the expected squared prediction error, as a function of structures included in the training set, is proportional to  $\Omega^{-1} \cdot D + 1$ . By minimizing this term, the expected error can be minimized. Specifically training structures that minimize  $\Omega^{-1} \cdot D$ , which is proportional to the error due to variance, need to be selected. Note that if  $\Omega = X_w^T X_w$  and  $D = I$ , Van de Walle and Ceder's approximation is recovered.

There are now two problems to address: The first is to calculate the matrix  $D$  so that it represents the universe of all test structures. Because this matrix is a compact,  $p \times p$  representation of the relevant information about the function domain, it will be referred to as the domain matrix.

The second problem is finding a computationally efficient way to select training structures so that  $\Omega^{-1} \cdot D$  is minimized. In the next two sections these problems will be addressed.

### III.3.1. Estimating the domain matrix

---

There are several methods available to estimate the domain matrix. In this section a few simple methods, and the situations in which each method might be appropriate, will be evaluated.

### III.3.1.1. Direct enumeration

Perhaps the simplest method of estimating the domain matrix is to enumerate a large set of unique structures in the test domain, and directly calculate the matrix. For cluster expansions with a large number of sites, such as those for infinite crystals, the challenge can be enumerating a set of unique structures that is large enough to be representative for a domain. In this section an efficient algorithm for accomplishing this in the case of cluster expansions on infinite crystals will be developed.

An infinite crystal is created by starting with a set of sites, known as the basis, and repeating the basis over an infinite Bravais lattice. The set of all sites in a crystal that is periodic in  $d$  dimensions can be defined by

$$\forall \vec{b}_j, \forall \vec{a} \in \mathbb{Z}^d : \vec{r}_{j,\vec{a}} = \vec{b}_j + \vec{a}^T L_{site} \quad \text{Equation 139}$$

where  $\vec{b}_j$  are the coordinates of the basis sites,  $\mathbb{Z}^d$  is the set of all  $d$ -dimensional integer vectors,  $\vec{r}_{j,\vec{a}}$  is the coordinate a crystal site, and  $L_{site}$  is a matrix whose rows are  $d$  linearly independent primitive vectors of the Bravais lattice. For a crystalline cluster expansion, every site has the same occupation variable domain and single-variable basis functions as the basis site from which it was generated. In this section it will be assumed that the smallest possible basis has been chosen to represent the crystal of sites. The corresponding Bravais lattice will be called the “site lattice”.

Often, low-energy structures (site decorations) are also crystalline. The Bravais lattice for a crystalline structure may be different from the “site lattice”, depending on the occupancies of the sites. The lattice for the occupied sites will be called the “structure lattice.” The site lattice has been defined so that it represents all lattice transformations that map sites onto symmetrically equivalent sites. Because of this, the lattice vectors for the structure lattice must be linear combinations of integer multiples of the lattice vectors for the site lattice. Mathematically, this can be written:

$$L_{structure} = S L_{site} \quad \text{Equation 140}$$

where the rows of  $L_{structure}$  are linearly independent primitive vectors for the structure lattice, and  $S$  is a  $d \times d$  matrix with all integer elements. The absolute value of the determinant of

$S$  is the ratio between the number of sites in the basis of the structure lattice and the number of sites in the basis of the site lattice. This means that the number of sites in the basis for the structure must be an integer multiple of the number of sites in the cluster expansion basis. This integer multiple will be referred to as the “size” of the structure lattice.

It is possible for two lattices to be equivalent. This occurs when the primitive vectors of one lattice can be expressed as a linear combination of integer multiples of the primitive vectors of the other, and both lattices have bases of the same size. This type of lattice equivalence will be represented by the symbol  $\sim$ , and defined mathematically as

$$L_1 \sim L_2 \text{ iff } \exists T \text{ s.t. } L_1 = TL_2 \quad \text{Equation 141}$$

where  $T$  is a unimodular matrix (a square matrix with integer elements and determinant of 1 or -1). Because  $T$  is unimodular, it is straightforward to derive the transitive property of lattice equivalence:

$$L_1 \sim L_2 \ \& \ L_2 \sim L_3 \Rightarrow L_1 \sim L_3 \quad \text{Equation 142}$$

Because of lattice equivalence, there are a finite number of possible different structure lattices with a given basis size.

For a cluster expansion defined on an infinite lattice, there may be symmetry operations in addition to translations. The cluster expansion space group will be defined as the group of all operations that leave the sites, allowed occupancies, and single-variable basis functions unchanged. Every one of the operations of the space group can be expressed as a “point” operation, which leaves the lattice origin unchanged, combined with a translation. Every space group has a finite number of point operations. Because the point operations must map lattice points to lattice points, they can be expressed as integer matrices operating on lattice vectors. These point operations may cause two lattices to be symmetrically equivalent. For example, if a cluster expansion is defined on a 2-dimensional square lattice, then the reflection operation that maps the x-coordinate to the y-coordinate and vice-versa is a point operation. Because of this operation, the structure lattice with primitive vectors of (2, 0) and (0, 1) is equivalent to the structure lattice with primitive vectors of (1, 0) and (0, 2).

A template for an algorithm to enumerate structures can now be written. Often the structures of interest have relatively small bases, so the algorithm starts by enumerating

structure lattices with a size of 1. This is repeated for lattices with a size of 2, then with size of 3, etc. For each size, all structure lattices at that size are enumerated with the constraint that no two enumerated lattices are lattice equivalent. Lattices that are symmetrically equivalent based on the point group of the cluster expansion are also removed. For each lattice, all possible decorations of the sites in the corresponding basis are enumerated, with the constraint that no two decorations representing symmetrically equivalent structures are enumerated. The algorithm is terminated at some pre-defined point, usually a size at which the computational expense outweighs the benefit of additional structures. In the following sections, the details of a rapid implementation of an algorithm that follows this template will be fleshed out.

### III.3.1.1.1. Enumerating structure lattices

Every structure lattice can be generated from a given site lattice using Equation 140, with a different matrix  $S$  for each structure lattice. Let two lattices  $L_1$  and  $L_2$  be defined by

$$\begin{aligned} L_1 &= S_1 L_{site} \\ L_2 &= S_2 L_{site} \end{aligned} \quad \text{Equation 143}$$

By Equation 140, and the fact that  $L_{site}$  is invertible, it follows that

$$L_1 \sim L_2 \text{ iff } S_1 \sim S_2 \quad \text{Equation 144}$$

The problem of finding a maximal set of lattices that are not lattice-equivalent can therefore be transformed to the problem of finding the maximal set of integer matrices that are not lattice-equivalent.

To proceed further, the concept of Hermite normal is introduced [93]. An integer matrix in Hermite normal form is lower-triangular<sup>6</sup>, with the requirements that the diagonal elements are all positive, no elements are negative, and the maximum element in each column is located on the diagonal. If two matrices in Hermite normal form are different from each other, then they are not lattice-equivalent. On the other hand, every integer matrix is lattice-equivalent to exactly one matrix in Hermite normal form. Because a matrix in Hermite normal form is diagonal, the size of a lattice represented by a matrix in Hermite normal form is given by the product of the diagonal elements. If  $H(S)$  represents an

---

<sup>6</sup> Equivalently, often upper triangular matrices are used.

operation that reduces an integer matrix to Hermite normal form, lattice-equivalence can be written as follows:

$$L_1 \sim L_2 \text{ iff } H(S_1) = H(S_2) \quad \text{Equation 145}$$

Enumerating all non-equivalent structure lattices of a given size is therefore the same as enumerating all matrices in Hermite normal form with a given determinant. The algorithm for enumerating all unique  $d$ -dimensional superlattices is as follows:

- 1) Let  $k$  be the desired size of the structure lattice.
- 2) Calculate all possible factorizations of  $k$  consisting of  $d$  factors.
- 3) For each factorization:
  - a. Calculate all unique permutations of the factorization.
  - b. For each permutation:
    - i. Place the factors, in the order of the permutation, along the diagonal of a  $d \times d$  matrix.
    - ii. Enumerate all possible ways of populating the lower triangular part of the matrix, such that the elements of the lower triangle are non-negative integers, and the maximum of each column lies along the diagonal.
    - iii. Use each enumerated matrix as the matrix  $S$  in Equation 141 to generate the superlattice vectors.

For example, the set of all symmetrically distinct superlattices with primitive cells that are three times the size of the base primitive cell can be generated from the following integer matrices:



$$S \in \left\{ \begin{array}{l} \left[ \begin{array}{ccc} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right], \left[ \begin{array}{ccc} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{array} \right], \left[ \begin{array}{ccc} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{array} \right], \\ \left[ \begin{array}{ccc} 3 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right], \left[ \begin{array}{ccc} 3 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{array} \right], \left[ \begin{array}{ccc} 3 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 0 & 1 \end{array} \right], \\ \left[ \begin{array}{ccc} 3 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right], \left[ \begin{array}{ccc} 3 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 0 & 1 \end{array} \right], \left[ \begin{array}{ccc} 3 & 0 & 0 \\ 2 & 1 & 0 \\ 2 & 0 & 1 \end{array} \right], \\ \left[ \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{array} \right], \left[ \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 1 & 1 \end{array} \right], \left[ \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 2 & 1 \end{array} \right], \\ \left[ \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{array} \right] \end{array} \right\} \quad \text{Equation 146}$$

The matrices in Equation 146 have been arranged in a logical order. The three different permutations of  $\{3,1,1\}$  are explored in order (for example, from greatest to least when written out as numbers in base  $k+1$ ). For each permutation the off-diagonal elements are explored in a similar order, with the constraint that no off-diagonal element can be larger than the diagonal element in the same column. With this approach, it is straightforward to rapidly calculate the index of each matrix in the set, and to calculate the matrix that belongs at each index. Let  $S_i$  be the matrix in Hermite normal form corresponding to index  $i$ , and let  $i(S)$  be the function that maps a given integer matrix to the appropriate index.

The possibility remains that the generated structure lattices are symmetrically equivalent due to the point operations of the cluster expansion. For example, if the cluster expansion has cubic symmetry, the structure lattices represented by the first, tenth, and last matrices in Equation 146 would be symmetrically equivalent. In general, the criterion for two structure lattices, given by  $L_1 = S_1 L_{site}$  and  $L_2 = S_2 L_{site}$ , to be symmetrically equivalent (denoted by  $\cong$ ) is

$$L_1 \cong L_2 \text{ iff } \exists F \in G_{pt} \text{ s.t. } S_1 F L_{site} \sim S_2 L_{site} \quad \text{Equation 147}$$

where  $G_{pt}$  is the set of integer matrices representing point operations for the cluster expansion. Remembering that  $H(S)$  is the operation that reduces an integer matrix to Hermite normal form, Equation 147 can be re-written:

$$L_1 \cong L_2 \text{ iff } \exists F \in G_{pt} \text{ s.t. } H(S_1 F) = H(S_2) \quad \text{Equation 148}$$

The full algorithm for generating symmetrically distinct structure lattices is as follows:

- 1) For a given size  $k$ , first calculate how many matrices exist in Hermite normal form with a determinant of  $k$ . Call this number  $N_k$ .
- 2) For the indices  $i = 1 \dots N_k$ :
  - a. Generate the corresponding matrix  $S_i$  in Hermite normal form.
  - b. For each cluster expansion point operation  $F_j$ :
    - i. Calculate  $j = i(H(S_i F_j))$ .
    - ii. If  $j < i$ , then a symmetrically equivalent matrix has already been generated; Set  $i = i + 1$ , and go to a).
  - c. Use Equation 140 is used to generate the new structure lattice.

This algorithm runs in  $O(N_k)$  time and uses only integer operations and very little memory. As an example of the performance of this algorithm, on a 2.0 GHz Pentium M processor all 9537 symmetrically distinct structure lattices up to size 64 for an FCC cluster expansion were found in 49 seconds.

### III.3.1.1.2. Enumerating structure basis decorations

For every structure lattice generated, it is necessary to generate all of the ways in which the basis sites for the structure can be decorated. Calculating the possible decorations is trivial: simply list all of the basis sites in some arbitrary order, and then loop in an ordered fashion over all possible combinations of occupation variables. Much like the matrix enumeration in the previous section, this provides a way to map every possible decoration to

a unique index, and for a given index to quickly calculate the corresponding decoration. Let the vector  $\vec{d}_i$  represent the decoration corresponding to index  $i$ , and  $i_d(\vec{d})$  be the function that maps a given decoration  $\vec{d}$  to a unique index.

It is still necessary to eliminate symmetrically equivalent decorations. Two decorations are symmetrically equivalent if one of the operations in the space group of the cluster expansion can be applied to a structure built from one decoration to arrive at the other decoration. Applying a symmetry operation to a decoration maps each site in the basis to another site in the basis. In this way, each symmetry operation is equivalent to a permutation of the basis sites. There are a finite number of permutations corresponding to symmetry operations in the cluster expansion space group. The set of all such permutations will be called  $G_{perm}$ . A statement can then be written that is analogous to Equation 148:

$$\vec{d}_1 \cong \vec{d}_2 \text{ iff } \exists F \in G_{perm} \text{ s.t. } i_d(\vec{d}_1) = i_d(F(\vec{d}_2)) \quad \text{Equation 149}$$

The enumeration algorithm proceeds similarly to the algorithm for the enumeration of structure lattices:

- 1) Calculate the total number of allowed decorations,  $N_d$ .
- 2) For  $i = 1 \dots N_d$ :
  - a. Generate the decoration  $\vec{d}_i$  corresponding to that index.
  - b. For all  $F \in G_{perm}$ 
    - i. Calculate  $j = i_d(F(\vec{d}_i))$ .
    - ii. If  $j < i$ , then a symmetrically equivalent matrix has already been generated; Set  $i = i + 1$ , and go to a).
  - c. Add the current decoration to the list of structures generated.

There are two more details to work out in the above algorithm. The first is that a given decoration might create a structure for which there is a smaller basis. For example, for a structure lattice of size  $k$ , if all sites in the basis are decorated with the same occupation variable, then the resulting structure can be expressed as one with a structure lattice of size 1. Such a structure would have already been generated when structures of size 1 were generated. To avoid generating duplicate structures, it is necessary to make sure that there is

no translation, other than translations in the structure lattice, which maps a structure with a given decoration onto itself. This can be accomplished by listing the full set of site lattice translations so that no two site lattice translations differ by a structure lattice translation. These translations are just special cases of the symmetry permutations listed earlier. Each of these site lattice translations are applied to each candidate decoration, and if any of them map the decoration onto itself then the decoration is ignored and the next decoration is considered.

The other remaining detail is the performance of the algorithm. The algorithm should run in  $O(N_d)$  time, where for a binary cluster expansion and a structure with  $m$  basis sites,  $N_d = 2^m$ . The exponential dependence on  $m$  this algorithm cannot be removed, but it may be possible to reduce the exponent. The iteration over decorations will often involve changing the occupation variable of only one site between decoration  $i$  and decoration  $i+1$ . Because of this, if decoration  $i$  is symmetrically equivalent to a previous decoration, there is a good chance that decoration  $i+1$  will be equivalent to another previous decoration. Evaluating all decorations until one is found that is symmetrically different from previous decorations may be time consuming. A more efficient approach would be to directly examine the permutations, and figure out the index of the next allowed decoration. It is then possible to skip directly to this index without generating the intermediate decorations. This version of the algorithm can run in as little as  $O(N_{d,u})$  time, where  $N_{d,u}$  is the number of symmetrically unique decorations. The advantage for structures with a large basis and high symmetry can be substantial.

For example, for a binary FCC cluster expansion, there are approximately  $2.8 \cdot 10^{12}$  possible decorations of structure lattices with a size of 31, but only about  $1.9 \cdot 10^9$  are symmetrically unique. This gives  $N_{d,u} \approx 1.58^m$ . Using the algorithms described in this thesis, all structures with a basis of up to 31 sites were found for this system in less than eight hours on a 2 GHz Pentium M processor.

### III.3.1.2. Analytical methods

An alternative to direct enumeration is to evaluate the domain matrix analytically. For this section, assume all structures as equally weighted. Each element of the domain matrix can be calculated as follows:

$$D_{\alpha\beta} = \frac{\sum_{i=1}^{N_{test}} X_{i\alpha} X_{i\beta}}{N_{test}} \quad \text{Equation 150}$$

Substituting Equation 32 into Equation 150 produces the following expression:

$$D_{\alpha\beta} = \frac{\sum_{i=1}^{N_{test}} \langle \varphi(\bar{c}_i) \rangle_{\alpha} \langle \varphi(\bar{c}_i) \rangle_{\beta}}{N_{test}} \quad \text{Equation 151}$$

Substituting Equation 27 yields

$$D_{\alpha\beta} = \frac{\sum_{i=1}^{N_{test}} \sum_{\bar{j} \in \alpha} \sum_{\bar{j}' \in \beta} \varphi_{\bar{j}}(\bar{c}_i) \varphi_{\bar{j}'}(\bar{c}_i)}{N_{\alpha} N_{\beta} N_{test}} \quad \text{Equation 152}$$

In the following sections, Equation 152 will be considered in two cases: when the test set is the domain of all possible structures, and when the test set consists only of structures at a given composition.

#### III.3.1.2.1. All possible structures

If the test set is the domain of all allowed structures,  $N_{test} = N_{\bar{c}}$ . If an orthonormal basis is assumed, Equation 152 can be combined with Equation 35 to yield the following expression:

$$D_{\alpha\beta} = \frac{\sum_{\bar{j} \in \alpha} \sum_{\bar{j}' \in \beta} \delta_{\bar{j}, \bar{j}'}}{N_{\alpha} N_{\beta}} \quad \text{Equation 153}$$

Evaluating the sums yields

$$D_{\alpha\beta} = \frac{\delta_{\alpha\beta}}{N_{\alpha}} \quad \text{Equation 154}$$

Equation 154 is an interesting result: the domain matrix is diagonal, and the only non-zero elements for the domain matrix are for the cluster orbits that contain a finite number of

clusters. For an infinite crystal, there is only one cluster orbit that fits this definition and has a finite number of sites for which  $j_k \neq 0$ . This is the orbit for the empty cluster, to which an index of zero is assigned. Thus the domain matrix has only one non-zero element,  $D_{00} = 1$ . Plugging this into Equation 137 yields

$$E \left( \frac{\left\| \bar{\hat{y}}_{T, test} - \bar{y}_{T, test} \right\|^2}{N_{test}} \right) = \sigma_w^2 \left[ (\Omega^{-1})_{00} + 1 \right] \quad \text{Equation 155}$$

This might seem to be a strange result, because it removes much of the information about the non-empty clusters. This is due to the fact that the distribution of all structures is sharply peaked at a point at which all correlations are zero except for the correlation for the empty cluster. Thus the prediction error, on average, will be dominated by errors in the ECI for the empty cluster. Equation 155 reflects the importance of getting the ECI for the empty cluster correct.

### III.3.1.2.2. Fixed composition

Although the entire universe of structures is sharply peaked around structures with zero correlations, the entire universe might not be of interest. For example, sometimes the structures being studied have a fixed composition of elements, such as 30% silver and 70% gold. In the limit of an infinite crystal, this is equivalent to assigning probabilities to each of the possible values of the occupation variables. Equation 152 can be re-written as

$$D_{\alpha\beta} = \frac{\sum_{\bar{j} \in \alpha} \sum_{\bar{j}' \in \beta} \langle \varphi_{\bar{j}}(\bar{c}_i) \varphi_{\bar{j}'}(\bar{c}_i) \rangle_{Test}}{N_\alpha N_\beta} \quad \text{Equation 156}$$

where  $\langle x_i \rangle_{Test}$  is used to indicate the average value of quantity  $x$  over the entire test set.

Combining Equation 156 with Equation 24 yields the following expression:

$$D_{\alpha\beta} = \frac{\sum_{\bar{j} \in \alpha} \sum_{\bar{j}' \in \beta} \left\langle \sum_{i=1}^{N_{test}} \prod_k \phi_{j_k, k}(c_{k,i}) \phi_{j'_k, k}(c_{k,i}) \right\rangle_{Test}}{N_\alpha N_\beta} \quad \text{Equation 157}$$

where, as a reminder,  $c_{k,i}$  is the occupation variable for the  $k$ th site and the  $i$ th structure,  $j_k$  is the  $k$ th element of  $\bar{j}$ , and  $\phi_{j_k, k}$  is the  $j_k$ th single-variable basis function for site  $k$ .

If  $p_{k,m}$  is defined as the probability that the occupation variable of the  $k$ th site is the  $m$ th allowed value, then in the limit of an infinite number of sites Equation 157 can be written:

$$D_{\alpha\beta} = \frac{\sum_{\bar{j} \in \alpha} \sum_{\bar{j}' \in \beta} \left( \prod_k \sum_{m=1}^{N_k} p_{k,m} \phi_{j_k,k}(c_{k,m}) \phi_{j'_k,k}(c_{k,m}) \right)}{N_\alpha N_\beta} \quad \text{Equation 158}$$

where  $c_{k,m}$  is now the  $m$ th allowed value for the occupation variable for the  $k$ th site.

If there are an infinite number of sites, but the prior guess only allows clusters with a finite number of sites to have non-zero ECI, then the odds of two clusters overlapping are essentially zero. Taking advantage of this fact, Equation 158 becomes:

$$D_{\alpha\beta} = \frac{\sum_{\bar{j} \in \alpha} \sum_{\bar{j}' \in \beta} \left( \prod_{k:j_k \neq 0} \sum_{m=1}^{N_k} p_{k,m} \phi_{j_k,k}(c_{k,m}) \right) \left( \prod_{k:j'_k \neq 0} \sum_{m=1}^{N_k} p_{k,m} \phi_{j'_k,k}(c_{k,m}) \right)}{N_\alpha N_\beta} \quad \text{Equation 159}$$

Because the terms in parenthesis are going to be the same for all clusters in the same orbit, Equation 159 can be written:

$$D_{\alpha\beta} = \left( \prod_{k:j_k \neq 0, \bar{j} \in \alpha} \sum_{m=1}^{N_k} p_{k,m} \phi_{j_k,k}(c_{k,m}) \right) \left( \prod_{k:j'_k \neq 0, \bar{j}' \in \beta} \sum_{m=1}^{N_k} p_{k,m} \phi_{j'_k,k}(c_{k,m}) \right) \quad \text{Equation 160}$$

Another way of writing Equation 160 is

$$D = \langle \bar{x} \rangle \langle \bar{x} \rangle^T \quad \text{Equation 161}$$

where

$$\langle \bar{x} \rangle_\alpha = \langle x_\alpha \rangle = \prod_{k:j_k \neq 0, \bar{j} \in \alpha} \sum_{m=1}^{N_k} p_{k,m} \phi_{j_k,k}(c_{k,m}) \quad \text{Equation 162}$$

This formula can be used to calculate the domain matrix at any composition. Consider the special case of a binary alloy, with the following domain and basis:

$$\begin{aligned} \forall k \quad c_k &\in \{-1, 1\} \\ \forall k \quad \phi_{0,k}(c_k) &= 1 \\ \forall k \quad \phi_{1,k}(c_k) &= c_k \end{aligned} \quad \text{Equation 163}$$

where occupation with species A corresponds to  $c_k = -1$  and occupation with species B corresponds to  $c_k = 1$ . Let the percentage of sites occupied by species A be fixed at  $p$ . For an infinite crystal, from Equation 160 becomes:

$$D_{\alpha\beta} = [1 - 2p]^{(n_\alpha + n_\beta)}, D_{00} = 1 \quad \text{Equation 164}$$

where  $n_\alpha$  is the number of sites in each cluster function  $\vec{j} \in \alpha$  for which  $j_k \neq 0$ , and  $n_\beta$  is defined similarly. From Equation 164, it can be seen that fixing the concentration at 50% leads to Equation 155. This is because the result for systems not constrained to a given composition is due to the sharply peaked density of structures at 50% composition.

In general, the distribution of compositions for structures in the test set is not sharply peaked at 50%. For example, suppose there is an equal interest in predicting the properties of structures at all compositions. The domain matrix for this situation can be calculated by integrating Equation 164 over all compositions:

$$D_{\alpha\beta} = \int_{p=0}^1 [1 - 2p]^{(n_\alpha + n_\beta)} dp = \begin{cases} \frac{1}{(n_\alpha + n_\beta + 1)} : n_\alpha + n_\beta \text{ even} \\ 0 : n_\alpha + n_\beta \text{ odd} \end{cases} \quad \text{Equation 165}$$

Noticing that along the diagonal of the matrix,  $n_\alpha + n_\beta$  is always even, some similarity to Van de Walle and Ceder's trace approximation becomes apparent.

The formulas derived in this section are exact in the limit of infinite crystal periodicity, and should work best for structures that do not necessarily have small crystal bases. In this sense, the analytical method may be considered complementary to the direct enumeration method.

### III.3.1.3. Statistical methods

There are some situations in which the distribution of structures on the test domain is a function of the occupation variables themselves. In general, there may be some probability distribution  $P(\vec{x}_i)$  representing the likelihood of a given structure appearing in the test set with

$$\sum_i P(\vec{x}_i) = 1 \quad \text{Equation 166}$$



If  $P(\bar{x}_i)$  (or a close approximation) is known, Equation 150 can be approximated by statistically sampling the test domain. For example, a Metropolis algorithm [73] might be used.

An example of when this approach could be useful would be if the thermodynamic properties of a material in a given range of temperature and chemical potential are of interest. For example, for a binary compound at a given temperature,  $T$ , with the chemical potential difference between species given by  $\mu$ , statistical mechanics states that in the grand canonical ensemble

$$P_{T,\mu_A}(\bar{x}_i) = \frac{e^{\frac{V(\bar{x}_i) - \mu N_A}{kT}}}{\sum_i e^{\frac{V(\bar{x}_i) - \mu N_A}{kT}}} \quad \text{Equation 167}$$

where  $N_A$  is the total number of nuclei of species A, and the total number of nuclei in the system is assumed to be fixed.

The probability function in Equation 167 can be used along with statistical sampling to estimate the domain matrix at a given temperature and chemical potential,  $D(T, \mu_A)$ . The key is that it is necessary to have a good estimate for  $V(\bar{x}_i)$ , which is the original problem. A reasonable estimate for this function might be the prior estimate. Equivalently, if some training data have already been generated, the most likely function in the posterior distribution could be used. Alternatively, and perhaps more accurately, the parameters of the function can be treated as random variables themselves. The sampling would then be over not only occupation variables, but over values for the ECI distributed according to the posterior distribution.

The domain matrix over a range of chemical potentials and temperatures can be estimated by numerically integrating over the domain matrices at fixed potentials and temperatures:

$$D_{[T_{\min}, T_{\max}], [\mu_{\alpha, \min}, \mu_{\alpha, \max}]} = \int_{T_{\min}}^{T_{\max}} \int_{\mu_{\alpha, \min}}^{\mu_{\alpha, \max}} D(T, \mu_A) d\mu_{\alpha} dT \quad \text{Equation 168}$$

This method for calculating the domain matrix, combined with Equation 137, provides a quantitative method for measuring the key property in a training set: the degree to which

structures are both thermodynamically likely and informative about the potential energy function.

### III.3.2. Selecting the training set

Once a domain matrix has been calculated, a training set can be generated to that gives a low value for the expected squared error, represented by Equation 137. Generally, the problem of building a training set can be viewed as one of choosing which structures to add to an existing set of training data. Equation 137 then becomes

$$E \left( \frac{\|\tilde{\hat{y}}_{T,test} - \bar{y}_{T,test}\|^2}{N_{test}} \right) = \sigma_w^2 \left[ \left( \Omega_{old} + X_{new}^T X_{new} \right)^{-1} \cdot D + 1 \right] \quad \text{Equation 169}$$

where

$$\Omega_{old} = X_{old}^T X_{old} + \Lambda \quad \text{Equation 170}$$

The matrix  $X_{old}$  is the set of input data that has already been added to the training set, and the matrix  $X_{new}$  is the set of input data that is being considered for addition to the training set. If  $\Omega_{old}$  is assumed to be invertible, which it generally should be as long as a good regularizing matrix  $\Lambda$  has been chosen, the Sherman-Morrison-Woodbury formula can be used to re-write Equation 182:

$$E \left( \frac{\|\tilde{\hat{y}}_{T,test} - \bar{y}_{T,test}\|^2}{N_{test}} \right) = \sigma_w^2 \left[ \left( \Omega_{old}^{-1} - \Omega_{old}^{-1} X_{new}^T \left( I + X_{new} \Omega_{old}^{-1} X_{new}^T \right)^{-1} X_{new} \Omega_{old}^{-1} \right) \cdot D + 1 \right] \quad \text{Equation 171}$$

Using the properties of the Frobenius inner product, Equation 171 can be re-written:

$$E \left( \frac{\|\tilde{\hat{y}}_{T,test} - \bar{y}_{T,test}\|^2}{N_{test}} \right) = \sigma_w^2 \left[ \Omega_{old}^{-1} \cdot D + 1 \right] - \sigma_w^2 \text{Tr} \left[ X_{new} \Omega_{old}^{-1} D \Omega_{old}^{-1} X_{new}^T \left( I + X_{new} \Omega_{old}^{-1} X_{new}^T \right)^{-1} \right] \quad \text{Equation 172}$$

The first term on the right is the expected squared prediction error of the known training set. The improvement due to the additional training samples is given by the second

term on the right. The problem of choosing new structures can therefore be reduced to the problem of finding a set of structures that minimizes the following expression:

$$-Tr \left[ X_{new} \Omega_{old}^{-1} D \Omega_{old}^{-1} X_{new}^T \left( I + X_{new} \Omega_{old}^{-1} X_{new}^T \right)^{-1} \right] \quad \text{Equation 173}$$

The expression in Equation 186 term can be calculated in  $O(p^2 + N_{new}^3)$  time, where  $p$  is the number of coefficients to be fit and  $N_{new}$  is the number of new training samples to be added to the training set. Finding sets of structures to minimize Equation 186 is a can be addressed with any number of well-known minimization heuristics.

The fact that different structures have different computational costs has not yet been addressed. Generally, it is desirable to find the best set of structures within a given computational cost. Van de Walle and Ceder suggest addressing this problem by finding the structure that maximizes the absolute value of  $\frac{|\Delta V(\bar{x}_i)|}{C(\bar{x}_i)}$ , where  $C(\bar{x}_i)$  is the computational cost of calculating the output for  $\bar{x}_i$ , and  $\Delta V(\bar{x}_i)$  is proportional the reduction in estimated variance by adding the correlations  $\bar{x}_i$  to the training set [74]. A similar approach could be used along with the formulas in this thesis, with the only difference being that Equation 1 would be used to calculate  $\Delta V(\bar{x}_i)$ .

An alternative, but related approach, is to treat this as a penalized minimization problem, similar to the problem of Tikhonov regularization. In this method, a penalty parameter  $\lambda$  is chosen, and the objective is to minimize following expression:

$$-Tr \left[ X_{new} \Omega_{old}^{-1} D \Omega_{old}^{-1} X_{new}^T \left( I + X_{new} \Omega_{old}^{-1} X_{new}^T \right)^{-1} \right] + \lambda C(X_{new}) \quad \text{Equation 174}$$

The parameter  $\lambda$  can be adjusted manually to balance the tradeoff between computational cost and improved predictive power. Once again, heuristic minimization algorithms can be used to find structures that give low values for Equation 174.

---

## III.4. Experiments

---

A good measure of the quality of a cluster expansion is the ability of the cluster expansion to predict properties of materials that were not included in the training set. When used to calculate thermodynamic averages, it is most important that the cluster expansion get the energy of low-energy structures correct. For phase diagrams, it is important that the cluster expansion correctly ranks the structure with the lowest formation energy at a given composition. Rather than focusing on specific applications, in this thesis a more general property of the cluster expansion that is important to many applications is tested: the ability to calculate the potential energy of structures that were not included in the training set.

### III.4.1. Sample data

---

To evaluate the various methods discussed in this thesis for generating cluster expansions, sample sets have been generated for three different binary material systems: Si-Ge, Ag-Au, and Pb-Ir. Each method was evaluated by selecting a training set from the much larger sample set and using the training set to fit a cluster expansion. The predictive power of the cluster expansion was then evaluated over the full sample set. To more accurately reflect the distribution of structures over the test domain, multiple distinct but symmetrically equivalent structures were allowed in the sample set. (This was accomplished by weighting each structure proportional to the number of distinct structures that were symmetrically equivalent to it, and then removing the symmetrically equivalent structures.) In this section the generation of the test and training data will be described.

#### III.4.1.1. Calculating input values

##### III.4.1.1.1. Generating the sample structures

Two of the systems tested (Ag-Au and Pb-Ir) were modeled by cluster expansions with face-centered cubic (FCC) symmetry. The training sets for these cluster expansions consisted of all structures with up to nine atoms in the structure basis. Using the methods described in this thesis, 1135 symmetrically distinct structures were found.

The remaining system, Si-Ge, was modeled by a cluster expansion with diamond cubic symmetry. All structures with up to fourteen atoms in the structure basis were included, for a total of 9631 structures.

#### III.4.1.1.2. Selecting training structures

For each set of input data, training sets with 15, 30, 45, 60, and 75 structures were evaluated. All training structures were selected from the sample set by using a greedy algorithm to find the structures that minimized Equation 1. The domain matrix was calculated directly from the entire sample set, and the regularization matrix was assumed to be zero for the purpose of selecting training structures. Because the regularization matrix was set to zero, it was necessary to ensure there were always at least as many training structures as clusters used for the structure selection algorithm. Otherwise, the matrix  $\Omega$  in Equation 1 is singular. To meet this requirement, the set of considered clusters for structure selection was increased as the set of training structures was increased.

#### **III.4.1.2. Calculating output values**

The cluster expansion method is generally used to model calculations that are otherwise computationally expensive. To avoid taking on an impractically large computational burden, the data sets used in this thesis were generated using fast multi-body potential models that are known to reproduce the properties of materials reasonably well. This section contains descriptions of the models used to generate the data. The General Utility Lattice Program (GULP) [94] was used to calculate the energies for the datasets using the potential models.

For all structures, the mean of the prior distribution for the potential energy was calculated as the linear average between the energies of the pure elements. This is equivalent to fitting the formation energies for the structures. For example, for all structures that were 25% silicon and 75% germanium, the mean for the prior distribution on the energy per atom was 25% of the energy per atom for pure silicon plus 75% of the energy per atom for pure germanium.

#### III.4.1.2.1. The embedded atom potential

The embedded atom potential is a simple multi-body potential that is well-suited for some metallic systems. The intuition behind the potential is that in addition to the pair-wise

interactions between nuclei, the nuclei also interact with the local electronic density due to nearby atoms. The local electronic density is calculated as a linear combination of contributions from nearby atoms, and the strength of the interaction with the electronic density is proportional to the square root of the density. The general form for the potential is:

$$V = \varepsilon_{ij} \left[ \frac{1}{2} \sum_i \sum_{j \neq i} \left( \frac{a_{ij}}{r_{ij}} \right)^{n_{ij}} - \sum_i c_i \sqrt{\sum_{j \neq i} \left( \frac{a_{ij}}{r_{ij}} \right)^{m_j}} \right] \quad \text{Equation 175}$$

where  $r_{ij}$  is the distance between nuclei  $i$  and  $j$ , and  $\varepsilon_{ij}$ ,  $c_i$ ,  $a_{ij}$ ,  $n_{ij}$ ,  $m_j$  are parameters that are dependent on the elements of nuclei  $i$  and  $j$ . In this thesis the parameters published by Sutton and Chen [95] have been used. Sutton and Chen list only self-interaction parameters for the elements. For interactions between different elements, the following combination rules used were used:

$$\begin{aligned} a_{ij} &= \frac{a_{ii} + a_{jj}}{2} \\ n_{ij} &= \frac{n_{ii} + n_{jj}}{2} \\ \varepsilon_{ij} &= \sqrt{\varepsilon_{ii} \varepsilon_{jj}} \end{aligned} \quad \text{Equation 176}$$

One of the challenges with the embedded atom model is that there is no distance at which the interaction between two atoms becomes zero. To make evaluation of the embedded atom method feasible, a cutoff radius may be chosen. Beyond this radius it is assumed that atoms have no interaction. In this thesis, a cutoff radius of 12 Å was used. The use of the cutoff radius can cause discontinuities in the potential energy surface, which may cause issues with convergence when searching for local minima.

The embedded-atom potential was used to generate data for two systems studied in this thesis. The first, face-centered cubic (FCC) Ag-Au, is an alloy between two elements with similar lattice parameters, and all systems relaxed to their local minima with forces converged to 1 meV / Å. The second system, FCC Pb-Ir, is an alloy of two elements with a significantly different FCC lattice parameters in their pure forms. In an attempt to resolve this lattice mismatch, there were more significant relaxations for Pb-Ir alloys. No effort was

made to check that the relaxed structures resembled the initial FCC structures, so there may have been significant reconstruction for some structures. There were 32 structures in which the forces did not converge within 1 meV / Å, but all converged within 10 meV / Å. It should be expected that the cluster expansion converges significantly better for the Ag-Au system than for the Pb-Ir system.

### III.4.1.2.2. The Tersoff potential

The Tersoff potential is a bond-order potential, in which the strength of the interaction between two atoms is dependent on the “order” of the bond between the atoms, which in turn is dependent on the positions of other nearby atoms [96]. It is a many-body potential that works well for diamond cubic materials such as silicon, germanium, and carbon in its diamond form. The form of the potential is given by:

$$V = \frac{1}{2} \sum_i \sum_{j \neq i} f_c(r_{ij}) (A_{ij} e^{-\lambda_{ij} r_{ij}} - b_{ij} B e^{-u_{ij} r_{ij}})$$

$$b_{ij} = \chi_{ij} \left( 1 + \beta_i^{n_i} \left[ \sum_{k \neq i, j} f_c(r_{ik}) \omega_{ik} \left( 1 + \frac{c_i^2}{d_i^2} - \frac{c_i^2}{d_i^2 + (h_i - \cos \theta_{ijk})^2} \right) \right]^{n_i} \right)^{-\frac{1}{2n_i}}$$

$$f_c(r_{ik}) = \begin{cases} 1, & r_{ij} < R_{ij} \\ \frac{1}{2} + \frac{1}{2} \cos \left( \pi \frac{(r_{ij} - R_{ij})}{S_{ij} - R_{ij}} \right), & R_{ij} < r_{ij} < S_{ij} \\ 0, & r_{ij} > S_{ij} \end{cases}$$

Equation 177

where all variables except  $r_{ij}$  and  $r_{ik}$  are element-specific parameters provided by Tersoff.

The Tersoff potential was used to calculate the energy of the Si-Ge structures. Because of the  $f_c(\cdot)$  term, there are no first-order discontinuities in this potential, and all calculations converged so that the maximum force was less than 1 meV / Å.

## III.4.2. Model selection

For all cluster expansions, parameters were chosen by attempting to minimize either the leave-one-out cross-validation (LOO CV) or generalized gross validation (GCV) scores. Selecting the set of parameters that minimizes one of these scores is in general a global

optimization problem which can not be analytically solved. The technique used to find a good set of parameters depended on the type of prior distribution.

For cluster selection, a rule was enforced that if a cluster was included in the fit, all subclusters must be included as well. This constraint can be compared to the constraint in the continuous prior distributions that clusters must have smaller prior widths than any of their subclusters. It was found that this constraint significantly improved the quality of the fit. For cluster expansions in which there were fewer than 30 candidate clusters, the set of clusters that minimized the score was found by an exhaustive search of all possible sets of clusters to include in the fit. For cluster expansions with more than 30 candidate structures, a Metropolis algorithm was run to search for the ground state. The Metropolis algorithm was run in multiple stages, with the number of attempted changes in each stage equal to 200 times the number of candidate clusters. Between stages, the temperature was dynamically adjusted to bring the ratio of successful flips closer to 10%. The algorithm was stopped when the ratio of successful flips was within 30% of 10%.

For the exponential, hierarchical, and constant width-generating functions, parameter selection was done in a two-stage process. The first stage of the process was a grid search for a local minimum on a logarithmic grid, in which neighboring grid points represented parameters that differed by a factor of 2. All parameters were initialized with a value of 1. The grid search was ended when the improvement in the score between neighboring points was less than  $1\text{E-}4$  eV. When the grid search was completed, a conjugate gradient algorithm was used to more finely resolve the local minimum. The conjugate gradient algorithm was stopped when the gradient of the score with respect to the natural log of the parameter values was less than  $1\text{E-}5$  eV.

For the individual width-generating function, the same method was used as the exponential, hierarchical, and constant width generating functions if there were fewer than six parameters. For situations in which there were six or more parameters, the multi-dimensional grid search rapidly becomes computationally expensive. In such situations, the grid search was skipped and only the conjugate gradient step was used.



### III.4.3. Cluster expansion options

To evaluate different methods for generating cluster expansions, a set of cluster expansions were generated by exploring all possible combinations of the options listed in Table 12. A total of 60,000 possible cluster expansions can be generated from the combinations of these options. The algorithm used to fit cluster expansions that use the Laplace prior requires that the input data be linearly independent, which restricted the number of cluster expansions that were successfully fit to 56,480.

Option	Allowed Values	Description
Data set	Pb-Ir, Ag-Au, Si-Ge	The binary material for which the input and output data were generated.
Nuclear positions	Initial, relaxed	Whether the nuclei should be allowed to relax to nearby local minima when calculating energies.
Training set size	15, 30, 45, 60, 75	The number of structures in the training set
Candidate cluster sets	{1, 0, 0}, {2, 2, 2}, {3, 0, 0}, {3, 2, 0}, {3, 3, 3}, {4, 3, 0}, {4, 4, 4}, {5, 0, 0}, {5, 4, 0}, {5, 5, 5}	The “candidate clusters” which were allowed to have non-zero ECI. Each set on the corresponds to the maximum distance between sites for pair clusters, triples, and quads, respectively. A value of “0” means none were included, a value of “1” means all clusters with no two sites further than next-nearest-neighbors were allowed, “2” means all clusters with no two sites further than next-nearest-neighbors were included, etc.

Add noise	True, false	Whether a random noise term, drawn from a Gaussian with standard deviation of 2 meV, should be added to the calculated energy.
Weight by symmetric degeneracy	True, false	Whether structures should be assigned a weight proportional to the number of symmetrically equivalent structures.
Prior distribution	Cluster selection, The Laplace prior, Gaussian priors, Iterated Gaussian priors	What type of prior distribution should be used. “Iterated Gaussian priors” is a method in which the Gaussian prior is used to fit the data, then all clusters with correlation coefficients below 0.1 meV were removed from consideration. This process was repeated until no clusters had correlation coefficients below 0.1 meV.
Width-generating function	Constant, independent, exponential, hierarchical	The width-generating functions are described in this thesis and were used to determine the width of Gaussian and Laplace prior distributions. For cluster selection, this option is not valid.
Regularized variable	ECI, correlation coefficients	Whether the prior distribution applied to the ECI or the correlation coefficients. For cluster selection, this option is not valid.
Cross-validation method	Leave-one-out cross-validation, generalized cross-validation	For the Laplace prior, the approximate generalized cross-validation score suggested by Tibshirani [86] was used.

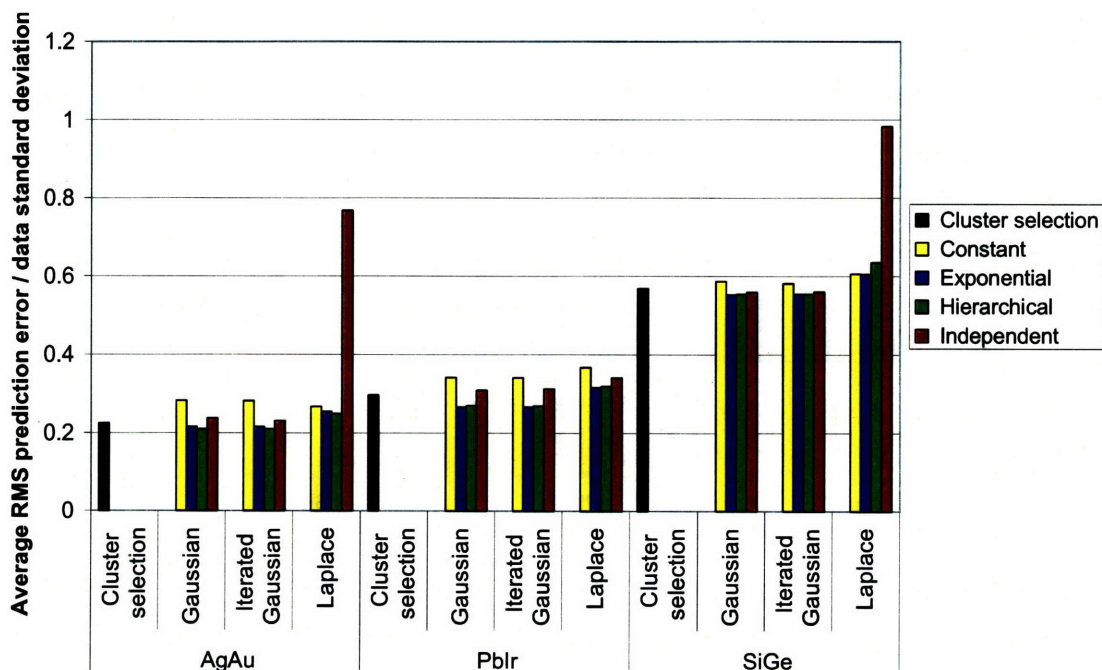
**Table 12** The different options used to generate cluster expansions evaluated in this thesis.

### III.4.4. Results

Because of the large number of cluster expansions generated, there are many ways in which the results can be analyzed. In this section the impact of the options and noteworthy combinations of options used to generate the cluster expansion will be examined.

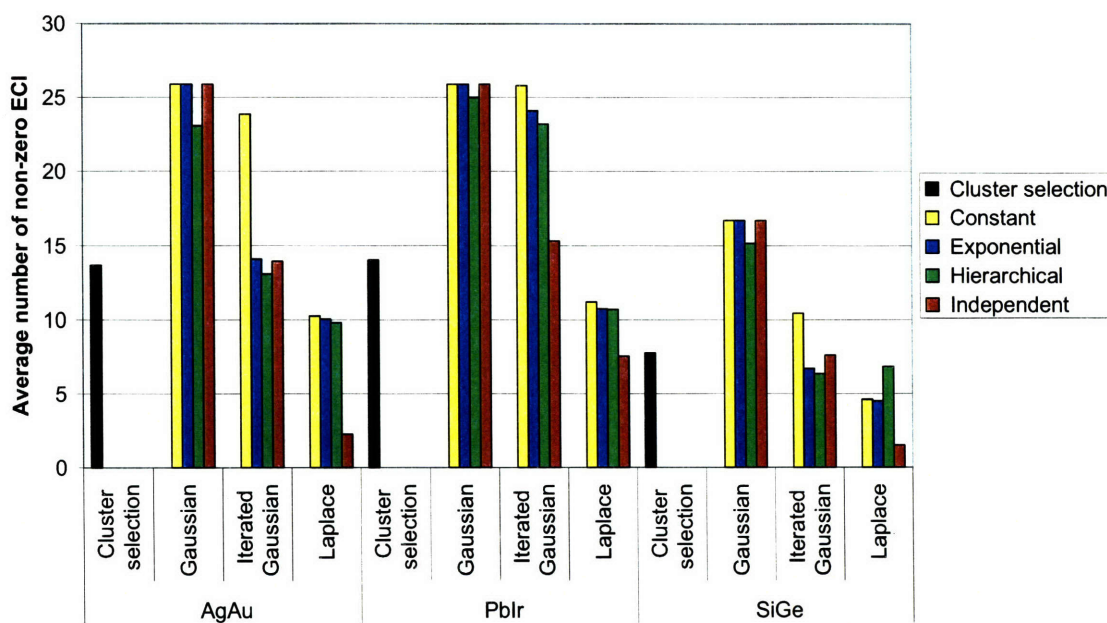
#### III.4.4.1. Prior distributions

Of the four different shapes of prior distributions considered, the combination of the Laplace prior with the independent width-generating function produced cluster expansions with anomalously poor predictive power (Figure 40). Because of the large number of parameters required for the independent width-generating function, minimization of the score typically relied on the conjugate gradient algorithm. However with the Laplace prior, the derivative of the score with respect to individual parameters is zero for parameters representing variables that are not in the active set, and it is discontinuous at the point at which the variables join the active set. It is possible that the conjugate gradient algorithm had difficulty navigating this surface for the Ag-Au and Si-Ge data sets.



**Figure 40** The average root-mean-square prediction error, relative to the data standard deviation, for different combinations of prior shapes and width-generating functions. The averages were taken over the calculated data with no additional noise added.

The speed at which the cluster expansion can determine the value of a given data point is determined by the number of clusters in the expansion with non-zero ECI. The cluster expansions based on the Gaussian prior tended to produce the most clusters with non-zero ECI (Figure 41). However, this number could be reduced significantly by using the iteration procedure to remove all ECI with values below 0.1 meV. For the Si-Ge and Ag-Au data sets, this method produced cluster expansions with about the same number of non-zero ECI as cluster selection. For the Pb-Ir data set, the iteration method did not trim the number of clusters as much, probably because the ECI for this data set were on average larger. Cluster expansions based on the Laplace prior tended to produce the fewest non-zero ECI. When combined with the independent width-generating function for the Ag-Au and Si-Ge data sets, very few clusters were assigned non-zero ECI. This helps explain why the prediction error of these cluster expansions is so poor, and is consistent with the conjugate gradient algorithm having a difficult time finding clusters that minimize the score.



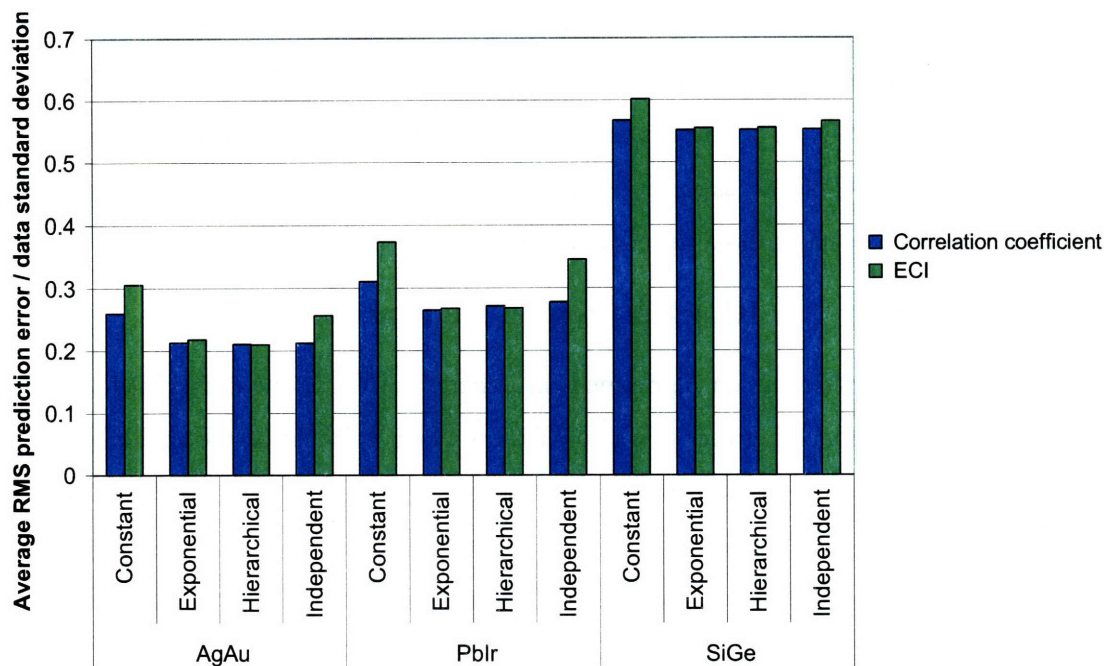
**Figure 41** The average number of cluster orbits with non-zero ECI for different combinations of prior shapes and width-generating functions. The averages were taken over the calculated data with no additional noise added.

Although the Laplace prior distribution produced the smallest cluster expansions, it also on average produced those with the worst predictive power. Some cluster expansions which

used the Laplace prior could not be fit because of constraints on the fitting algorithm. Because of the discontinuous nature of the Laplace prior, it is more general more difficult to work with analytically. For these reasons, the remainder of the results presented in this thesis will exclude all cluster expansions that relied on the Laplace prior. This is more likely to give meaningful data across a variety of cluster expansion options.

#### **III.4.4.2. Regularized variable**

The prior distribution can be defined on either the ECI or the correlation coefficient, with the only difference between the two of them being a factor of the multiplicity of the cluster orbit. Cluster expansions with both types of prior distributions were evaluated, and the predictive power of these expansions is given in Figure 42. There is little difference between defining a prior distribution on the ECI or the correlation coefficient, with the most significant impact occurring for the independent and constant width-generating functions. The result for the independent width-generating function is particularly interesting, because for this width-generating function the multiplicity factor should not affect the space of allowed prior distributions. Thus it is likely that the difference is due to the cross-validation minimization procedure starting at different initial points and navigating differently-scaled cross-validation score surfaces. This highlights the importance of evaluating different initial conditions and convergence criteria, especially for the constant and independent width-generating functions.



**Figure 42** The average root-mean-square prediction error, relative to the data standard deviation, for different width-generating functions applied to either the correlation coefficients or the ECI. The averages were taken over the calculated data with no additional noise added.

### III.4.4.3. Data sets

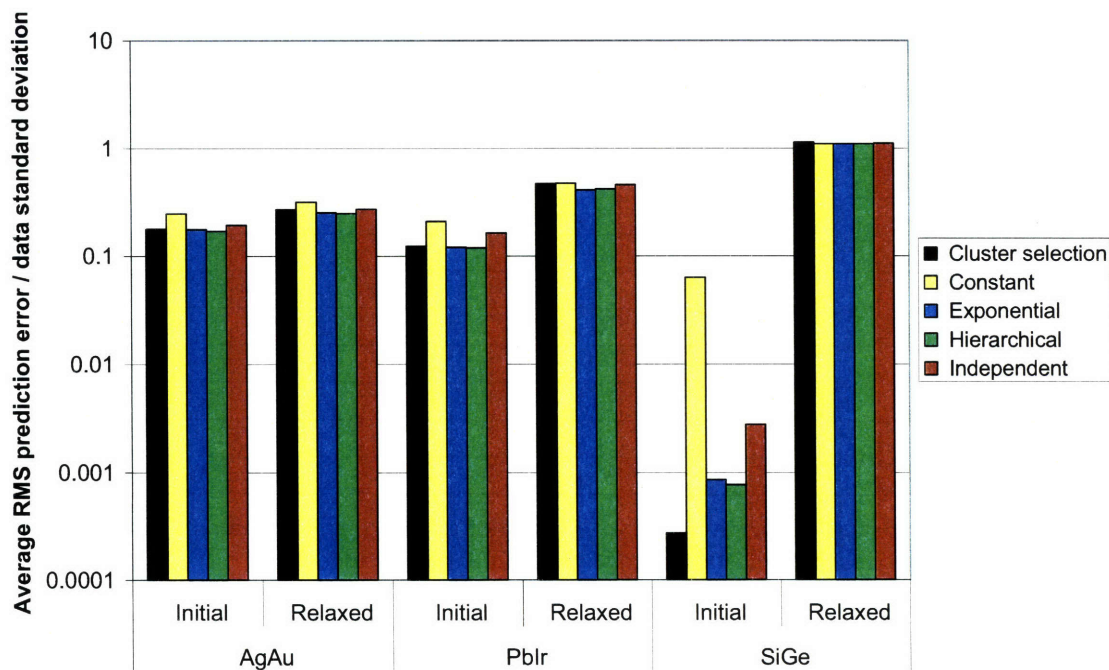
The standard deviations of the formation energies per atom in the data sets used for the cluster expansion are given in Table 13. The standard deviations vary significantly, with the Pb-Ir system having the largest standard deviation by far. This system was chosen because of the large difference between the parameters in the embedded atom potential for Pb and Ir, so it is not surprising that there is so much variation in the formation energies. On the other hand, the Si-Ge system is almost perfectly described by the mean of the prior distribution, resulting in very little standard deviation in the formation energies. An interesting result is that in both the Si-Ge and Pb-Ir systems, the relaxed structures had lower standard deviation than the unrelaxed systems. The near-linear correlation between energies and composition in these relaxed systems suggests that volume effects may significantly contribute to the energetics.



Data Set	Test set weights	Nuclear Positions	Standard Deviation (eV / atom)
Si-Ge	Weighted	Initial	0.0029
		Relaxed	0.0020
	Unweighted	Initial	0.0033
		Relaxed	0.0024
Ag-Au	Weighted	Initial	0.0110
		Relaxed	0.0123
	Unweighted	Initial	0.0131
		Relaxed	0.0144
Pb-Ir	Weighted	Initial	0.6294
		Relaxed	0.4662
	Unweighted	Initial	0.7392
		Relaxed	0.5332

**Table 13** The standard deviations of the formation energies in the different test sets used. Values are given for the formation energies without any noise added.

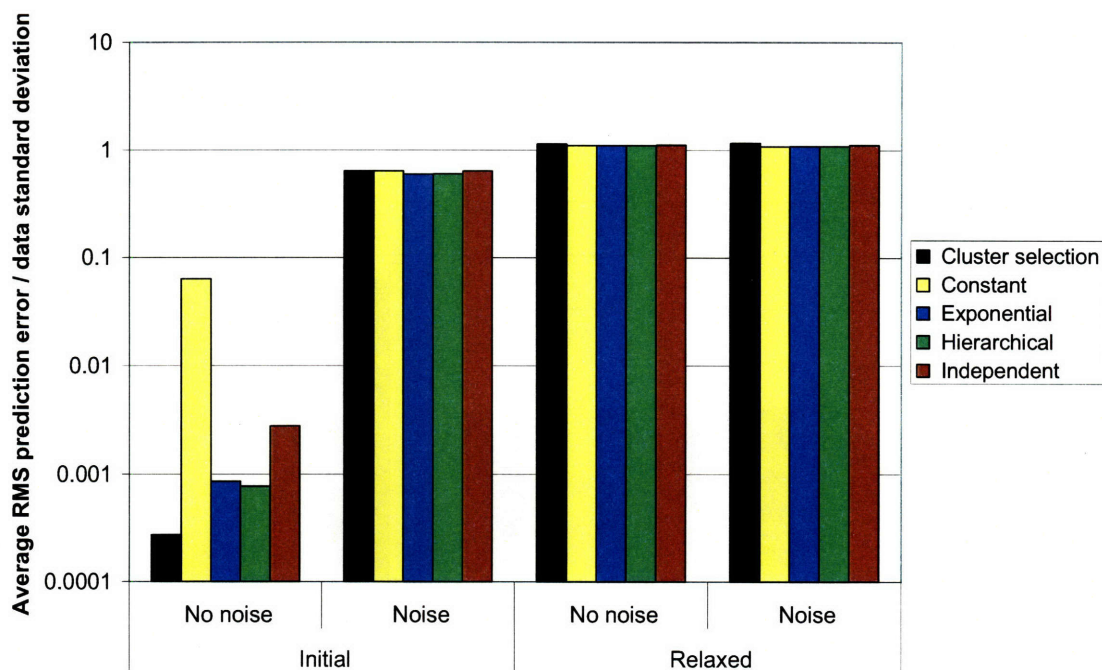
Because the standard deviations for the different data sets vary so widely, results will in general be presented relative to the standard deviation for the given test set. If the cluster expansion is replaced by a simple function that returns the average formation energy of the test set, the root-mean-squared (RMS) prediction error of the cluster expansion should be equal to the standard deviation of the test set. In general, cluster expansions should perform much better than this. An unexpected result, likely due to the low standard deviation for the relaxed Si-Ge system, is that on average the RMS prediction error for the cluster expansions on the relaxed Si-Ge system was close to the standard deviation of the formation energies (Figure 43). In contrast, the cluster expansion performed very well on the Si-Ge system with the nuclei in ideal diamond-cubic lattice positions.



**Figure 43** The average root-mean-square prediction error relative to the standard deviation of the data. Values are given for the formation energies without any noise added. A logarithmic scale is used.

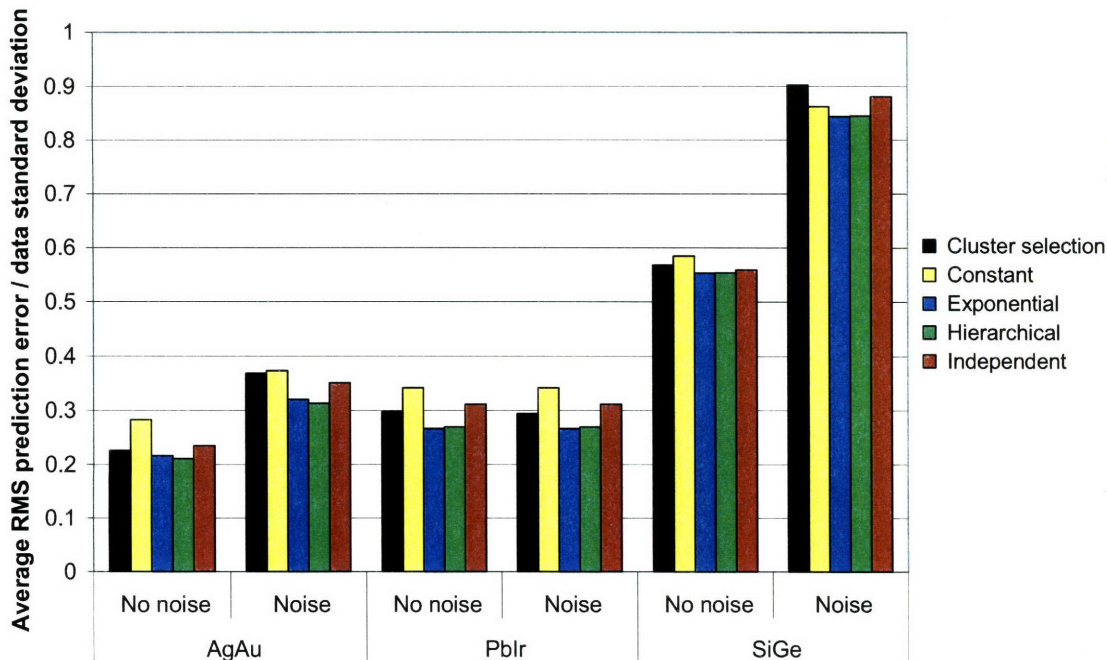
One possible explanation for the unexpectedly poor fits for the relaxed Si-Ge data is that in the relaxation process, there is a certain amount of noise that cannot be modeled by the clusters considered for the fit. Because the standard deviation of the relaxed Si-Ge system is so small, this noise becomes relatively more apparent. To support this hypothesis, the results for Si-Ge fit to the calculated data, and the calculated data plus manually added noise (drawn from a Gaussian of with 2 meV), are shown in Figure 44. The addition of the noise makes the results appear more similar to the results of the relaxed structure.





**Figure 44** The average root-mean-square prediction error relative to the standard deviation of the data for the Si-Ge system. Results are plotted for both nuclei in their initial and relaxed positions, and for the data with and without the addition of noise drawn from a 2 meV wide Gaussian. A logarithmic scale is used.

The addition of the noise is clearly most significant in Si-Ge, as opposed to the others in which the noise is smaller relative to the inherent fluctuations in the data (Figure 45). Within a given data set, the addition of noise generally makes the fit worse overall, and seems to affect the cluster selection method the most. In particular, the addition of noise seemed to be most detrimental to the cluster selection method for the Ag-Au data set. This may be due to the fact that the noise is neither too small to be very noticeable (as in Pb-Ir), nor so large that it overwhelms the data (as in Si-Ge).



**Figure 45** The effect of adding noise to the formation energies on the RMS prediction error, relative to the standard deviation, in each of the data sets.

#### III.4.4.4. Weighting

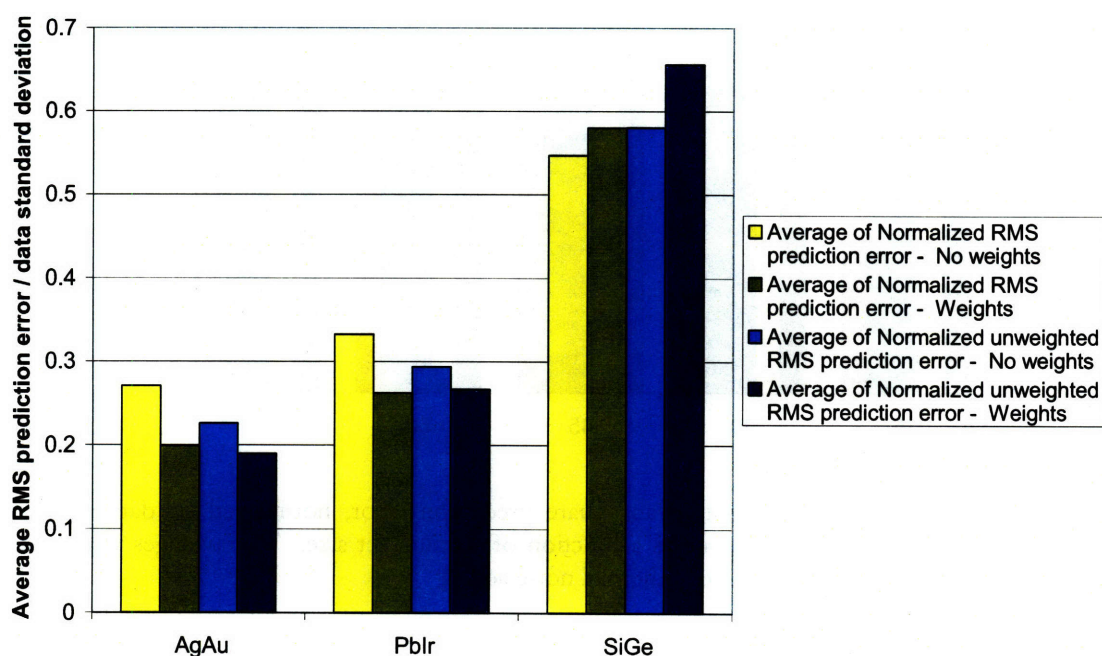
Often all training structures for a cluster expansion are given equal weight. For a variety of reasons, structures may be assigned unequal weights in an attempt to improve the predictive power of the fit. For example, it might make sense to assign a weight to a structure in the training set that is proportional to the number of symmetrically equivalent structures in the test domain. A structure with very high symmetry will generally have fewer distinct symmetrically equivalent structures in the test domain, whereas a structure with low symmetry will have many distinct symmetrically equivalent structures.

Cluster expansions were generated from both the situation in which all training structures had equal weights and the situation in which all training structures are given weights proportional to the number of symmetrically equivalent structures. These cluster expansions were then evaluated against two test sets:

1. A test set in which all structures up to a certain size were included, even if multiple structures were symmetrically equivalent.

2. A test set in which structures were removed so that no two structures in the set were symmetrically equivalent.

The predictive power of these cluster expansions are shown in Figure 46. In both the FCC training sets, the weighted fits outperform the unweighted fits, whereas in the Si-Ge training set the unweighted fit does better. However, the data do not paint a clear and consistent picture of how assigning weights to the structures affects the quality of the fit.



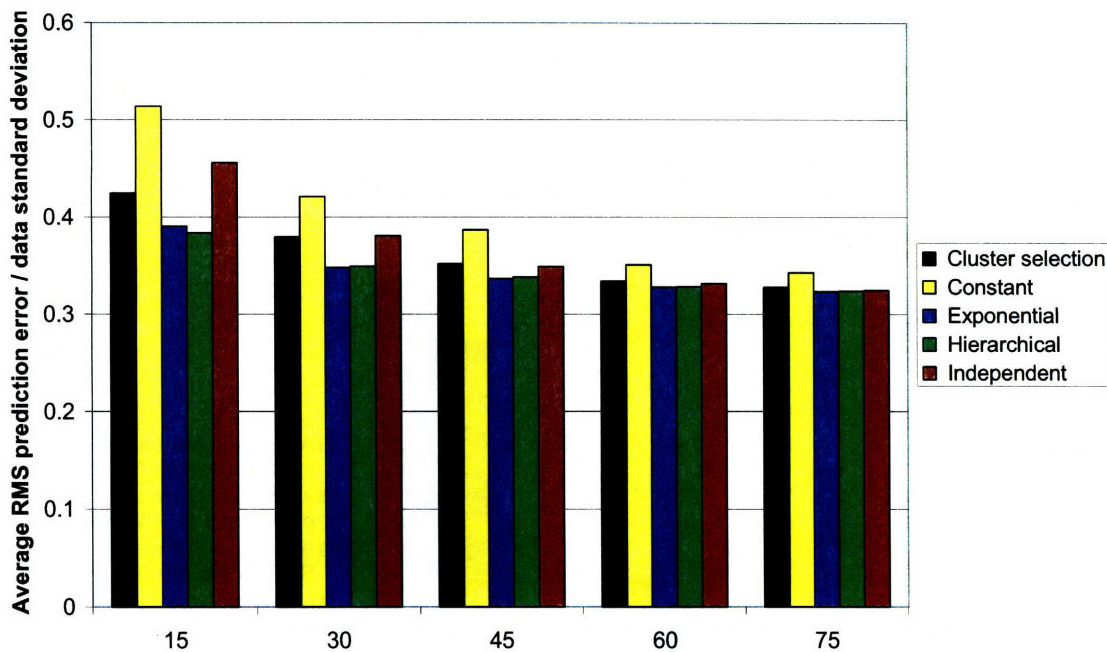
**Figure 46** The average root mean square prediction error, normalized by data standard deviation, for cluster expansions generated from both weighted and unweighted training sets. The unweighted prediction error was evaluated on test sets in which only one representative from each orbit of symmetrically equivalent structures was included. Values are given for the formation energies without any noise added.

### III.4.4.5. Convergence

Because training data can be so expensive to generate, one of the key challenges for a cluster expansion is to rapidly converge to a model with good predictive power. The average predictive power of cluster expansions, as a function of training set size, is shown in Figure



47. The average is taken over all data sets, for which the trends were similar. In this chart the advantages of the more physically meaningful width-generating functions start to become more apparent.



**Figure 47** The average root mean square prediction error, normalized by data standard deviation, for cluster expansions as a function of training set size. The averages were taken over the calculated data with no additional noise added.

A more detailed breakdown of the relative predictive power of the cluster expansions as a function of training set size, the set of candidate clusters, and the width-generating function used is given in Table 16 - Table 21. Some of the key results are summarized in Table 14, Table 15, and Table 22.

The average predictive power for cluster expansions built using the exponential and hierarchical width-generating functions is the best across all data sets (Table 14). The exponential and hierarchical width-generating functions perform similarly to each other, which is not surprising considering their similar form. Cluster selection and the independent width-generating function were the next best on average, about 10% worse than the exponential and hierarchical width-generating functions. The constant width-generating

function performed the worst, about 20% behind the exponential and hierarchical width-generating functions.

		Cluster selection	Constant	Exponential	Hierarchical	Independent
Ag-Au	No noise	1.072	1.345	1.026	1.000	1.114
	Noise	1.177	1.191	1.021	1.000	1.122
Pb-Ir	No noise	1.117	1.280	1.000	1.013	1.169
	Noise	1.103	1.281	1.000	1.011	1.168
Si-Ge	No noise	1.027	1.059	1.000	1.001	1.011
	Noise	1.071	1.022	1.000	1.001	1.044
All sets		1.091	1.191	1.004	1.000	1.101

**Table 14** The average relative root mean square prediction error of cluster expansions generated using each of the width-generating functions, for each of the data sets.

The cluster selection method most frequently had the best predictive power, followed by the hierarchical, exponential, and constant width-generating functions (Table 15). The independent width-generating function rarely had the best predictive power.

		Cluster selection	Constant	Exponential	Hierarchical	Independent
Ag-Au	No noise	50%	16%	14%	14%	6%
	Noise	38%	14%	12%	36%	0%
Pb-Ir	No noise	34%	18%	18%	30%	0%
	Noise	32%	18%	20%	30%	0%
Si-Ge	No noise	14%	28%	16%	10%	32%
	Noise	6%	20%	44%	22%	8%
Average		29.00%	19.00%	20.67%	23.67%	7.67%

**Table 15** The percentage sets of candidate clusters for which the given width-generating function produced the cluster expansion with the lowest average root mean square prediction error.



Training set size	Pair cutoff	Triple cutoff	Quad cutoff	RMS prediction error relative to row minimum					Minimum RMS prediction error / data standard deviation
				Cluster selection	Constant	Exponential	Hierarchical	Independent	
15	NN	None	None	1.000	1.002	1.002	1.000	1.000	0.962
	3NN	3NN	3NN	1.046	1.000	1.048	1.044	1.025	0.471
	4NN	None	None	1.000	1.024	1.025	1.029	1.037	0.197
	4NN	3NN	None	1.000	1.020	1.020	1.005	1.009	0.212
	4NN	4NN	4NN	3.950	3.615	1.000	1.055	2.054	0.156
	5NN	4NN	None	1.000	1.052	1.021	1.065	1.064	0.065
	5NN	5NN	5NN	1.151	7.932	3.734	1.000	4.296	0.072
	6NN	None	None	1.005	1.000	1.004	1.021	1.022	0.065
	6NN	5NN	None	1.604	2.986	1.000	1.172	2.134	0.058
6NN	6NN	6NN	1.593	8.153	2.511	1.000	6.315	0.074	
Average for training set of 15				1.219	1.635	1.122	1.000	1.351	0.238
30	NN	None	None	1.000	1.000	1.000	1.001	1.001	0.881
	3NN	3NN	3NN	1.025	1.000	1.011	1.008	1.012	0.452
	4NN	None	None	1.000	1.016	1.016	1.011	1.015	0.188
	4NN	3NN	None	1.001	1.010	1.008	1.002	1.000	0.208
	4NN	4NN	4NN	1.016	1.038	1.016	1.028	1.000	0.128
	5NN	4NN	None	1.000	1.018	1.014	1.022	1.019	0.063
	5NN	5NN	5NN	2.505	7.256	1.000	1.102	2.495	0.059
	6NN	None	None	1.000	1.009	1.010	1.023	1.027	0.064
	6NN	5NN	None	1.011	1.013	1.010	1.000	1.049	0.040
6NN	6NN	6NN	1.621	10.342	1.000	1.087	4.134	0.053	
Average for training set of 30				1.057	1.402	1.000	1.005	1.119	0.215
45	NN	None	None	1.000	1.000	1.000	1.000	1.000	0.855
	3NN	3NN	3NN	1.044	1.000	1.013	1.015	1.017	0.448
	4NN	None	None	1.000	1.013	1.011	1.010	1.013	0.174
	4NN	3NN	None	1.001	1.011	1.009	1.002	1.000	0.198
	4NN	4NN	4NN	1.000	1.090	1.072	1.084	1.031	0.109
	5NN	4NN	None	1.000	1.018	1.021	1.031	1.023	0.060
	5NN	5NN	5NN	1.249	5.743	1.000	1.117	1.266	0.051
	6NN	None	None	1.000	1.007	1.007	1.023	1.017	0.060
	6NN	5NN	None	1.002	1.000	1.014	1.007	1.015	0.038
6NN	6NN	6NN	2.269	11.488	1.088	1.000	3.055	0.039	
Average for training set of 45				1.029	1.312	1.000	1.002	1.042	0.205
60	NN	None	None	1.000	1.000	1.000	1.000	1.000	0.834
	3NN	3NN	3NN	1.040	1.000	1.012	1.015	1.016	0.434
	4NN	None	None	1.000	1.008	1.009	1.008	1.010	0.165
	4NN	3NN	None	1.000	1.007	1.008	1.001	1.003	0.186
	4NN	4NN	4NN	1.000	1.037	1.027	1.034	1.008	0.108
	5NN	4NN	None	1.000	1.005	1.007	1.021	1.012	0.059



Training set size	Pair cutoff	Triple cutoff	Quad cutoff	RMS prediction error relative to row minimum					Minimum RMS prediction error / data standard deviation
				Cluster selection	Constant	Exponential	Hierarchical	Independent	
	5NN	5NN	5NN	1.059	1.153	1.002	1.000	1.062	0.046
	6NN	None	None	1.000	1.004	1.004	1.014	1.013	0.059
	6NN	5NN	None	1.000	1.001	1.009	1.009	1.019	0.037
	6NN	6NN	6NN	1.796	11.295	1.007	1.000	1.667	0.034
Average for training set of 60				1.018	1.180	1.000	1.001	1.013	0.197
75	NN	None	None	1.000	1.000	1.000	1.000	1.000	0.819
	3NN	3NN	3NN	1.033	1.000	1.004	1.005	1.010	0.428
	4NN	None	None	1.000	1.007	1.007	1.008	1.009	0.159
	4NN	3NN	None	1.000	1.004	1.004	1.001	1.000	0.184
	4NN	4NN	4NN	1.000	1.039	1.027	1.034	1.013	0.109
	5NN	4NN	None	1.000	1.010	1.009	1.015	1.015	0.057
	5NN	5NN	5NN	1.056	1.030	1.000	1.014	1.025	0.041
	6NN	None	None	1.000	1.008	1.004	1.012	1.015	0.059
	6NN	5NN	None	1.000	1.019	1.025	1.025	1.033	0.036
	6NN	6NN	6NN	1.866	9.353	1.043	1.000	1.347	0.032
Average for training set of 75				1.018	1.138	1.000	1.000	1.007	0.194
Average over all trials				1.072	1.345	1.026	1.000	1.114	0.210

**Table 16** The relative predictive powers for the **Ag-Au** data set, **with no noise** added, for the five different width-generating functions as a function of training set size and candidate cluster sets. This table may be read in the following way: In each row, the first four columns specify the combination of training set size and candidate cluster set. The next five columns are the average RMS prediction error for each of the different width-generating functions, expressed relative to the minimum RMS prediction error for the row. Green cells are the minimum values for the row, and blue cells are within 3% of the minimum. Red cells are the worst value for the row, as long as the worst value is not within 3% of the minimum. The final column gives the value, relative to the data standard deviation, of the minimum RMS prediction error for the row.



Training set size	Pair cutoff	Triple cutoff	Quad cutoff	RMS prediction error relative to row minimum					Minimum RMS prediction error / data standard deviation
				Cluster selection	Constant	Exponential	Hierarchical	Independent	
15	NN	None	None	1.000	1.001	1.001	1.000	1.000	0.988
	3NN	3NN	3NN	1.034	1.000	1.042	1.038	1.017	0.513
	4NN	None	None	1.000	1.012	1.012	1.008	1.016	0.271
	4NN	3NN	None	1.000	1.023	1.020	1.013	1.020	0.323
	4NN	4NN	4NN	2.374	2.079	1.000	1.057	1.842	0.289
	5NN	4NN	None	1.067	1.076	1.036	1.000	1.041	0.220
	5NN	5NN	5NN	1.985	2.650	1.773	1.000	2.229	0.227
	6NN	None	None	1.020	1.002	1.000	1.005	1.013	0.201
	6NN	5NN	None	1.456	1.434	1.037	1.000	1.459	0.203
6NN	6NN	6NN	1.650	2.871	1.519	1.000	2.472	0.216	
Average for training set of 15				1.242	1.333	1.084	1.000	1.264	0.350
30	NN	None	None	1.001	1.000	1.000	1.001	1.001	0.904
	3NN	3NN	3NN	1.028	1.000	1.006	1.000	1.005	0.484
	4NN	None	None	1.000	1.007	1.007	1.004	1.010	0.267
	4NN	3NN	None	1.000	1.009	1.009	1.003	1.002	0.303
	4NN	4NN	4NN	1.095	1.128	1.008	1.000	1.075	0.243
	5NN	4NN	None	1.018	1.034	1.010	1.000	1.021	0.194
	5NN	5NN	5NN	2.404	2.417	1.000	1.021	1.916	0.202
	6NN	None	None	1.000	1.002	1.002	1.002	1.006	0.190
	6NN	5NN	None	1.044	1.051	1.018	1.000	1.029	0.191
6NN	6NN	6NN	3.068	2.797	1.037	1.000	2.129	0.210	
Average for training set of 30				1.237	1.222	1.005	1.000	1.141	0.320
45	NN	None	None	1.000	1.000	1.000	1.000	1.000	0.883
	3NN	3NN	3NN	1.030	1.000	1.009	1.004	1.010	0.473
	4NN	None	None	1.000	1.006	1.007	1.004	1.007	0.250
	4NN	3NN	None	1.000	1.006	1.008	1.006	1.001	0.279
	4NN	4NN	4NN	1.037	1.110	1.003	1.000	1.036	0.219
	5NN	4NN	None	1.000	1.016	1.005	1.007	1.009	0.190
	5NN	5NN	5NN	1.545	2.151	1.044	1.000	1.678	0.190
	6NN	None	None	1.000	1.007	1.006	1.008	1.006	0.189
	6NN	5NN	None	1.021	1.029	1.002	1.000	1.021	0.187
6NN	6NN	6NN	2.859	2.519	1.023	1.000	1.639	0.189	
Average for training set of 45				1.155	1.175	1.005	1.000	1.086	0.306
60	NN	None	None	1.001	1.000	1.000	1.001	1.001	0.855
	3NN	3NN	3NN	1.035	1.000	1.012	1.013	1.010	0.457
	4NN	None	None	1.000	1.004	1.005	1.004	1.005	0.242
	4NN	3NN	None	1.000	1.007	1.008	1.008	1.003	0.263
	4NN	4NN	4NN	1.017	1.050	1.000	1.003	1.014	0.217
	5NN	4NN	None	1.000	1.016	1.004	1.001	1.006	0.188



Training set size	Pair cutoff	Triple cutoff	Quad cutoff	RMS prediction error relative to row minimum					Minimum RMS prediction error / data standard deviation
				Cluster selection	Constant	Exponential	Hierarchical	Independent	
	5NN	5NN	5NN	1.186	1.358	1.004	1.000	1.268	0.188
	6NN	None	None	1.000	1.006	1.004	1.005	1.005	0.189
	6NN	5NN	None	1.024	1.022	1.009	1.000	1.019	0.183
	6NN	6NN	6NN	2.916	2.400	1.005	1.000	1.571	0.186
Average for training set of 60				1.136	1.113	1.001	1.000	1.054	0.298
75	NN	None	None	1.000	1.000	1.000	1.001	1.000	0.837
	3NN	3NN	3NN	1.023	1.000	1.002	1.002	1.006	0.453
	4NN	None	None	1.000	1.003	1.004	1.003	1.004	0.239
	4NN	3NN	None	1.000	1.005	1.005	1.007	1.001	0.259
	4NN	4NN	4NN	1.000	1.033	1.008	1.011	1.003	0.216
	5NN	4NN	None	1.000	1.013	1.006	1.007	1.007	0.185
	5NN	5NN	5NN	1.079	1.205	1.000	1.002	1.138	0.185
	6NN	None	None	1.000	1.005	1.004	1.004	1.004	0.187
	6NN	5NN	None	1.015	1.009	1.004	1.000	1.007	0.180
	6NN	6NN	6NN	2.490	2.119	1.011	1.000	1.465	0.184
Average for training set of 75				1.100	1.085	1.000	1.000	1.038	0.293
Average over all trials				1.177	1.191	1.021	1.000	1.122	0.313

**Table 17** The relative predictive powers for the **Ag-Au** data set, **with noise** added, for the five different width-generating functions as a function of training set size and candidate cluster sets. This table may be read in the following way: In each row, the first four columns specify the combination of training set size and candidate cluster set. The next five columns are the average RMS prediction error for each of the different width-generating functions, expressed relative to the minimum RMS prediction error for the row. Green cells are the minimum values for the row, and blue cells are within 3% of the minimum. Red cells are the worst value for the row, as long as the worst value is not within 3% of the minimum. The final column gives the value, relative to the data standard deviation, of the minimum RMS prediction error for the row.



Training set size	Pair cutoff	Triple cutoff	Quad cutoff	RMS prediction error relative to row minimum					Minimum RMS prediction error / data standard deviation
				Cluster selection	Constant	Exponential	Hierarchical	Independent	
15	NN	None	None	1.000	1.010	1.010	1.020	1.020	0.558
	3NN	3NN	3NN	1.054	1.000	1.008	1.059	1.047	0.323
	4NN	None	None	1.000	1.027	1.022	1.058	1.059	0.392
	4NN	3NN	None	1.000	1.009	1.003	1.036	1.092	0.247
	4NN	4NN	4NN	1.000	2.485	1.051	1.489	2.290	0.266
	5NN	4NN	None	1.032	1.055	1.000	1.383	1.163	0.188
	5NN	5NN	5NN	1.879	2.616	1.259	1.000	2.467	0.275
	6NN	None	None	1.000	1.049	1.027	1.075	1.075	0.376
	6NN	5NN	None	1.649	1.864	1.062	1.000	1.519	0.264
6NN	6NN	6NN	1.498	2.831	1.232	1.000	2.537	0.259	
Average for training set of 15				1.113	1.420	1.000	1.031	1.367	0.334
30	NN	None	None	1.000	1.004	1.004	1.007	1.007	0.519
	3NN	3NN	3NN	1.008	1.000	1.003	1.003	1.014	0.302
	4NN	None	None	1.000	1.018	1.013	1.030	1.031	0.352
	4NN	3NN	None	1.004	1.007	1.006	1.000	1.022	0.221
	4NN	4NN	4NN	1.023	1.237	1.016	1.000	1.022	0.197
	5NN	4NN	None	1.027	1.011	1.003	1.000	1.012	0.176
	5NN	5NN	5NN	2.697	3.358	1.000	1.038	2.075	0.172
	6NN	None	None	1.000	1.014	1.003	1.027	1.036	0.351
	6NN	5NN	None	1.054	1.037	1.000	1.149	1.030	0.180
6NN	6NN	6NN	2.569	3.634	1.000	1.029	3.076	0.179	
Average for training set of 30				1.219	1.352	1.000	1.019	1.223	0.266
45	NN	None	None	1.000	1.002	1.002	1.004	1.004	0.508
	3NN	3NN	3NN	1.006	1.000	1.004	1.002	1.009	0.297
	4NN	None	None	1.000	1.009	1.006	1.016	1.018	0.330
	4NN	3NN	None	1.004	1.005	1.003	1.000	1.018	0.219
	4NN	4NN	4NN	1.037	1.000	1.057	1.051	1.053	0.160
	5NN	4NN	None	1.014	1.001	1.006	1.000	1.012	0.176
	5NN	5NN	5NN	1.728	3.025	1.012	1.000	1.257	0.154
	6NN	None	None	1.000	1.007	1.001	1.016	1.026	0.333
	6NN	5NN	None	1.020	1.013	1.000	1.006	1.002	0.173
6NN	6NN	6NN	2.522	3.718	1.000	1.042	2.595	0.147	
Average for training set of 45				1.133	1.281	1.000	1.004	1.116	0.251
60	NN	None	None	1.000	1.001	1.001	1.002	1.002	0.489
	3NN	3NN	3NN	1.003	1.000	1.004	1.003	1.005	0.290
	4NN	None	None	1.000	1.003	1.001	1.008	1.012	0.325
	4NN	3NN	None	1.003	1.004	1.004	1.000	1.010	0.217
	4NN	4NN	4NN	1.011	1.000	1.057	1.051	1.033	0.151
5NN	4NN	None	1.007	1.005	1.006	1.000	1.011	0.175	



Training set size	Pair cutoff	Triple cutoff	Quad cutoff	RMS prediction error relative to row minimum					Minimum RMS prediction error / data standard deviation
				Cluster selection	Constant	Exponential	Hierarchical	Independent	
	5NN	5NN	5NN	1.113	1.125	1.006	1.000	1.134	0.148
	6NN	None	None	1.000	1.012	1.009	1.020	1.023	0.312
	6NN	5NN	None	1.000	1.005	1.009	1.002	1.005	0.176
	6NN	6NN	6NN	2.064	3.563	1.000	1.019	1.454	0.136
Average for training set of 60				1.061	1.147	1.000	1.001	1.036	0.244
75	NN	None	None	1.000	1.001	1.001	1.002	1.002	0.480
	3NN	3NN	3NN	1.006	1.000	1.002	1.002	1.007	0.282
	4NN	None	None	1.000	1.006	1.005	1.010	1.013	0.312
	4NN	3NN	None	1.002	1.003	1.003	1.000	1.007	0.213
	4NN	4NN	4NN	1.009	1.000	1.022	1.018	1.021	0.150
	5NN	4NN	None	1.003	1.005	1.005	1.000	1.009	0.174
	5NN	5NN	5NN	1.055	1.087	1.000	1.008	1.079	0.139
	6NN	None	None	1.003	1.003	1.000	1.008	1.011	0.310
	6NN	5NN	None	1.004	1.000	1.008	1.002	1.006	0.174
6NN	6NN	6NN	1.852	3.519	1.015	1.000	1.219	0.128	
Average for training set of 75				1.047	1.139	1.000	1.000	1.020	0.237
Average over all trials				1.117	1.280	1.000	1.013	1.169	0.266

**Table 18** The relative predictive powers for the **Pb-Ir** data set, **with no noise** added, for the five different width-generating functions as a function of training set size and candidate cluster sets. This table may be read in the following way: In each row, the first four columns specify the combination of training set size and candidate cluster set. The next five columns are the average RMS prediction error for each of the different width-generating functions, expressed relative to the minimum RMS prediction error for the row. Green cells are the minimum values for the row, and blue cells are within 3% of the minimum. Red cells are the worst value for the row, as long as the worst value is not within 3% of the minimum. The final column gives the value, relative to the data standard deviation, of the minimum RMS prediction error for the row.



Training set size	Pair cutoff	Triple cutoff	Quad cutoff	RMS prediction error relative to row minimum					Minimum RMS prediction error / data standard deviation
				Cluster selection	Constant	Exponential	Hierarchical	Independent	
15	NN	None	None	1.000	1.010	1.010	1.019	1.020	0.559
	3NN	3NN	3NN	1.055	1.000	1.009	1.058	1.048	0.323
	4NN	None	None	1.000	1.027	1.022	1.057	1.059	0.393
	4NN	3NN	None	1.000	1.009	1.003	1.037	1.093	0.247
	4NN	4NN	4NN	1.266	2.471	1.000	1.377	2.283	0.267
	5NN	4NN	None	1.030	1.055	1.000	1.381	1.155	0.189
	5NN	5NN	5NN	1.615	2.614	1.254	1.000	2.490	0.276
	6NN	None	None	1.000	1.050	1.027	1.075	1.075	0.376
	6NN	5NN	None	1.338	1.879	1.070	1.000	1.524	0.262
6NN	6NN	6NN	2.059	2.807	1.221	1.000	2.488	0.262	
Average for training set of 15				1.137	1.426	1.000	1.026	1.372	0.333
30	NN	None	None	1.000	1.004	1.004	1.007	1.007	0.519
	3NN	3NN	3NN	1.008	1.000	1.003	1.003	1.014	0.302
	4NN	None	None	1.000	1.018	1.013	1.030	1.031	0.353
	4NN	3NN	None	1.004	1.007	1.006	1.000	1.022	0.221
	4NN	4NN	4NN	1.024	1.238	1.017	1.000	1.019	0.197
	5NN	4NN	None	1.030	1.011	1.003	1.000	1.013	0.177
	5NN	5NN	5NN	2.636	3.339	1.000	1.023	2.092	0.173
	6NN	None	None	1.000	1.014	1.003	1.027	1.036	0.350
	6NN	5NN	None	1.050	1.038	1.000	1.151	1.030	0.180
6NN	6NN	6NN	2.193	3.619	1.000	1.025	3.082	0.180	
Average for training set of 30				1.191	1.351	1.000	1.018	1.225	0.266
45	NN	None	None	1.000	1.002	1.002	1.004	1.004	0.508
	3NN	3NN	3NN	1.006	1.000	1.004	1.002	1.009	0.297
	4NN	None	None	1.000	1.009	1.006	1.016	1.018	0.330
	4NN	3NN	None	1.004	1.005	1.003	1.000	1.018	0.219
	4NN	4NN	4NN	1.012	1.000	1.056	1.050	1.055	0.160
	5NN	4NN	None	1.015	1.001	1.006	1.000	1.012	0.176
	5NN	5NN	5NN	1.459	3.032	1.015	1.000	1.273	0.154
	6NN	None	None	1.000	1.007	1.001	1.016	1.026	0.333
	6NN	5NN	None	1.019	1.012	1.000	1.005	1.000	0.173
6NN	6NN	6NN	1.905	3.693	1.000	1.025	2.334	0.148	
Average for training set of 45				1.079	1.280	1.000	1.003	1.102	0.252
60	NN	None	None	1.000	1.001	1.001	1.002	1.002	0.489
	3NN	3NN	3NN	1.003	1.000	1.004	1.003	1.005	0.290
	4NN	None	None	1.000	1.003	1.001	1.008	1.012	0.325
	4NN	3NN	None	1.003	1.004	1.004	1.000	1.010	0.217
	4NN	4NN	4NN	1.011	1.000	1.060	1.054	1.035	0.151
	5NN	4NN	None	1.007	1.005	1.007	1.000	1.011	0.175



Training set size	Pair cutoff	Triple cutoff	Quad cutoff	RMS prediction error relative to row minimum					Minimum RMS prediction error / data standard deviation
				Cluster selection	Constant	Exponential	Hierarchical	Independent	
	5NN	5NN	5NN	1.103	1.120	1.004	1.000	1.135	0.148
	6NN	None	None	1.000	1.012	1.009	1.020	1.023	0.312
	6NN	5NN	None	1.000	1.006	1.009	1.003	1.005	0.176
	6NN	6NN	6NN	1.834	3.560	1.000	1.018	1.430	0.136
Average for training set of 60				1.047	1.147	1.000	1.001	1.035	0.244
75	NN	None	None	1.000	1.001	1.001	1.002	1.002	0.480
	3NN	3NN	3NN	1.006	1.000	1.002	1.002	1.007	0.282
	4NN	None	None	1.000	1.006	1.005	1.010	1.013	0.313
	4NN	3NN	None	1.002	1.003	1.003	1.000	1.007	0.213
	4NN	4NN	4NN	1.009	1.000	1.021	1.017	1.021	0.150
	5NN	4NN	None	1.003	1.005	1.005	1.000	1.009	0.174
	5NN	5NN	5NN	1.053	1.088	1.000	1.008	1.081	0.140
	6NN	None	None	1.003	1.003	1.000	1.008	1.011	0.310
	6NN	5NN	None	1.004	1.000	1.008	1.002	1.007	0.174
6NN	6NN	6NN	1.746	3.507	1.013	1.000	1.273	0.128	
Average for training set of 75				1.042	1.139	1.000	1.000	1.023	0.237
Average over all trials				1.103	1.281	1.000	1.011	1.168	0.266

**Table 19** The relative predictive powers for the **Pb-Ir** data set, **with noise** added, for the five different width-generating functions as a function of training set size and candidate cluster sets. This table may be read in the following way: In each row, the first four columns specify the combination of training set size and candidate cluster set. The next five columns are the average RMS prediction error for each of the different width-generating functions, expressed relative to the minimum RMS prediction error for the row. Green cells are the minimum values for the row, and blue cells are within 3% of the minimum. Red cells are the worst value for the row, as long as the worst value is not within 3% of the minimum. The final column gives the value, relative to the data standard deviation, of the minimum RMS prediction error for the row.



Training set size	Pair cutoff	Triple cutoff	Quad cutoff	RMS prediction error relative to row minimum					Minimum RMS prediction error / data standard deviation
				Cluster selection	Constant	Exponential	Hierarchical	Independent	
15	NN	None	None	1.002	1.000	1.000	1.002	1.002	0.874
	3NN	3NN	3NN	1.120	1.000	1.065	1.106	1.084	0.645
	4NN	None	None	1.002	1.001	1.001	1.001	1.000	0.579
	4NN	3NN	None	1.012	1.001	1.000	1.004	1.009	0.572
	4NN	4NN	4NN	1.206	1.593	1.022	1.000	1.071	0.589
	5NN	4NN	None	1.022	1.000	1.003	1.001	1.000	0.470
	5NN	5NN	5NN	1.292	1.845	1.023	1.000	1.116	0.469
	6NN	None	None	1.001	1.007	1.007	1.007	1.000	0.475
	6NN	5NN	None	1.071	1.051	1.044	1.000	1.064	0.472
6NN	6NN	6NN	1.259	1.930	1.000	1.001	1.186	0.459	
Average for training set of 15				1.076	1.198	1.003	1.000	1.034	0.568
30	NN	None	None	1.000	1.000	1.000	1.001	1.001	0.921
	3NN	3NN	3NN	1.027	1.000	1.017	1.019	1.025	0.703
	4NN	None	None	1.000	1.001	1.002	1.002	1.001	0.576
	4NN	3NN	None	1.000	1.000	1.001	1.001	1.002	0.570
	4NN	4NN	4NN	1.019	1.007	1.011	1.009	1.000	0.525
	5NN	4NN	None	1.002	1.000	1.003	1.005	1.007	0.477
	5NN	5NN	5NN	1.000	1.261	1.017	1.024	1.145	0.431
	6NN	None	None	1.000	1.004	1.003	1.005	1.003	0.486
	6NN	5NN	None	1.013	1.002	1.000	1.002	1.010	0.445
6NN	6NN	6NN	1.588	1.743	1.042	1.000	1.224	0.447	
Average for training set of 30				1.047	1.074	1.002	1.000	1.027	0.562
45	NN	None	None	1.000	1.000	1.000	1.000	1.000	0.900
	3NN	3NN	3NN	1.025	1.000	1.018	1.019	1.019	0.692
	4NN	None	None	1.000	1.000	1.000	1.001	1.000	0.569
	4NN	3NN	None	1.000	1.002	1.003	1.004	1.002	0.570
	4NN	4NN	4NN	1.036	1.000	1.062	1.069	1.005	0.491
	5NN	4NN	None	1.013	1.000	1.008	1.009	1.012	0.481
	5NN	5NN	5NN	1.000	1.011	1.056	1.060	1.007	0.395
	6NN	None	None	1.001	1.003	1.002	1.004	1.000	0.472
	6NN	5NN	None	1.001	1.001	1.003	1.007	1.000	0.434
6NN	6NN	6NN	1.227	1.504	1.000	1.053	1.102	0.456	
Average for training set of 45				1.013	1.030	1.000	1.006	1.000	0.553
60	NN	None	None	1.000	1.000	1.000	1.000	1.000	0.918
	3NN	3NN	3NN	1.014	1.000	1.004	1.006	1.014	0.702
	4NN	None	None	1.000	1.000	1.000	1.000	1.001	0.565
	4NN	3NN	None	1.001	1.002	1.002	1.003	1.000	0.564
	4NN	4NN	4NN	1.029	1.000	1.043	1.046	1.000	0.476
	5NN	4NN	None	1.005	1.000	1.002	1.005	1.002	0.467



Training set size	Pair cutoff	Triple cutoff	Quad cutoff	RMS prediction error relative to row minimum					Minimum RMS prediction error / data standard deviation
				Cluster selection	Constant	Exponential	Hierarchical	Independent	
	5NN	5NN	5NN	1.009	1.014	1.049	1.048	1.000	0.374
	6NN	None	None	1.002	1.002	1.001	1.002	1.000	0.474
	6NN	5NN	None	1.007	1.003	1.001	1.004	1.000	0.443
	6NN	6NN	6NN	1.042	1.017	1.002	1.000	1.071	0.402
Average for training set of 60				1.007	1.000	1.006	1.007	1.004	0.540
75	NN	None	None	1.000	1.000	1.000	1.000	1.000	0.927
	3NN	3NN	3NN	1.011	1.000	1.003	1.004	1.012	0.721
	4NN	None	None	1.000	1.001	1.001	1.001	1.000	0.566
	4NN	3NN	None	1.001	1.001	1.001	1.002	1.000	0.562
	4NN	4NN	4NN	1.024	1.006	1.038	1.041	1.000	0.456
	5NN	4NN	None	1.003	1.000	1.000	1.003	1.000	0.465
	5NN	5NN	5NN	1.009	1.008	1.011	1.029	1.000	0.369
	6NN	None	None	1.001	1.002	1.002	1.003	1.000	0.480
	6NN	5NN	None	1.008	1.000	1.004	1.005	1.004	0.444
6NN	6NN	6NN	1.009	1.050	1.024	1.027	1.000	0.364	
Average for training set of 75				1.004	1.003	1.005	1.007	1.000	0.537
Average over all trials				1.027	1.059	1.000	1.001	1.011	0.554

**Table 20** The relative predictive powers for the Si-Ge data set, with noise added, for the five different width-generating functions as a function of training set size and candidate cluster sets. This table may be read in the following way: In each row, the first four columns specify the combination of training set size and candidate cluster set. The next five columns are the average RMS prediction error for each of the different width-generating functions, expressed relative to the minimum RMS prediction error for the row. Green cells are the minimum values for the row, and blue cells are within 3% of the minimum. Red cells are the worst value for the row, as long as the worst value is not within 3% of the minimum. The final column gives the value, relative to the data standard deviation, of the minimum RMS prediction error for the row.



Training set size	Pair cutoff	Triple cutoff	Quad cutoff	RMS prediction error relative to row minimum					Minimum RMS prediction error / data standard deviation
				Cluster selection	Constant	Exponential	Hierarchical	Independent	
15	NN	None	None	1.004	1.000	1.000	1.001	1.001	1.080
	3NN	3NN	3NN	1.071	1.000	1.032	1.055	1.058	0.895
	4NN	None	None	1.014	1.003	1.004	1.000	1.004	0.859
	4NN	3NN	None	1.035	1.000	1.019	1.002	1.023	0.835
	4NN	4NN	4NN	1.278	1.136	1.020	1.000	1.211	0.875
	5NN	4NN	None	1.029	1.016	1.002	1.000	1.030	0.795
	5NN	5NN	5NN	1.757	1.087	1.025	1.000	1.197	0.859
	6NN	None	None	1.020	1.006	1.003	1.003	1.000	0.859
	6NN	5NN	None	1.174	1.142	1.016	1.000	1.154	0.849
6NN	6NN	6NN	1.947	1.110	1.000	1.027	1.246	0.883	
Average for training set of 15				1.220	1.040	1.003	1.000	1.082	0.887
30	NN	None	None	1.002	1.000	1.000	1.000	1.000	1.082
	3NN	3NN	3NN	1.024	1.000	1.008	1.011	1.020	0.940
	4NN	None	None	1.007	1.001	1.001	1.000	1.001	0.836
	4NN	3NN	None	1.010	1.005	1.000	1.000	1.005	0.820
	4NN	4NN	4NN	1.015	1.046	1.000	1.003	1.027	0.829
	5NN	4NN	None	1.013	1.002	1.000	1.003	1.015	0.812
	5NN	5NN	5NN	1.240	1.115	1.009	1.000	1.288	0.755
	6NN	None	None	1.016	1.000	1.001	1.001	1.005	0.798
	6NN	5NN	None	1.036	1.039	1.000	1.013	1.034	0.805
6NN	6NN	6NN	1.482	1.162	1.000	1.016	1.413	0.807	
Average for training set of 30				1.077	1.033	1.000	1.003	1.073	0.850
45	NN	None	None	1.002	1.000	1.000	1.000	1.000	1.029
	3NN	3NN	3NN	1.008	1.000	1.000	1.002	1.012	0.921
	4NN	None	None	1.004	1.000	1.000	1.001	1.001	0.825
	4NN	3NN	None	1.002	1.004	1.000	1.001	1.003	0.817
	4NN	4NN	4NN	1.006	1.009	1.000	1.003	1.007	0.822
	5NN	4NN	None	1.022	1.005	1.000	1.001	1.012	0.807
	5NN	5NN	5NN	1.055	1.048	1.000	1.006	1.053	0.750
	6NN	None	None	1.000	1.003	1.003	1.004	1.002	0.769
	6NN	5NN	None	1.015	1.022	1.000	1.002	1.022	0.775
6NN	6NN	6NN	1.298	1.159	1.000	1.003	1.305	0.791	
Average for training set of 45				1.039	1.023	1.000	1.002	1.039	0.831
60	NN	None	None	1.001	1.000	1.000	1.001	1.001	1.033
	3NN	3NN	3NN	1.005	1.000	1.000	1.001	1.006	0.928
	4NN	None	None	1.003	1.000	1.000	1.001	1.001	0.819
	4NN	3NN	None	1.000	1.001	1.000	1.001	1.003	0.806
	4NN	4NN	4NN	1.003	1.000	1.007	1.011	1.000	0.810
5NN	4NN	None	1.008	1.007	1.000	1.000	1.005	0.792	



Training set size	Pair cutoff	Triple cutoff	Quad cutoff	RMS prediction error relative to row minimum					Minimum RMS prediction error / data standard deviation
				Cluster selection	Constant	Exponential	Hierarchical	Independent	
	5NN	5NN	5NN	1.044	1.024	1.000	1.001	1.026	0.746
	6NN	None	None	1.000	1.001	1.001	1.002	1.000	0.762
	6NN	5NN	None	1.017	1.019	1.000	1.007	1.021	0.763
	6NN	6NN	6NN	1.018	1.070	1.000	1.010	1.122	0.805
Average for training set of 60				1.009	1.011	1.000	1.002	1.017	0.827
75	NN	None	None	1.001	1.000	1.000	1.001	1.001	1.020
	3NN	3NN	3NN	1.009	1.000	1.005	1.005	1.008	0.941
	4NN	None	None	1.001	1.002	1.000	1.000	1.000	0.816
	4NN	3NN	None	1.003	1.001	1.000	1.001	1.003	0.806
	4NN	4NN	4NN	1.000	1.005	1.002	1.004	1.005	0.790
	5NN	4NN	None	1.006	1.007	1.000	1.000	1.004	0.789
	5NN	5NN	5NN	1.037	1.015	1.000	1.000	1.013	0.737
	6NN	None	None	1.001	1.001	1.000	1.001	1.000	0.757
	6NN	5NN	None	1.010	1.009	1.000	1.004	1.010	0.769
6NN	6NN	6NN	1.000	1.098	1.053	1.056	1.104	0.743	
Average for training set of 75				1.001	1.007	1.000	1.001	1.008	0.822
Average over all trials				1.071	1.022	1.000	1.001	1.044	0.844

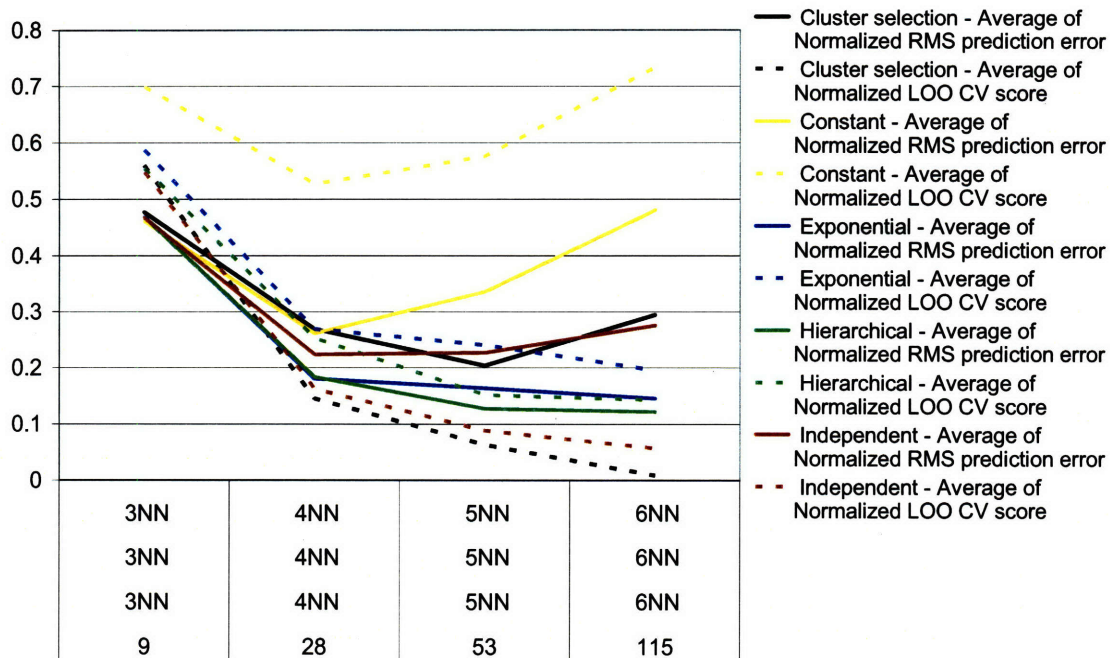
**Table 21** The relative predictive powers for the Si-Ge data set, with no noise added, for the five different width-generating functions as a function of training set size and candidate cluster sets. This table may be read in the following way: In each row, the first four columns specify the combination of training set size and candidate cluster set. The next five columns are the average RMS prediction error for each of the different width-generating functions, expressed relative to the minimum RMS prediction error for the row. Green cells are the minimum values for the row, and blue cells are within 3% of the minimum. Red cells are the worst value for the row, as long as the worst value is not within 3% of the minimum. The final column gives the value, relative to the data standard deviation, of the minimum RMS prediction error for the row.

Although the cluster selection method most frequently had the best predictive power, it was not as consistent as the exponential and hierarchical width-generating functions (Table 22). The exponential function was particularly impressive, getting within 3% of the best predictive power across a wide variety of cluster expansions 89% of the time.

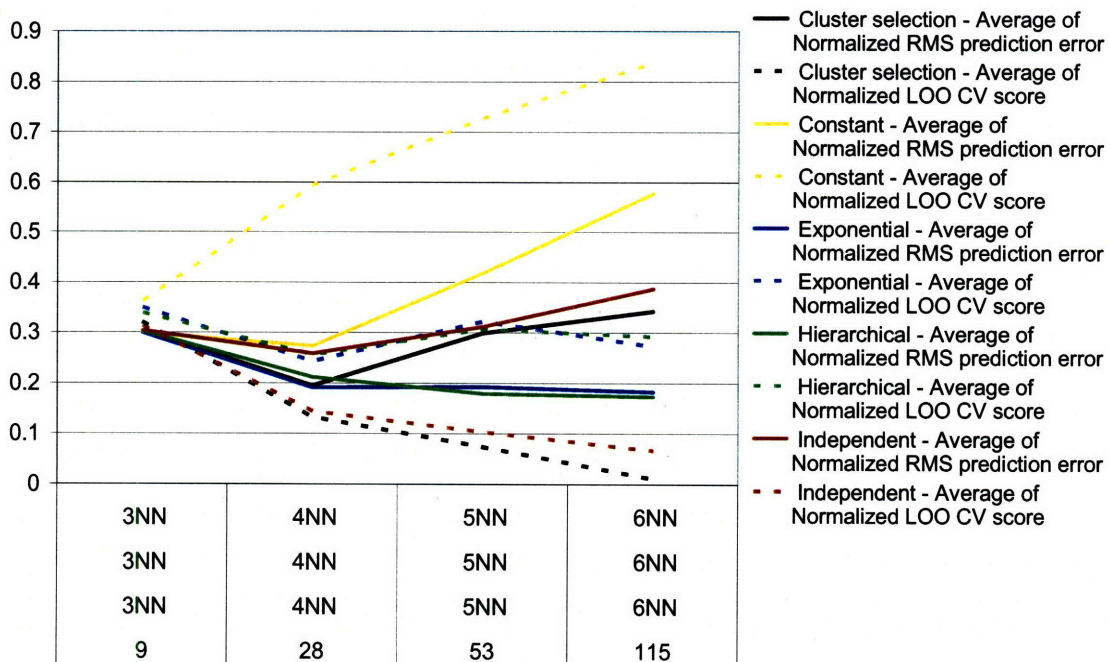
		Cluster selection	Constant	Exponential	Hierarchical	Independent
Ag-Au	No noise	68%	68%	88%	78%	68%
	Noise	64%	62%	86%	96%	70%
Pb-Ir	No noise	70%	68%	88%	76%	56%
	Noise	70%	68%	90%	80%	56%
Si-Ge	No noise	82%	84%	84%	86%	82%
	Noise	74%	76%	96%	96%	76%
Average		71.33%	71.00%	88.67%	85.33%	68.00%

**Table 22** The percentage sets of candidate clusters for which the given width-generating function produced a cluster expansion with root mean square prediction error within 3% of the best.

The above averages somewhat understate the performance of the exponential and hierarchical width-generating functions. The average performance of a cluster expansion across a variety of candidate cluster sets is not as interesting as the best performance of a cluster expansion. The average RMS prediction error for cluster expansions is plotted as a function of candidate cluster set size in Figure 48. As the number of candidate clusters is increased, the prediction error decreases for both the exponential and hierarchical width-generating functions, with the exception for a slight up tick in the Si-Ge data set. On the other hand, the RMS prediction error *increases* past a certain point for cluster expansions based on cluster selection and the constant and independent width-generating functions. For the largest candidate cluster size, which in theory should be capable of generating the best cluster expansions, the hierarchical and exponential width-generating functions are significantly better than the others in all data sets.

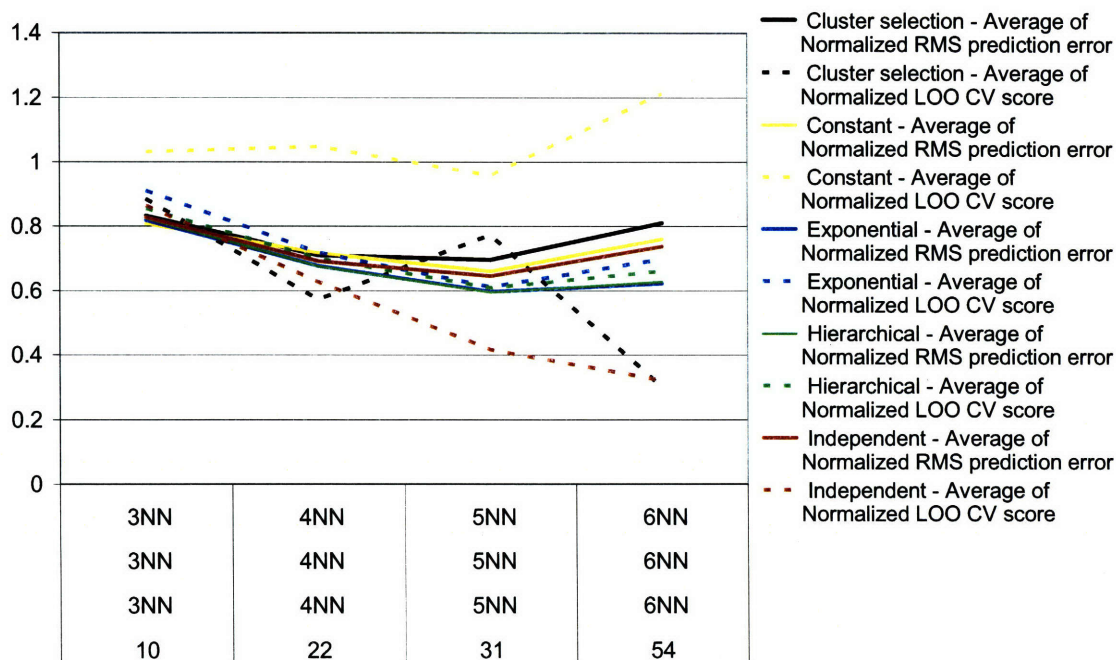


(a)



(b)





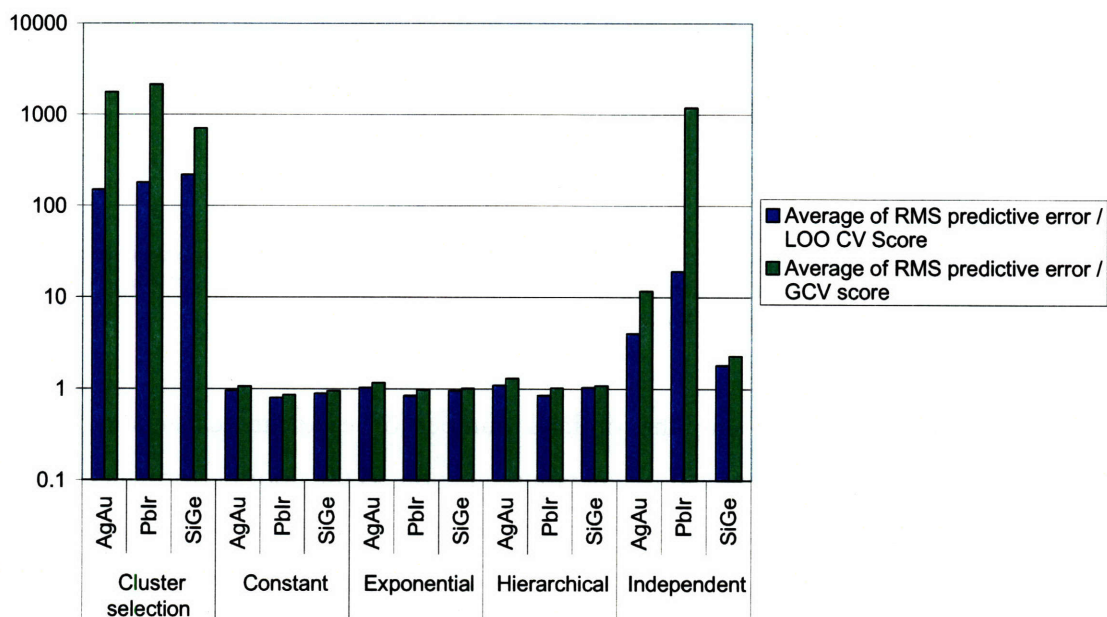
(c)

**Figure 48** A comparison of the average root mean square prediction error and the leave-one-out cross-validation score for four different sets of candidate clusters. All values are normalized relative to the standard deviation of the data. Each column on the bottom of the chart represents a set of candidate clusters and indicates, from top to bottom, the maximum distance between sites in a two-point cluster, a three point cluster, and a four point cluster, and the number of total candidate clusters. The data sets are (a) Ag-Au, (b) Pb-Ir, (c) Si-Ge.

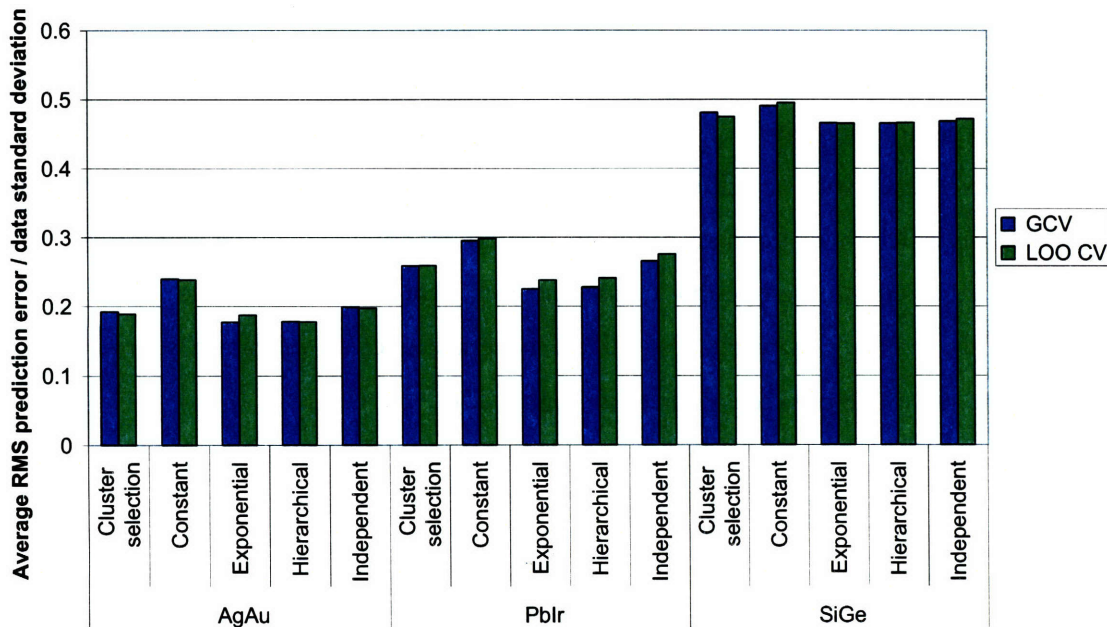
It might seem that the sensible thing to do is find the candidate cluster set with the lowest cross-validation score, and use that set. This approach is problematic for cluster selection and the independent width-generating function, because the cross-validation score decreases with increasing candidate cluster set size, even as the actual prediction error is increasing. There is no way, based on the cross-validation score, to determine the optimal set of candidate clusters. If an overly small or overly large set is used, a cluster expansion with poor predictive power will be produced. The cross-validation score is more useful for cluster expansions based on the constant width-generating function. Although the score tends to overestimate the prediction error, it is a reasonable indicator of the trend in prediction error.

The problem with the cross-validation score for cluster selection and the independent width-generating function is illustrated in Figure 49, in which the average ratio between the

root mean square prediction error and the cross-validation score is plotted. For the three width-generating functions with five or fewer independent parameters, the cross-validation score is a reliable indicator of the prediction error. For the remaining two, the cross-validation score tends to significantly underestimate the actual prediction error, especially for situations in which there are many candidate clusters. It should be noted that this problem is not isolated to the leave-one-out cross-validation score; the generalized cross-validation score shows similar behavior. Overall the generalized cross-validation score performed similarly to leave-one-out cross-validation (Figure 50).



**Figure 49** The average ratio between the root mean square prediction error and the leave-one-out cross-validation (LOO CV) and generalized cross-validation (GCV) scores. The scale is logarithmic.



**Figure 50** The average root mean square prediction error for cluster expansions in which the parameters are chosen to minimize the leave-one-out cross-validation (LOO CV) or generalized cross-validation (GCV) scores.

The problem of choosing the optimal number of candidate clusters is not very important for the independent width-generating function, as the poor performance and large computational expense of this function make it unlikely that it will ever be used. However cluster selection is a widely used method, and it performs relatively well under the constraint that all subclusters of every selected cluster must also be selected. Although no hard rule exists for choosing the correct number of candidate clusters, a heuristic can be developed by studying the behavior of the optimal number of candidate clusters as a function of training set size. If the number of candidate clusters is large, but no larger than the number of training structures, reasonably good results can be expected (Table 23). In general, this heuristic tends to produce candidate sets that are a little too large.



Data set	Training set size	Best number of candidate clusters	Heuristic number of candidate clusters	Average RMS prediction error for largest set relative to best set.	Average RMS prediction error for heuristic set relative to best set.
Ag-Au	15	7	13	1.744	1.106
	30	19	28	3.028	1.651
	45	19	28	2.734	1.470
	60	19	53	2.693	1.214
	75	19	53	2.351	1.109
Pb-Ir	15	13	13	2.383	1.000
	30	13	28	2.358	1.111
	45	28	28	1.988	1.000
	60	28	53	1.739	1.074
	75	53	53	1.569	1.000
Si-Ge	15	12	12	1.758	1.000
	30	22	15	1.476	1.071
	45	31	31	1.333	1.000
	60	54	31	1.066	1.066
	75	54	54	1.000	1.000

**Table 23** A comparison of different methods for selecting the best candidate structure set for the cluster selection method. The rightmost column represents the average RMS prediction error for candidate sets selected by using the heuristic that the largest set smaller than the training set was chosen. The errors are expressed relative to the error for the best candidate set.

#### III.4.4.6. Discussion

The cross-validation scores are reasonably successful in estimating the predictive power of cluster expansions generated using the exponential, hierarchical, and constant width-generating functions. These functions also have the fewest number of parameters to be determined. On the other hand, for cluster selection and the independent width-generating function, which have as many parameters as there are candidate cluster orbits, the cross-validation scores become poor predictors of the predictive power of the cluster expansion as the number of parameters increases. This leads to a situation in which improving the cross-validation score does not necessarily improve the predictive power of the cluster expansion, and cluster expansions with poor predictive power are selected. These data suggest that the correlation between the predictive power and the cross-validation score is best maintained when there are few parameters in the width-generating function. This is consistent with the



idea of over fitting the cross-validation score: the more degrees of freedom are allowed, the more likely over fitting becomes. For this reason, it is suggested that width-generating functions with few parameters should be preferred.

The constant width-generating function has the fewest parameters, but it performs poorly for the cluster expansions evaluated. The reason it performs poorly is not likely due to a problem with the cross-validation scores; the scores predict its poor performance. The constant width-generating function represents a prior distribution that is inconsistent with the physical expectation for these systems. The physical expectation is that as clusters increase in size, their ECI are expected to approach zero. This expectation is captured in both the hierarchical and exponential width-generating functions, which perform significantly better than the constant width-generating function. The expectation is also captured, in a different way, in the cluster selection method as long as the constraint is imposed that every subcluster of a selected cluster must also be selected. If this constraint is turned off, the RMS prediction error for the cluster-selection method increases on average by 45%, 44%, and 5% for the Ag-Au, Pb-Ir, and Si-Ge data sets respectively.

Based on the tests in this section, the hierarchical and exponential width-generating functions have the following advantages over traditional cluster selection methods:

- They more frequently produce cluster expansions with very good predictive power.
- The cross-validation scores for these methods are more reliable indicators of the true predictive power.
- They tend to get better (or at least not get much worse) as the number of candidate clusters is increased.
- They produce more consistently good cluster expansions from small training sets.
- Model selection by searching for a low cross-validation score involves a search over an approximately five-dimensional continuous space, as opposed to a search over a binary space of as many dimensions as there are candidate clusters.

Although the exponential and hierarchical width-generating functions have been successful for the data sets tested, other width-generating functions may be more successful for other data sets. The better the prior distribution represents the true values, the better the cluster expansion should converge. The results in this thesis suggest a recipe for a generally good cluster expansion:

1. Use the Gaussian prior, or iterated Gaussian prior if the number of non-zero ECI is a concern.
2. Select a width-generating function that penalizes large clusters and has few parameters. Functions similar to the exponential or hierarchical functions might work well. A possible advantage to the hierarchical width-generating function is that in the tests in this thesis, it converged much more quickly to a local minimum than the exponential function.
3. Use the generalized cross-validation score instead of the leave-one-out cross-validation score for improved speed.
4. Select a reasonably large number of candidate clusters.
5. Weighting the training structures may or may not help.
6. Decide whether to apply the prior to correlation coefficients or ECI.
7. Select a reasonable initial guess for the parameters for your width-generating function. You may want to try several different guesses to see how the cross-validation scores change.
8. Use a local minimization algorithm to find a set of parameters that produces a low cross-validation score.

The combination of the Gaussian prior, the hierarchical width-generating function, and the generalized cross-validation score may be the fastest way to generate cluster expansions with consistently good predictive power.

One of the advantages of using the Gaussian prior and a simple width-generating function is that the mathematics of fitting the cluster expansion are relatively straightforward to implement. Nonetheless, considering the historical popularity of cluster selection, it can be expected that this method will continue to be used. If it is used, it is strongly

recommended that the clusters are selected with a constraint that forces all subclusters of selected clusters to also be selected, and that the set of candidate clusters is smaller than the number of training structures. It is recommended that the generalized cross-validation score is used instead of the leave-one-out cross-validation score, as the generalized cross-validation score is faster and produces results that are approximately the same on average.

---

## III.5. Further Considerations

---

This chapter has provided a basic foundation for the application of Bayesian learning to the problem of potential fitting, with a specific focus on cluster expansions. With the fundamentals of this method in place, it is worth exploring some other potential benefits of the Bayesian approach.

### III.5.1. The posterior distribution

---

The key to the Bayesian approach for building a cluster expansion is that the fit cluster expansion is chosen to be the maximum of the posterior distribution. Other properties of the posterior distribution, such as its shape and width, have been for the most part ignored in this chapter. (The discussions of the variance of the ECI and prediction errors touched upon these properties indirectly.) There is potentially valuable information in these properties. For example, the posterior distribution can be used to calculate confidence intervals for predictions made by the cluster expansion. Indirectly, it can be used to calculate confidence intervals on material properties as well.

It is often the case that the maximum of the posterior distribution is also its mean. For this reason, there may be little additional value in calculating the “average” cluster expansion. However, many material properties are non-linear functions of cluster-expanded values. For example, imagine there is a material property  $A$  that is the function of the output vector  $\bar{y}_{test}$  over the entire test domain.

$$A = A(\bar{y}_{test}) = A(V(X_{test})) \quad \text{Equation 178}$$

The average value for  $A$  can be written as an average over the posterior distribution of  $\bar{y}_{test}$ :

$$\langle A \rangle = \int P(V(\cdot)) A(V(X_{test})) dV(\cdot) \quad \text{Equation 179}$$

where the integral is over all possible functions  $V(\cdot)$  and  $P(V(\cdot))$  is the posterior distribution over functions. In the cluster expansion, a given function  $V(\cdot)$  is represented by the vector  $\bar{V}$ , and Equation 179 becomes

$$\langle A \rangle = \int P(\vec{v}) A(X_{test}, \vec{v}) d\vec{v} \quad \text{Equation 180}$$

Unless  $A(\vec{y}_{test})$  is a linear function (which it generally is not), it is worth noting that in general

$$\langle A \rangle \neq A\left(\int [P(\vec{v}) X_{test}, \vec{v}] d\vec{v}\right) \quad \text{Equation 181}$$

For the calculation of the average, or even most likely, values of non-linear properties it may be best to calculate the distribution in the values of the property from the posterior distribution of cluster expansions.

### III.5.2. Continuous potential energy surfaces

---

The focus of this chapter has been on the discrete cluster expansion potential model because it is both useful and relatively simple to analyze. However much of the theory developed in this chapter also applies to potential energy functions defined over a continuous domain of position variables. The mean for the prior distribution for such functions might be empirically derived potential functions, such as the embedded atom potential. This would reduce the problem to that of learning the difference between the actual potential energy surface and the empirical potential function. The norm of this difference is a natural regularization term. Kernel-based methods could then be employed to find the maximum of the posterior distribution. With this general approach, the Bayesian method may be of similar benefit to the development of continuous potential functions as it is to the cluster expansion. Further development and evaluation of this approach is beyond the scope of this thesis.

### III.5.3. Applying the Bayesian cluster expansion to surfaces

---

The cluster expansion tends to work best when applied to systems with high symmetry, such as crystals with infinite periodicity in all directions.<sup>7</sup> Because of symmetry only the ECI for orbits of symmetrically equivalent cluster functions, rather than the ECI for every cluster function individually, need to be determined. There are many situations, however, in which

---

<sup>7</sup> Of course, no crystal is truly infinite. However, when determining the properties of a bulk material, the surface is usually assumed to be so far away relative to the distance between atoms that it has a negligible effect on the local physics. In this situation, infinite periodicity is a safe and usually highly accurate approximation.

symmetry is broken, significantly increasing the number of ECI that need to be determined. For example, the region of interest for a particular material might be the material surface. At the surface, it is typically only safe to assume infinite periodicity in the directions parallel to the surface. This greatly complicates the cluster expansion, as atoms on the surface need to be treated differently from atoms one layer from the surface, which are different from atoms two layers from the surface, etc. A similar problem arises when studying nanowires and nanoparticles.

A common approximation in these situations is as follows: beyond a certain distance from the surface, atoms are approximately symmetrically equivalent. For example, it might be safe to approximate that an atom  $n$  layers from the surface is equivalent to an atom  $n + 1$  layers from the surface if  $n = 10$ . The smaller  $n$ , the fewer unique cluster orbits there are, and the easier it is to fit the cluster expansion. There is a trade-off between making  $n$  small to reduce the computational cost, and making  $n$  large to increase the accuracy of the cluster expansion.

In this problem there is a prior expectation that the further an atom is from the surface, the more it will resemble an atom in the bulk. Phrased this way, surfaces seem to be natural candidates for a Bayesian treatment. The key, as always, is how to express the prior expectation in a way that is both realistic and allows for efficient estimation of the ECI.

To construct a prior for a surface, it is necessary to find a way to directly translate physical intuition into a probability distribution. The expected absolute difference between the ECI for a given cluster function and the ECI for the same cluster function in the bulk limit should decrease as a function of the distance from the surface. If a Gaussian distribution is assumed, the following expression can be written:

$$\tilde{V}_{\bar{j}} \sim \mathbf{N}(V_{\bar{j},Bulk}, \sigma^2(r_{\bar{j}})) \quad \text{Equation 182}$$

where  $\tilde{V}_{\bar{j}}$  is the ECI for some cluster labeled  $\bar{j}$ ,  $V_{\bar{j},Bulk}$  is the ECI for an equivalent cluster in the bulk limit, and  $r_{\bar{j}}$  is some measure of the distance between cluster  $\bar{j}$  and the surface.  $\mathbf{N}(V_{\bar{j},Bulk}, \sigma^2(r_{\bar{j}}))$  is a normal distribution with mean  $V_{\bar{j},Bulk}$  and variance,  $\sigma^2$ , which is a function of  $r_{\bar{j}}$ . It would be expected that  $\sigma^2(r_{\bar{j}}) \rightarrow 0$  as  $r_{\bar{j}} \rightarrow \infty$ .

Now consider another cluster function that is symmetrically distinct from cluster function  $\bar{j}$  due to the presence of the surface, but is symmetrically equivalent in the bulk limit. Label this cluster function  $\bar{i}$ . The counterpart to Equation 182 is

$$\tilde{V}_{\bar{i}} \sim \mathbf{N}(V_{\bar{i},Bulk}, \sigma^2(r_{\bar{i}})) \quad \text{Equation 183}$$

where  $V_{\bar{i},Bulk} = V_{\bar{j},Bulk}$ . Combining Equation 182 with Equation 183, and using the known properties of the normal distribution, results in:

$$\tilde{V}_{\bar{i}} - \tilde{V}_{\bar{j}} \sim \mathbf{N}(\mathbf{0}, \sigma^2(r_{\bar{i}}) + \sigma^2(r_{\bar{j}})) \quad \text{Equation 184}$$

Equation 184 says that for any two cluster functions that are equivalent in the bulk, the difference between their ECI should get smaller (approaching zero) the farther they are from the surface. An advantage to Equation 184 is that  $V_{\bar{i},Bulk}$  has disappeared, leaving  $\sigma^2(r)$  as the only unknown. Equation 184 should begin to look familiar. If it is added to the Gaussian prior distribution for the ECI (Equation 70), the new prior becomes:

$$P(V_{\Delta}(\cdot) = V_{opt}(\cdot) - V_p(\cdot)) \propto \prod_j e^{\frac{-V_{j,\Delta}^2}{2\sigma_j^2}} \prod_{k \in J_{bulk}} e^{\frac{-(V_{j,\Delta} - V_{k,\Delta})^2}{4(\sigma^2(r_j) + \sigma^2(r_k))}} \quad \text{Equation 185}$$

where  $J_{bulk}$  is the set of all cluster functions that would be symmetrically equivalent to  $\bar{j}$  in the bulk.

The minimization problem in Equation 71 becomes

$$\hat{V}_{Bayes}(\cdot) = V_p(\cdot) + \arg \min_{V_{\Delta}(\cdot)} \left( \begin{aligned} & \frac{1}{2} \sum_i \frac{1}{\sigma_i^2} (V_{\Delta}(\bar{c}_i) - y_{\Delta,i})^2 \\ & + \frac{1}{2} \sum_{\bar{j}} \frac{V_{\bar{j},\Delta}^2}{\sigma_{\bar{j}}^2} + \frac{1}{4} \sum_{\bar{j}} \sum_{k \in J_{bulk}} \frac{(V_{\bar{j},\Delta} - V_{k,\Delta})^2}{\sigma^2(r_{\bar{j}}) + \sigma^2(r_{\bar{k}})} \end{aligned} \right) \quad \text{Equation 186}$$

Equation 186 can be expanded to get



$$\hat{V}_{Bayes}(\cdot) = V_p(\cdot) + \arg \min_{V_\Delta(\cdot)} \left( \begin{aligned} & \frac{1}{2} \sum_i \frac{1}{\sigma_i^2} (V_\Delta(\bar{c}_i) - y_{\Delta,i})^2 \\ & + \frac{1}{2} \sum_{\bar{j}} V_{\bar{j},\Delta}^2 \left( \frac{1}{\sigma_{\bar{j}}^2} + \sum_{\bar{k} \in j_{bulk}} \frac{1}{\sigma^2(r_{\bar{j}}) + \sigma^2(r_{\bar{k}})} \right) \\ & - \frac{1}{2} \sum_{\bar{j}} \sum_{\bar{k} \in j_{bulk}} V_{\bar{j},\Delta} V_{\bar{k},\Delta} \left( \frac{1}{\sigma^2(r_{\bar{j}}) + \sigma^2(r_{\bar{k}})} \right) \end{aligned} \right) \quad \text{Equation 187}$$

Equation 187 can be written concisely in matrix-vector notation:

$$\hat{V}_{Bayes}(\bar{x}) = V_p(\bar{x}) + \bar{x} \cdot \arg \min_{\bar{\beta}_\Delta} \left( \frac{1}{2} (\bar{y}_\Delta - X\bar{\beta}_\Delta)^T W (\bar{y}_\Delta - X\bar{\beta}_\Delta) + \bar{\beta}_\Delta^T \Lambda \bar{\beta}_\Delta \right) \quad \text{Equation 188}$$

where  $\bar{y}_\Delta$ ,  $X$ ,  $W$ , and  $\bar{\beta}_\Delta$  are defined as earlier. The only thing that is different between Equation 188 and Equation 74 is the definition of the matrix  $\Lambda$ . Previously, only considered the case in which  $\Lambda$  is a diagonal matrix was considered. Now  $\Lambda$  is defined by:

$$\begin{aligned} \Lambda_{\alpha\alpha} &= \frac{1}{m_\alpha^2} \left( \frac{1}{\sigma_\alpha^2} + \sum_{\gamma \neq \alpha} \frac{1}{\sigma^2(r_\alpha) + \sigma^2(r_\gamma)} \right) \\ \Lambda_{\alpha\gamma} &= \Lambda_{\gamma\alpha} = \frac{-1}{m_\alpha m_\gamma} \left( \frac{1}{\sigma^2(r_\alpha) + \sigma^2(r_\gamma)} \right) \text{ for} \\ & \alpha \neq \gamma, \alpha_{bulk} = \gamma_{bulk} \end{aligned} \quad \text{Equation 189}$$

where  $\alpha_{bulk}$  and  $\gamma_{bulk}$  are bulk-material equivalents to orbits  $\alpha$  and orbit  $\gamma$ , and  $m_\alpha$  and  $m_\gamma$  are the multiplicities of cluster function orbits  $\alpha$  and  $\gamma$  respectively. The multiplicities are necessary because Equation 189 is written in terms of correlation coefficients  $\beta_\alpha$ , where  $\beta_\alpha = m_\alpha V_\alpha$ . The diagonal variance terms  $\sigma_\alpha^2$  can be parameterized in one of the ways discussed previously. All that remains is to develop a parameterized expression for  $\sigma^2(r_\alpha)$ .

The first step in developing an expression for  $\sigma^2(r_\alpha)$  is to define  $r_\alpha$ . This is a measure of how far cluster functions in orbit  $\alpha$  are from the surface. Using the intuition that having even one site on the surface can significantly affect the ECI for a cluster function, let  $r_\alpha$  be defined as the distance between the surface and the site in the cluster orbit that is nearest to the surface. Therefore, for any clusters orbits containing sites on the surface,  $r_\alpha = 0$ .

It is helpful to list some desired properties of the function  $\sigma^2(r_\alpha)$ . As  $r_\alpha \rightarrow \infty$ , it is expected that  $\sigma^2(r_\alpha) \rightarrow 0$ , due to the clusters becoming nearly indistinguishable from bulk clusters. The function should be monotonically decreasing, and it should be capable of having a large, distinct peak at  $r_\alpha = 0$ , where it is expected that the ECI may differ substantially from bulk ECI. As usual, it is desirable to keep the number of parameters to a minimum and retain a form that is invariant to length and energy scales. A simple form that fits this description is

$$\sigma^2(r_\alpha) = \lambda_1 e^{-\lambda_2 r_\alpha} \quad \text{Equation 190}$$

where there are only two unknown positive parameters,  $\lambda_1$  and  $\lambda_2$ .

From this point on, the problem is identical to problems that have been treated previously in this chapter. The solution to Equation 188 is still Equation 81, and determining the parameters for the prior distribution is still done using a model-selection method such as minimizing a cross-validation score.

This is a mathematically simple, but powerful technique for fitting cluster expansions for surfaces, nanoparticles, or nanowires. It is not hard to imagine developing a similar method for cluster expansions near crystal defects. The problem of having to determine the ECI for a large number of cluster orbits is mitigated by placing realistic soft constraints on the ECI that reduce the effective number of degrees of freedom in the model. These constraints are more flexible than methods that impose artificial symmetry on the problem, as they allow ECI for slightly different cluster functions to be slightly different, but discourage them from being very different.

### III.5.3.1. Example: Si-Ge nanowire

As a simple test of the above approach, it has been applied to the calculation of the energy for a Si-Ge (silicon-germanium) nanowire. For bulk Si-Ge, including all interactions up to second-nearest-neighbor pairs in the cluster expansion requires four cluster orbits. For a two nanometer wide nanowire in the  $\langle 111 \rangle$  direction, including all interactions up to second-nearest-neighbor pairs requires eighty-six cluster orbits. This material, with dilute germanium composition (5-18%), is of interest in the field of thermoelectrics. Here it is used as a test system for the above method.

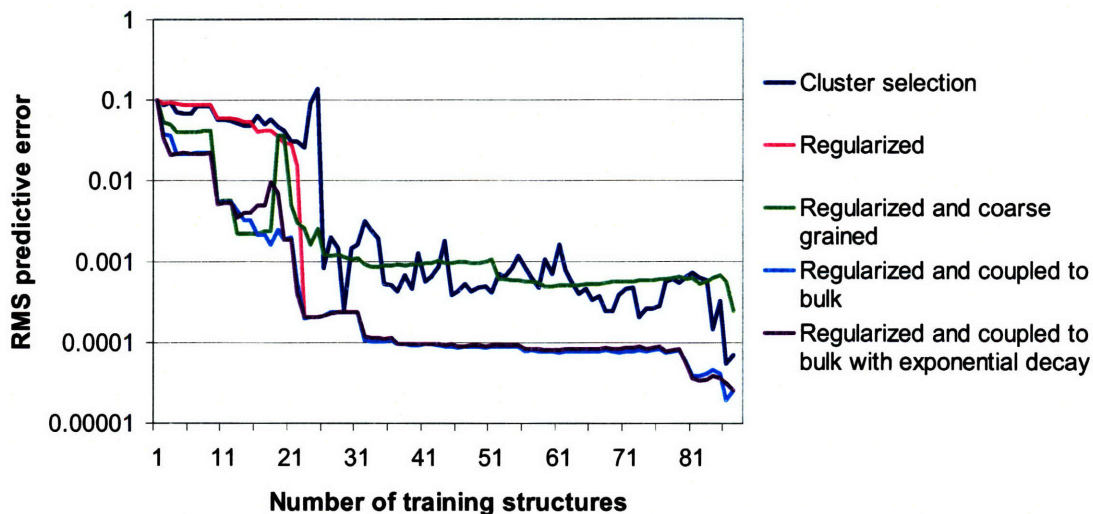
The sample data was generated as follows: using the Tersoff potential, a set of energies were calculated for eighty-eight nanowire structures.<sup>8</sup> The first eighteen structures in the set are high-symmetry structures that were generated manually, and the remainder of the set was generated automatically to minimize the expected prediction error over the composition range of interest.

To evaluate cluster expansion methodologies on this data, one structure at a time was drawn from the sample data and added to the training set. Cluster expansions were then trained on this set using each of the methodologies of interest, and for each methodology the root-mean-square (RMS) prediction error was calculated for the structures left out of the training set. The methodologies tested are as follows:

- 1) **Cluster selection:** A Monte Carlo algorithm is used to find a subset of clusters that gives a low leave-one-out cross-validation score. The ECI for the clusters in this subset are determined by using a least-squares fit, and the remaining ECI are set to zero.
- 2) **Regularized:** The ECI are determined using regularization with the exponential prior. The parameters of the prior were determined by using gradient descent to minimize the leave-one-out cross-validation score.
- 3) **Regularized and coarse grained:** The ECI were determined in the same way as for the regularized fit, but the ECI for all cluster functions that would share a common bulk orbit ( $j_{Bulk}$ ) were constrained to be the same as long as the clusters had no surface sites.
- 4) **Regularized and coupled to the bulk:** The ECI were determined by solving Equation 188, with  $\sigma_\alpha^2$  given by the exponential prior and  $\sigma^2(r_\alpha) = \lambda_1$  for some parameter  $\lambda_1$ . The parameters for the priors were determined by using gradient descent to minimize the leave-one-out cross-validation score.
- 5) **Regularized and coupled to the bulk with exponential decay:** This is the same as the previous method, but with  $\sigma^2(r_\alpha) = \lambda_1 e^{-\lambda_2 r_\alpha}$ .

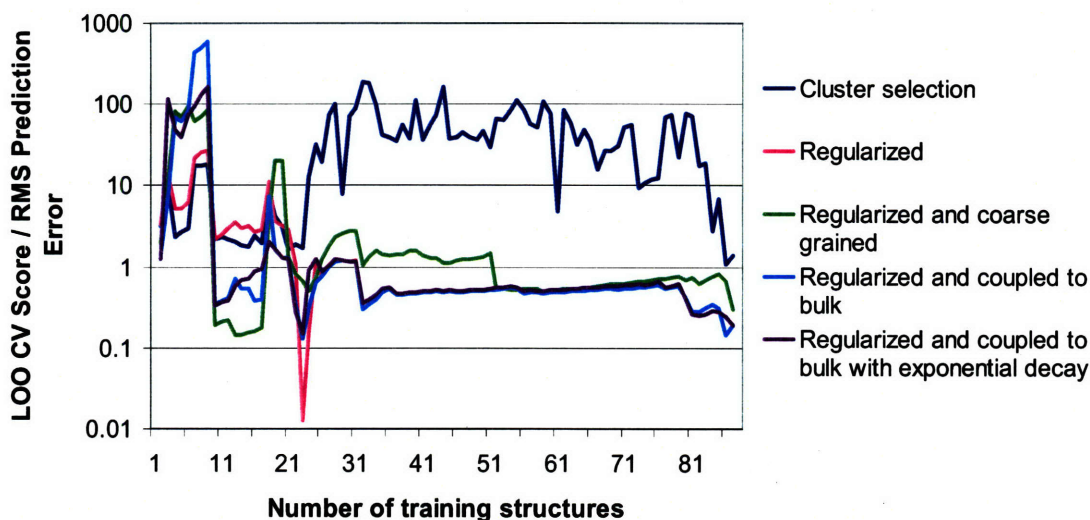
---

<sup>8</sup> Thanks to John Reed at Sandia National Laboratories for calculating these energies.



**Figure 51** The root-mean-square prediction error for cluster expansions fit using different methods, plotted logarithmically as a function of the size of the training set.

The results for these calculations are given in Figure 51. It can be seen that cluster selection does relatively poorly, with a few situations in which the expansion had a low leave-out CV score but terrible predictive power. This is despite the fact that this methodology was more time-consuming than any of the others. It could be argued that the Metropolis algorithm used did not find the optimal set of clusters, but in general this is a risk of the cluster selection technique when there is a large set of candidate clusters. It could also be argued that the number of candidate clusters was too large, given the number of structures included in the training set. It is possible that if the set of candidate structures were reduced, this method would produce better results. Unfortunately it is not always clear which clusters should be included in the set of candidate clusters. It should be noted that despite the poor predictive power of the cluster selection technique, it generally had the highest computational cost and produced the lowest leave-one-out cross-validation scores (Figure 52).



**Figure 52** The ratio between the actual root-mean-square prediction error and the leave-one-out cross-validation score for the different methods evaluated. A value of 1 would be ideal.

Regularization with the exponential prior does much better. As expected, the predictive power of the regularized fit converges much more smoothly than cluster selection as structures are added to the training set. Still, regularization struggles with a small number of structures.

Regularization with coarse graining does much better than straight regularization for a small number of structures, but it appears to hit a wall and actually ends up doing worse than cluster selection for a large number of structures. This is likely because of the constraint that all non-surface clusters of the same type must have the same ECI. This constraint is overly restrictive, and there is no value for the common ECI that can make up for the fact that the “true” ECI should be slightly different from each other.

The best methods are the two in which there is regularization with a coupling between clusters as described above. Both methods return similar results, regardless of the form used for  $\sigma^2(r_\alpha)$ . This is somewhat surprising, but it may be due to the fact that for this particular system, in which there are not many layers, there ends up being little practical difference between the methods. These methods tend to capture the best of both worlds: for a small number of structures, they perform about as well as coarse-graining, beating the other two methods by nearly an order of magnitude. For a large number of structures they

perform as well as regularization, beating cluster selection and coarse graining by nearly an order of magnitude.

The best techniques have difficulty improving significantly after about forty structures are added to the training set. This could be due to inherent noise in the training data, or due to multi-body interactions that cannot be captured by a cluster expansion that only goes up to next-nearest-neighbor pairs.

It is interesting to observe that all techniques appear to get significantly better after about 20 structures are added to the training set. This is the point at which the machine-generated structures start to be included. A likely explanation is that the manually created structures insufficiently sampled some regions of phase space, and the variance-reduction algorithm recognized this and filled these regions in with the structures near structure #20. As these regions filled in, the predictive power of all methodologies significantly improved.

It is encouraging that the results are generally consistent with what would be expected. The Bayesian approach should enable faster and more accurate generation of cluster expansions for surfaces, nanowires, and nanoparticles. A similar approach may also work well for certain types of crystal defects.

## IV. References





1. Carrette, L., K.A. Friedrich, and U. Stimming, *Fuel cells: Principles, types, fuels, and applications*. Chemphyschem, 2000. **1**(4): p. 162-193.
2. *Fuel Cells* (<http://www.eere.energy.gov/hydrogenandfuelcells/fuelcells/>). 2004, United States Department of Energy.
3. Song, C.S., *Fuel processing for low-temperature and high-temperature fuel cells - Challenges, and opportunities for sustainable development in the 21st century*. Catalysis Today, 2002. **77**(1-2): p. 17-49.
4. Acres, G.J.K., *Recent advances in fuel cell technology and its applications*. Journal of Power Sources, 2001. **100**(1-2): p. 60-66.
5. *Hydrogen, Fuel Cells, & Infrastructure Technology Program Multi-Year Research, Development and Demonstration Plan*. 2005, United States Department of Energy.
6. Satyapal, S., et al., *The US Department of Energy's National Hydrogen Storage Project: Progress towards meeting hydrogen-powered vehicle requirements*. Catalysis Today, 2007. **120**(3-4): p. 246-256.
7. Schlapbach L, Z.A., *Hydrogen-storage materials for mobile applications*. Nature, 2001. **414**: p. 353-358.
8. Zhou, L., *Progress and problems in hydrogen storage methods*. Renewable & Sustainable Energy Reviews, 2005. **9**(4): p. 395-408.
9. Zuttel, A., *Hydrogen storage methods*. Naturwissenschaften, 2004. **91**(4): p. 157-172.
10. Seayad, A.M. and D.M. Antonelli, *Recent advances in hydrogen storage in metal-containing inorganic nanostructures and related materials*. Advanced Materials, 2004. **16**(9-10): p. 765-777.
11. Langmi, H.W., et al., *Hydrogen adsorption in zeolites A, X, Y and RHO*. Journal of Alloys and Compounds, 2003. **356**: p. 710-715.
12. Tang, C.C., et al., *Catalyzed collapse and enhanced hydrogen storage of BN nanotubes*. Journal of the American Chemical Society, 2002. **124**(49): p. 14550-14551.
13. Rowsell, J.L.C., et al., *Hydrogen sorption in functionalized metal-organic frameworks*. Journal of the American Chemical Society, 2004. **126**(18): p. 5666-5667.
14. Panella, B. and M. Hirscher, *Hydrogen physisorption in metal-organic porous crystals*. Advanced Materials, 2005. **17**(5): p. 538-+.
15. Lin, X., et al., *High H-2 adsorption by coordination-framework materials*. Angewandte Chemie-International Edition, 2006. **45**(44): p. 7358-7364.

16. Grochala, W. and P.P. Edwards, *Thermal decomposition of the non-interstitial hydrides for the storage and production of hydrogen*. Chemical Reviews, 2004. **104**(3): p. 1283-1315.
17. Schuth, F., B. Bogdanovic, and M. Felderhoff, *Light metal hydrides and complex hydrides for hydrogen storage*. Chemical Communications, 2004(20): p. 2249-2258.
18. Alapati, S.V., J.K. Johnson, and D.S. Sholl, *Predicting reaction equilibria for destabilized metal hydride decomposition reactions for reversible hydrogen storage*. Journal of Physical Chemistry C, 2007. **111**(4): p. 1584-1591.
19. Alapati, S.V., J.K. Johnson, and D.S. Sholl, *Using first principles calculations to identify new destabilized metal hydride reactions for reversible hydrogen storage*. Physical Chemistry Chemical Physics, 2007. **9**(12): p. 1438-1452.
20. Chen, P., et al., *Interaction of hydrogen with metal nitrides and imides*. Nature, 2002. **420**(6913): p. 302-304.
21. Mueller, T. and G. Ceder, *A density functional theory study of hydrogen adsorption in MOF-5*. Journal of Physical Chemistry B, 2005. **109**(38): p. 17974-17983.
22. Mueller, T. and G. Ceder, *Effective interactions between the N-H bond orientations in lithium imide and a proposed ground-state structure*. Physical Review B, 2006. **74**(13).
23. Nijkamp, M.G., et al., *Hydrogen Storage using physisorption - materials demand*. Appl. Phys. A, 2001. **72**: p. 619-623.
24. Becher, M., et al., *Hydrogen storage in carbon nanotubes*. Comptes Rendus Physique, 2003. **4**(9): p. 1055-1062.
25. James, S.L., *Metal-organic frameworks*. Chemical Society Reviews, 2003. **32**(5): p. 276-288.
26. Rosi, N.L., et al., *Hydrogen Storage in Microporous Metal-Organic Frameworks*. Science, 2003. **300**: p. 1127.
27. Zhao, X.B., et al., *Hysteretic adsorption and desorption of hydrogen by nanoporous metal-organic frameworks*. Science, 2004. **306**(5698): p. 1012-1015.
28. Pan, L., et al., *Microporous metal organic materials: Promising candidates as sorbents for hydrogen storage*. Journal of the American Chemical Society, 2004. **126**(5): p. 1308-1309.
29. Li, H., et al., *Design and synthesis of an exceptionally stable and highly porous metal-organic framework*. Nature, 1999. **402**(6759): p. 276-279.
30. Eddaoudi, M., et al., *Modular chemistry: Secondary building units as a basis for the design of highly porous and robust metal-organic carboxylate frameworks*. Accounts of Chemical Research, 2001. **34**(4): p. 319-330.

31. Hohenberg, P. and W. Kohn, *Phys. Rev.*, 1964. **136**: p. 864.
32. Zhang, Y.K., W. Pan, and W.T. Yang, *Describing van der Waals Interaction in diatomic molecules with generalized gradient approximations: The role of the exchange functional*. *Journal of Chemical Physics*, 1997. **107**(19): p. 7921-7925.
33. Perezjorda, J.M. and A.D. Becke, *A Density-Functional Study of Van-Der-Waals Forces - Rare-Gas Diatomics*. *Chemical Physics Letters*, 1995. **233**(1-2): p. 134-137.
34. Wesolowski, T.A., et al., *Comparative study of benzene center dot center dot center dot X (X = O-2, N-2, CO) complexes using density functional theory: The importance of an accurate exchange-correlation energy density at high reduced density gradients*. *Journal of Physical Chemistry A*, 1997. **101**(42): p. 7818-7825.
35. Patton, D.C. and M.R. Pederson, *Application of the generalized-gradient approximation to rare-gas dimers*. *Physical Review A*, 1997. **56**(4): p. R2495-R2498.
36. Tran, F., et al., *Physisorption of molecular hydrogen on polycyclic aromatic hydrocarbons: A theoretical study*. *Journal of Physical Chemistry B*, 2002. **106**(34): p. 8689-8696.
37. Wu, X., et al., *Towards extending the applicability of density functional theory to weakly bound systems*. *Journal of Chemical Physics*, 2001. **115**(19): p. 8748-8757.
38. Perdew, J.P. and W. Yue, *Accurate and simple analytic representation of the electron-gas energy*. *Physical Review B (Condensed Matter)*, 1992. **45**(23): p. 13244-9.
39. Perdew, J.P., K. Burke, and M. Ernzerhof, *Generalized gradient approximation made simple*. *Physical Review Letters*, 1996. **77**(18): p. 3865-3868.
40. Blochl, P.E., *Projector Augmented-Wave Method*. *Physical Review B*, 1994. **50**(24): p. 17953-17979.
41. Kresse, G. and D. Joubert, *From ultrasoft pseudopotentials to the projector augmented-wave method*. *Physical Review B*, 1999. **59**(3): p. 1758-1775.
42. Rowsell, J.L.C. and O.M. Yaghi, *Description of the Structure of MOF-5 (Private Communication)*. 2003.
43. Huber, K.P. and G. Herzberg, *Constants of Diatomic Molecules*. *Molecular Spectra and Molecular Structure*, ed. G. Herzberg. 1979, New York City: Van Nostrand Reinhold Company.
44. Sagara, T., J. Klassen, and E. Ganz, *Computational study of hydrogen binding by metal-organic framework-5*. *Journal of Chemical Physics*, 2004. **121**(24): p. 12543-12547.
45. Rowsell, J.L.C., et al., *Gas adsorption sites in a large-pore metal-organic framework*. *Science*, 2005. **309**(5739): p. 1350-1354.

46. Spencer, E.C., et al., *Determination of the hydrogen absorption sites in ZnO(1,4-benzenedicarboxylate) by single crystal neutron diffraction*. Chemical Communications, 2006(3): p. 278-280.
47. Rowsell, J.L.C., J. Eckert, and O.M. Yaghi, *Characterization of H-2 binding sites in prototypical metal-organic frameworks by inelastic neutron scattering*. Journal of the American Chemical Society, 2005. **127**(42): p. 14904-14910.
48. Yildirim, T. and M.R. Hartman, *Direct observation of hydrogen adsorption sites and nanocage formation in metal-organic frameworks*. Physical Review Letters, 2005. **95**(21).
49. McQuarrie, D.A., *Statistical Mechanics*. 2000, Sausalito, CA: University Science Books.
50. Pace, E.L. and A.R. Siebert, *Heat of Adsorption of Parahydrogen and Orthodeuterium on Graphon*. Physical Chemistry, 1959. **63**: p. 1398-1400.
51. Mattera, L., et al., *Selective Adsorption of H-1(2) and H-2(2) on the (0001) Graphite Surface*. Surface Science, 1980. **93**(2-3): p. 515-525.
52. Rowsell, J.L.C. and O.M. Yaghi, *Effects of functionalization, catenation, and variation of the metal oxide and organic linking units on the low-pressure hydrogen adsorption properties of metal-organic frameworks*. Journal of the American Chemical Society, 2006. **128**(4): p. 1304-1315.
53. Kaye, S.S. and J.R. Long, *Hydrogen storage in the dehydrated Prussian blue analogues M-3 Co(CN)(6) (2) (M = Mn, Fe, Co, Ni, Cu, Zn)*. Journal of the American Chemical Society, 2005. **127**(18): p. 6506-6507.
54. Wong-Foy, A.G., A.J. Matzger, and O.M. Yaghi, *Exceptional H-2 saturation uptake in microporous metal-organic frameworks*. Journal of the American Chemical Society, 2006. **128**(11): p. 3494-3495.
55. Hubner, O., et al., *On the interaction of dihydrogen with aromatic systems*. Journal of Physical Chemistry A, 2004. **108**(15): p. 3019-3023.
56. Leng, H., T. Ichikawa, and H. Fujii, *Hydrogen storage properties of Li-Mg-N-H systems with different ratios of LiH/Mg(NH<sub>2</sub>)<sub>2</sub>*. Journal of Physical Chemistry B, 2006. **110**(26): p. 12964-12968.
57. Meisner, G.P., et al., *Hydrogen release from mixtures of lithium borohydride and lithium amide: A phase diagram study*. Journal of Physical Chemistry B, 2006. **110**(9): p. 4186-4192.
58. Juza, R. and K. Opp, *Zur Kenntnis des Lithiumimides*. Zeitschrift fuer Anorganische und Allgemeine Chemie, 1951. **266**: p. 325-330.
59. Ohoyama, K., et al., *Revised crystal structure model of Li<sub>2</sub>NH by neutron powder diffraction*. Journal of the Physical Society of Japan, 2005. **74**(1): p. 483-487.

60. Noritake, T., et al., *Crystal structure and charge density analysis of Li<sub>2</sub>NH by synchrotron X-ray diffraction*. Journal of Alloys and Compounds, 2005. **393**(1-2): p. 264-268.
61. Herbst, J.F. and L.G. Hector, Jr., *Energetics of the Li amide/Li imide hydrogen storage reaction*. Physical Review B, 2005. **72**(12).
62. Zhang, C.J., M. Dyer, and A. Alavi, *Quantum delocalization of hydrogen in the Li<sub>2</sub>NH crystal*. Journal of Physical Chemistry B, 2005. **109**(47): p. 22089-22091.
63. Balogh, M.P., et al., *Crystal structures and phase transformation of deuterated lithium imide, Li<sub>2</sub>ND*. Journal of Alloys and Compounds, 2006. **420**(1-2): p. 326-336.
64. Magyari-Kope, B., V. Ozolins, and C. Wolverton, *Theoretical prediction of low-energy crystal structures and hydrogen storage energetics in Li[<sub>sub 2</sub>]NH*. Physical Review B (Condensed Matter and Materials Physics), 2006. **73**(22): p. 220101-4.
65. Kohn, W. and L.J. Sham, *Self-consistent equations including exchange and correlation effects*. Phys. Rev. A, 1965. **140**: p. 1133.
66. Forman, R.A., *Order-Disorder Transition in Lithium Imide, Li<sub>2</sub>NH*. The Journal of Chemical Physics, 1971. **55**(4): p. 1987-1988.
67. Kojima, Y. and Y. Kawai, *IR characterizations of lithium imide and amide*. Journal of Alloys and Compounds, 2005. **395**(1-2): p. 236-239.
68. de Fontaine, D., *Configurational Thermodynamics of Solid Solutions*, in *Solid State Physics*, D. Turnbull and F. Seitz, Editors. 1979, Academic Press. p. 73-273.
69. Sanchez, J.M., F. Ducastelle, and D. Gratias, *Generalized cluster description of multicomponent systems*. Physica, 1984. **128A**: p. 334-350.
70. Ceder, G., *The first-principles calculation of phase stability and thermodynamic properties*. Current Opinions in Solid State & Materials Science, 1998. **6**: p. 533-537.
71. Kresse, G. and J. Furthmuller, Phys. Rev. B, 1996. **54**: p. 11169.
72. Monkhorst, H.J. and J.D. Pack, *Special Points for Brillouin-Zone Integrations*. Physical Review B, 1976. **13**(12): p. 5188-5192.
73. Metropolis, N., et al., *Equations of state calculations by fast computing machines*. Journal of Chemical Physics, 1953. **21**(6): p. 1087-1092.
74. Van de Walle, A. and G. Ceder, *Automating First-Principles Phase Diagram Calculations*. J. of Phase Equilibria, 2002. **23**(4): p. 348-359.
75. Gonze, X., et al., *First-principles computation of material properties: the ABINIT software project*. Computational Materials Science, 2002. **25**(3): p. 478-492.

76. Korling, M. and J. Haglund, *Cohesive and Electronic-Properties of Transition-Metals - the Generalized Gradient Approximation*. Physical Review B, 1992. **45**(23): p. 13293-13297.
77. Fuchs, M., et al., *Pseudopotential study of binding properties of solids within generalized gradient approximations: The role of core-valence exchange-correlation*. preprint, 1997.
78. Fuchs, M., et al., *Cohesive properties of group-III nitrides: A comparative study of all-electron and pseudopotential calculations using the generalized gradient approximation*. Physical Review B, 2002. **65**(24).
79. Perlitz, H.A., E., *A redetermination of the crystal structure of lithium*. Philosophical Magazine, Serie 1940. **30**: p. 55-63.
80. M. Born, R.O., *Zur Quantentheorie der Molekeln*. Annalen der Physik, 1927. **389**(20): p. 457-484.
81. de Fontaine, D., *Cluster Approach to Order-Disorder Transformations in Alloys*, in *Solid State Physics*, H. Ehrenreich and D. Turnbull, Editors. 1994, Academic Press. p. 33-176.
82. Tikhonov, A.N. and V.Y. Arsenin, *Solutions of Ill-Posed Problems*. Scripta Series in Mathematics. 1977, Washington, D. C.: John Wiley & Sons.
83. Lindley, D.V. and A.F.M. Smith, *Bayes Estimates for the Linear Model*. Journal of the Royal Statistical Society. Series B (Methodological), 1972. **34**(1): p. 1-41.
84. Denison, D.G.T., et al., *Bayesian Methods for Nonlinear Classification and Regression*. Wiley Series in Probability and Statistics, ed. P. Bloomfield, et al. 2002, Chichester, U. K.: John Wiley & Sons.
85. Bayes, M. and M. Price, *An Essay towards Solving a Problem in the Doctrine of Chances. By the Late Rev. Mr. Bayes, F. R. S. Communicated by Mr. Price, in a Letter to John Canton, A. M. F. R. S.* Philosophical Transactions (1683-1775), 1763. **53**: p. 370-418.
86. Tibshirani, R., *Regression shrinkage and selection via the Lasso*. Journal of the Royal Statistical Society Series B-Methodological, 1996. **58**(1): p. 267-288.
87. Efron, B., et al., *Least angle regression*. Annals of Statistics, 2004. **32**(2): p. 407-451.
88. Drautz, R. and A. Diaz-Ortiz, *Obtaining cluster expansion coefficients in ab initio thermodynamics of multicomponent lattice-gas systems*. Physical Review B, 2006. **73**(22).
89. Laks, D.B., et al., *Efficient cluster expansion for substitutional systems*. Physical Review B, 1992. **46**(19): p. 12587-12605.
90. Henderson, H.V. and S.R. Searle, *On Deriving the Inverse of a Sum of Matrices*. Siam Review, 1981. **23**(1): p. 53-60.



91. Golub, G.H., M. Heath, and G. Wahba, *Generalized Cross-Validation as a Method for Choosing a Good Ridge Parameter*. *Technometrics*, 1979. **21**(2): p. 215-223.
92. Hastie, T.J. and R.J. Tibshirani, *Generalized Additive Models*. First ed. Monographs on Statistics and Applied Probability, ed. D.R. Cox, et al. 1990, Bury St. Edmunds: Chapman and Hall.
93. Domich, P.D., R. Kannan, and L.E. Trotter, *Hermite Normal-Form Computation Using Modulo Determinant Arithmetic*. *Mathematics of Operations Research*, 1987. **12**(1): p. 50-59.
94. Gale, J.D. and A.L. Rohl, *The General Utility Lattice Program (GULP)*. *Molecular Simulation*, 2003. **29**(5): p. 291-341.
95. Sutton, A.P. and J. Chen, *Long-Range Finnis Sinclair Potentials*. *Philosophical Magazine Letters*, 1990. **61**(3): p. 139-146.
96. Tersoff, J., *Modeling Solid-State Chemistry - Interatomic Potentials for Multicomponent Systems*. *Physical Review B*, 1989. **39**(8): p. 5566-5568.