

Abstract

Thermoelectric inorganic clathrates of type I are under intense experimental and theoretical study. Partly due to its embodiment of the phonon glass electron crystal property but also because the diversity of elements which can make up the structure. This allows for a fine tuning of thermoelectrical properties, producing high zT values. Density functional theory calculations and experimental work sample only a small subset of materials and compositions. In this thesis, cluster expansions based on the compressive sensing algorithm is developed and used together with Monte Carlo simulations. This method allow the investigation of configurational dependency on temperature and composition. The clathrates $\text{Ba}_8\text{Ga}_{16}\text{Ge}_{30}$, $\text{Ba}_8\text{Ga}_{16}\text{Si}_{30}$, $\text{Ba}_8\text{Al}_{16}\text{Ge}_{30}$ and $\text{Ba}_8\text{Al}_{16}\text{Si}_{30}$ are studied as well as their non-stoichiometric versions. Calculated site occupancy factors are in good agreement with experimental work. The experimental works can be understood with stoichiometry and vacancies. Secondly, it is found that the band gap decreases as the disorder increases for all clathrates. Finally the cluster expansion in conjunction with the Monte Carlo method produces ground state structures that accurately predicts the energy which is confirmed with density functional theory calculations.

Acknowledgements

I would like to express my gratitude to my supervisor, Paul Erhart, for his guidance throughout this thesis. It has been invaluable. Secondly I would like to thank Erik Fransson. His discussions and company on lunch breaks are much appreciated. Finally I would like to give a big thanks to my family who have been a great support during these years at Chalmers.

Mattias, Gothenburg 2015-03-23

Contents

1	Introduction	1
1.1	Thermoelectrics	1
1.2	Thermoelectric materials	4
1.2.1	Half heusler phases	4
1.2.2	Layered materials	4
1.2.3	Clathrates	4
1.2.4	Zintl concept	5
2	Method	9
2.1	Cluster Expansion	10
2.2	Compressive sensing	11
2.3	Cross-Validation	13
2.4	The Monte Carlo Method	13
2.4.1	The Metropolis Method	14
3	Ba₈Ga₁₆Ge₃₀	18
3.1	Fitting the Ba ₈ Ga ₁₆ Ge ₃₀ system	18
3.2	Results	19
4	Other binaries	26
4.1	Result	26
5	Quaternaries	34
6	Discussion	36
6.1	Site occupancy factors	36
6.1.1	Ba ₈ Ga ₁₆ Ge ₃₀	36
6.1.2	Ba ₈ Ga ₁₆ Si ₃₀	37
6.1.3	Ba ₈ Al ₁₆ Ge ₃₀	38
6.1.4	Ba ₈ Al ₁₆ Si ₃₀	39
6.1.5	Rules for occupancy	39

6.1.6	General conclusions for all binaries	40
6.2	Outlook	40
7	Conclusion	41
	Bibliography	45

1

Introduction

The world's demand for energy continues to grow. Combusting more fossil fuels to meet these demands increases human environmental impact and the depletion rate of these fuels. Another way to improve energy output without increasing the fuel consumption is by increasing the efficiency of electric generators. This can be accomplished by scavenging waste heat with thermoelectric generators. There are countless processes which release waste heat. Thermoelectrics have no moving parts, they are silent, reliable and require no maintenance.

The cradle of thermoelectricity was in 1821 when Baltic German physicist Thomas Johann Seebeck discovered the seebeck effect.[1] He observed that a compass needle would be deflected if it was placed near a closed loop formed from two different conductors when one of the junctions was heated. After that the field of thermoelectricity took a slow start. It had a temporary uprising in 1850 due to advances in thermodynamics. In the 1950s with the development of synthetic semiconductors the basic science of thermoelectrics started to establish itself and a few semiconductors were manufactured that had a relatively high rate of conversion compared to today. [1, 2] In the following three decades, the development of thermoelectrics was slow but steady. It found niche applications in e.g. space probes.[1, 2] For deep space missions, far away from the sun, radioisotope thermoelectric generators (RTGs) where the heat source come from radioactive decay allow for electric power without maintenance for extended periods of time. In the Voyager spacecraft, launched in 1977 RTG provided Voyager with energy for longer than 17 years. [1]

1.1 Thermoelectrics

A thermoelectric generator will, when put between a hot and a cold place, i.e. under the influence of a thermal gradient, generate a current. The maximum efficiency, i.e. how

much heat is transformed into electricity, of a thermoelectric material is given by [1],

$$\eta = \frac{(T_H - T_C)}{T_H} \frac{(1 + ZT)^{1/2} - 1}{(1 + ZT)^{1/2} + T_C/T_H} \quad (1.1)$$

Where T_H and T_C are the absolute temperatures of the hot and cold reservoirs. ZT is a figure of merit which involves both the n-type and p-type properties of the thermoelements[1]. It is common practice to instead use zT , the figure of merit for a single material. The first part of Eq(1.1) is the Carnot limit and is, according to Carnot's theorem, the limit for heat engines and is defined as $1 - T_C/T_H$. For $T_C = 300$ K, η as a function of ZT is visualized in figure 1.1. The figure of merit, zT is given by

$$zT = \frac{S^2 \sigma T}{\kappa} \quad (1.2)$$

It is dimensionless and can be used to compare the efficiency of thermoelectrics. S is the Seebeck coefficient, σ is the electrical conductivity and κ is the thermal conductivity.

The basic physics of a thermoelectric material is that application of a thermal gradient creates a potential gradient. This can be explained by considering that the charge carriers are free to move like gas molecules while transporting charge and heat. When a temperature gradient is applied to a material the charge carriers near the hot reservoir will move faster than the charge carriers near the cold reservoir resulting in a diffusion of charge carriers from the hot to the cold reservoir producing an electrostatic potential. This is known as the Seebeck effect.

When designing a thermoelectric material, it is essential to maximize the zT value. The challenge is that many of the properties of the thermoelectric have conflicting effects on the figure of merit. The basic properties of a thermoelectric is:

- **Seebeck effect** The Seebeck coefficient, S is directly related to the conversion of temperature gradients to electrical potential gradient. It is defined by,

$$S = \frac{V_2 - V_1}{T_2 - T_1} \quad (1.3)$$

Where $V_2 - V_1$ is the potential between point 1 and 2 and $T_2 - T_1$ is the temperature difference. A high Seebeck coefficient is desired. To this end it is optimal to have only one type of charge carriers. A mix of charge carriers will make both electrons and holes diffusing to the cold end resulting in a lower voltage difference. For metals or degenerate semiconductors the Seebeck coefficient is given by,

$$S = \frac{8\pi^2 k_B^2}{3eh^2} m^* T \left(\frac{\pi}{3n} \right)^{2/3} \quad (1.4)$$

where n is the carrier concentration, m^* is the effective mass of the carrier, e is the electron charge, h is Planck's constant and k_B is Boltzmann's constant.

- **Electrical conductivity - Ohm's law.**

Ohm's law states that the current through a conductor between two points is proportional to the potential difference between two points.

$$\mathbf{J} = \sigma \mathbf{E} \quad (1.5)$$

where \mathbf{J} is the current density, σ is the conductivity and \mathbf{E} is the electric field. For a thermoelectric which purpose is to create a current it is clear that a high conductivity is wanted. The electrical conductivity is related to the carrier concentration by

$$\sigma = ne\mu \quad (1.6)$$

where μ is the carrier mobility.

- **Thermal conductivity - Fourier's law.**

Fourier's law shows that the local heat flux density \mathbf{q} is dependent on the thermal conductivity κ and the temperature gradient ∇T ,

$$\mathbf{q} = -\kappa \nabla T \quad (1.7)$$

The thermal conductivity comes from two sources: 1) charge carriers transporting heat (κ_e) and 2) phonons traveling through the lattice (κ_l) thus we have $\kappa = \kappa_e + \kappa_l$. The electronic contribution κ_e is proportional to the electrical conductivity σ as stated by the Wiedemann-Franz law,

$$\kappa_e = \sigma LT \quad (1.8)$$

where L is the Lorenz number. Now zT can be reformulated in a more detailed way:

$$zT = \frac{S^2 \sigma T}{\kappa} = \frac{S^2 \sigma T}{\sigma LT + \kappa_l} \quad (1.9)$$

As stated previously a high flux of electrons is desirable, which is dependent on the temperature difference. If heat is transferred by means of thermal conduction this will decrease the efficiency of the thermoelectric ergo a low thermal conductivity is required.

The carrier concentration have a conflicting effect of the figure of merit since a high carrier concentration decreases the Seebeck coefficient but increases the electrical conductivity. Another conflicting property is the effective mass of the charge carriers as larger effective masses produce high Seebeck coefficient but low electrical conductivity.

Heavy carriers will move with slower velocities and therefore have smaller mobilities.

A good thermoelectric is a material with a low lattice thermal conductivity and high electric conductivity. Slack *et al.* [3] suggested that solids with weakly bonded atoms which can “rattle” in their positions may scatter phonons, leading to a lower thermal conductivity. However, rattlers scattering phonons may equally well scatter electrons. To prevent this, one must find a material where the rattlers scatter phonons but do not disturb charge carriers. Such a material is called a phonon glass-electron crystal (PGEC). A PGEC material with a high Seebeck coefficient is likely to be a very good thermoelectric material.

1.2 Thermoelectric materials

Here is a short description of different interesting thermoelectric materials. With the exception of clathrates, there will not much else said about the other materials.

1.2.1 Half heusler phases

Half heusler phases are interesting thermoelectric materials. They can exhibit many inhomogeneities such as interstitials, small Heusler inclusions and multiple Half-Heusler phases.[4]. Inhomogeneities are often interesting for potential thermoelectric material. Having many of them can allow a fine tuning to achieving great thermoelectric materials.

1.2.2 Layered materials

The concept behind layered materials is that instead of designing a material to both have low thermal conductivity and high thermopower, combination of two or more nanoblocks are used, creating a complex crystals. One example of this is the layered CoO_2 [5].

1.2.3 Clathrates

Inorganic clathrates or sometimes called Zintl clathrates[6] are cage-like structures, which enclose bonded atoms (guest atoms) and can be composed of elements from groups 13 and 14 of the periodic table (host atoms). The guest atoms, often loosely bonded in the cages, can act as rattlers, scattering phonons. Clathrates are thereby practical realizations of the so-called phonon glass-electron crystal concept [3].

Clathrates are classified according to crystal symmetry. In the present study the focus is on type I clathrates, which adopt space group $Pm\bar{3}n$. In this structure the unit cell contains 46 host atoms that leaves 8 voids or cages for the guest atoms to occupy. The cages are two small dodecahedron and 6 larger tetrakaidecahedron, where the centers of the cages correspond to Wyckoff sites 2a and 6d respectively. The host atoms are found at sites 6c, 16i and 24k (see figure 1.2) The site occupance factors (sof) of the Wyckoff sites are an important description of the clathrates. It is a quantity which

can be experimentally observed and it is related to the local order. If the group 13 and 14 atoms were uniformly distributed the sof's would be 35% group 13 and 65% group 14.

The host elements from group 13 and 14 assume a certain ratio (16:30). This stoichiometry can be explained by the Zintl concept.

1.2.4 Zintl concept

Much of the behaviour can be understood with the Zintl concept [7]. In this model, four electrons must be available for each tetrahedrally bonded host atom. The guest atom is assumed to be ionic and donate its valance electrons to the framework[7]. This means that with 46 framework atoms, there must be 184 electrons available for the framework bonding. For $\text{Ba}_8\text{Ga}_{16}\text{Ge}_{30}$ each Ba guest atom donates two electrons and the the number of valancies for Ga and Ge is 3 and 4. This gives a total of 184 electrons and the compound is expected to be a semiconductor. Deviation from 16:30 ratio will result in doping in this model. Exchanging a Ge atom for a Ga atom will result in one less electron in the framework and the material is assumed p-doped. A Ga to Ge exchange leads to an excess electron and n-doping.

Because of the interesting properties of clathrates, they have been investigated in numerous papers. Christensen *et al.* 2007 [8] experimented on three different samples of $\text{Ba}_8\text{Al}_{16}\text{Ge}_{30}$. Each was prepared in a different manner, producing very different site occupancy factors. The experiment showed that one of the samples had a thermal conductivity 3 times smaller than another. They claim that controlling the host structure chemistry is key to manipulating electrical properties and thermal conductivity.

This example is an example of understanding the chemical ordering is key in understanding how the thermoelectric will perform. Experiments and density functional theory (DFT) calculations only sample a small subset of materials and configurations. The main objective of this thesis is to elucidate the relation between chemical order in type-I clathrates and certain thermodynamic properties. To this end cluster expansions[9] has been developed using a compressive sensing algorithm[10], which was then used in Monte Carlo simulations. This approach made it possible to sample configurational dependency on temperature and composition.

The main focus of this thesis has been on the quasi-binary systems $\text{Ba}_8\text{Ga}_{16}\text{Ge}_{30}$, $\text{Ba}_8\text{Ga}_{16}\text{Si}_{30}$, $\text{Ba}_8\text{Al}_{16}\text{Ge}_{30}$ and $\text{Ba}_8\text{Al}_{16}\text{Si}_{30}$. The main results are the site occupancy factors which are in good agreement with experimental results for all binaries. This validates the method used, allowing to extend the calculations to other properties. The temperature dependence of the band gap is calculated. The findings show that band gap is lowered with increasing disorder and shows a large increase for decreasing temperatures for all binaries. This result is significant because of the exponential dependence of the charge carrier concentration[11] and thus the electrical conductivity, as seen in Eq.(1.6). The lattice parameter show a significant dependence on the ordering of the clathrate. All material show a decrease in lattice parameter for increasing order, except for $\text{Ba}_8\text{Ga}_{16}\text{Si}_{30}$ showing an increase in lattice parameter for increasing order.

Finally the compressive sampling cluster expansion method in conjunction with the Monte Carlo method produces ground state structures and accurately predicts the energy which are confirmed by DFT calculations. These ground state configurations are significant because they represent prototypical structures that embody the key chemical features in a high symmetry structure. They can thereby serve as starting points for further theoretical investigations.

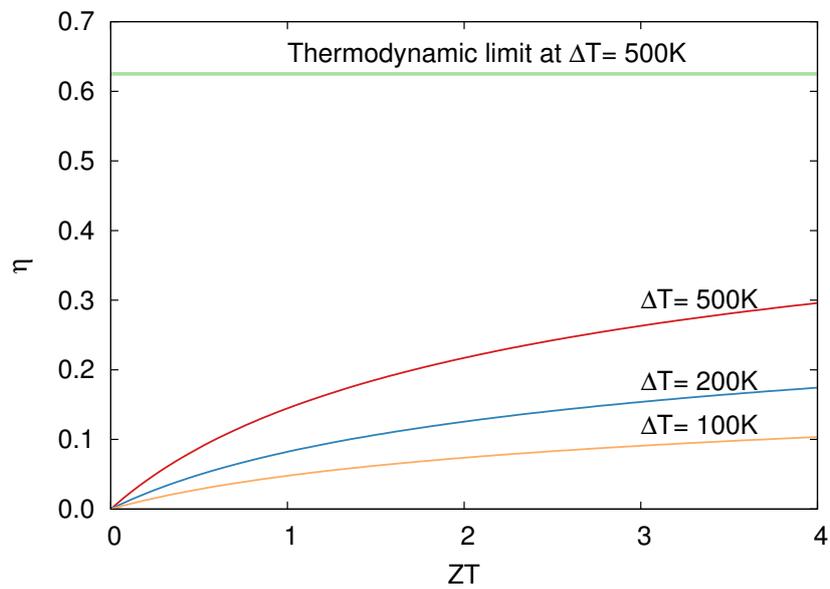


Figure 1.1: The conversion efficiency as a function of ZT with $T_C = 300K$ for different temperature differences. The thermodynamic limit is given by Carnot's theorem which is $1 - T_C/T_H$

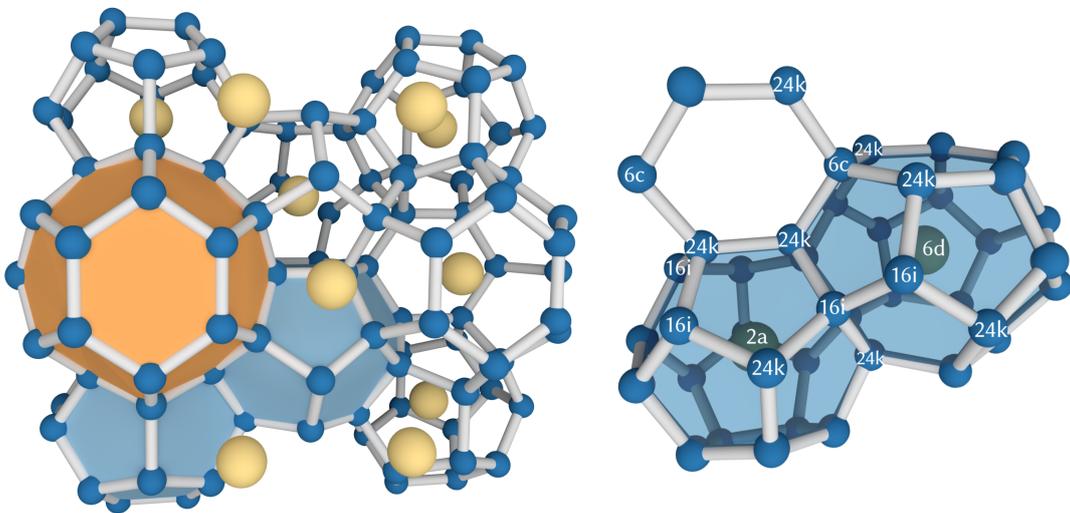


Figure 1.2: Left) Crystal structure of clathrate-I. Right) The different cages of the clathrates with some of the Wyckoff positions shown. The cage in the upper right is the larger tetrakaidecahedron cage with 14 sides, the cage on the bottom left is a smaller dodecahedron cage with 12 sides.

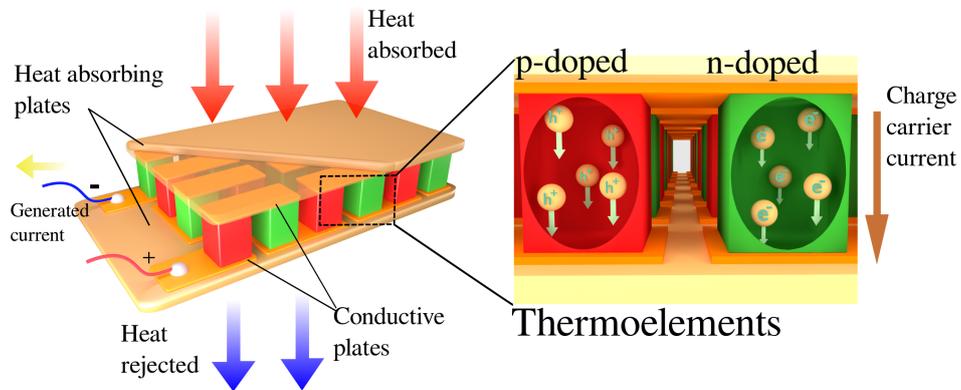


Figure 1.3: The left hand of the figure show the thermoelectric device. It consists of a series of thermoelectric elements with p- and n-doping alternating. On the right hand of the figure is a closer look on the thermoelements, a heat flow gives rise to a current of charge carrier, the setup of alternating thermoelements will result that each thermo-couple will act as a charge source in a series.

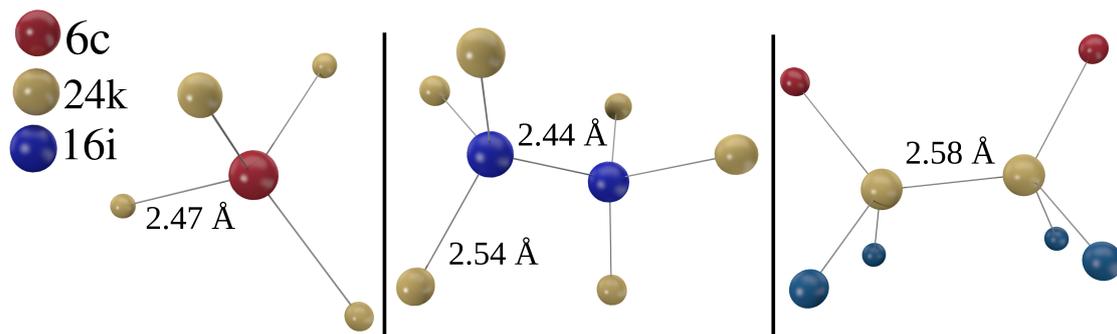


Figure 1.4: Atomic environment of different crystal positions of the cage.

2

Method

Many solid-state structures can be described as substitutional A/B systems, where the sites of the crystal lattice are occupied by A and B atoms in different configurations. For simplicity ternary and higher order systems are ignored, but can without much effort be generalized to include more atom types. If one is interested in the ground state configuration or to sample finite-temperature thermodynamic averages you would in principle be forced to know the energy of the 2^N possible configurations for placing A and B atoms on N lattice sites. For ab-initio calculations this present a problem in that the computational effort for quantum-mechanical total energy calculations increases rapidly with the number of atoms in the unit cell so you are limited to small or few configurations. This makes ab initio calculations unsuitable for determining the ground state configuration, the energy of random configurations or for construction phase diagrams.

One of the simplest models for calculating the configurational dependency of a quantity, q , is the Ising model,

$$q(\sigma) = - \sum_{i \neq j} J \sigma_i \sigma_j \quad (2.1)$$

Which is a sum over the nearest neighbours of lattice sites i and j , J is an effective interaction energy and σ_i , σ_j is the atomic spin. For a binary system, one may assign values of +1 and -1 for atom type A and B respectively. Of course, only considering nearest neighbours might not pick up all the subtleties a material may exhibit since atomic interactions seldom extend only to the nearest neighbours. This leads to a version of an extended Ising model where you consider pairs of atoms inside a certain radius, r_{cut} ,

$$q(\sigma) = - \sum_{r < r_{cut}} J_{i,j} \sigma_i \sigma_j \quad (2.2)$$

One may now in this fashion add the effect of all imaginable clusters,

$$q(\sigma) = \sum_i J_i \sigma_i + \sum_{i \neq j} J_{i,j} \sigma_i \sigma_j + \sum_{i > j > k} J_{i,j,k} \sigma_i \sigma_j \sigma_k + \dots \quad (2.3)$$

If this expansion is carried out to infinity it will calculate the exact energy for a certain choice of parameters [9]. This expansion of clusters is called the Cluster Expansion. The generalisation of which is described in the next section.

2.1 Cluster Expansion

One approach to combat the weakness of direct ab initio calculations for different configurations is the cluster expansion (CE). A configurationally dependent quantity, q , is given by the CE as:

$$q(\sigma) = \sum_{\alpha} m_{\alpha} J_{\alpha} \langle \Gamma_{\alpha'}(\sigma) \rangle_{\alpha} \quad (2.4)$$

where the following definitions according to Ref.[12] were used:

- A configuration σ is represented by a vector σ_i indicating which type of atom occupies lattice site i . The occupation value can take any value from 0 to $M_i - 1$ where M_i is the number of atom types can occupy site i .
- α is a cluster such as singlets, pairs, triplets etc. A cluster can be described by a vector of elements α_i that can take any value from 0 to $M_i - 1$ where 0 indicates that site i doesn't belong to the cluster.
- The sum is over all possible clusters that are mutually distinct.
- The average $\langle \dots \rangle_{\alpha}$ is over all clusters α' that are equivalent by symmetry to cluster α .
- $\Gamma_{\alpha'}(\sigma)$ are cluster functions. They are selected to be of the form

$$\Gamma_{\alpha}(\sigma) = \prod_i \gamma_{\alpha_i, M_i}(\sigma_i) \quad (2.5)$$

where $\gamma_{\alpha_i, M_i}(\sigma_i)$ satisfies $\gamma_0(\sigma_i) = 1$ and the following orthogonality condition

$$\sum_{\sigma_i=0}^{M_i-1} \gamma_{\alpha_i, M_i}(\sigma_i) \gamma_{\beta_i, M_i}(\sigma_i) = \begin{cases} \neq 0 & \text{if } \alpha_i = \beta_i \\ 0 & \text{otherwise} \end{cases} \quad (2.6)$$

- m_{α} is the cluster count, indicating the number of clusters equivalent to α by symmetry
- J_{α} are the coefficients to be determined and are also called effective cluster interaction (ECI)

When all clusters α are considered in the sum, the cluster expansion is able to represent any function $q(\sigma)$ if the $\Gamma_\alpha(\sigma)$ form an orthogonal basis for the space of configurations, which is true if the orthogonality condition, (2.6) is satisfied [12].

The choice for $\gamma_{\alpha_i, M_i}(\sigma_i)$ is valid for any function satisfying eq.(2.6). This thesis uses the same choice as in Ref.[12] which is the following:

$$\gamma_{\alpha_i, M_i}(\sigma_i) = \begin{cases} 1 & \text{if } \alpha_i = 0 \\ -\cos\left(2\pi \left[\frac{\alpha_i}{2}\right] \frac{\sigma_i}{M}\right) & \text{if } \alpha_i > 0 \text{ and odd} \\ -\sin\left(2\pi \left[\frac{\alpha_i}{2}\right] \frac{\sigma_i}{M}\right) & \text{if } \alpha_i > 0 \text{ and even} \end{cases} \quad (2.7)$$

where $[\dots]$ is the "round up" operation and α_i and σ_i can range from 0 to $M_i - 1$

2.2 Compressive sensing

The fitting process in this thesis seeks to reproduce DFT data, taken from a data set [13]. It is common practice to divide a data set into a training set and one validation set. The training set is used to fit the parameters, that solution can then be used to predict the data from the validation set. If the prediction is good for the validation set, one can assume that it will be good for other configurations that are not part of the data set. When considering the fitting process the problem is equivalent to solving the matrix equation,

$$\bar{\Pi} \vec{J} = \vec{E} \quad (2.8)$$

Where \vec{E} contains the training data and \vec{J} are the coefficients to be fitted and $\bar{\Pi}$ contains the configurations in such a way that $\Pi_{i,j}$ is the value of $m_\alpha \langle \Gamma_{\alpha'}(\sigma) \rangle_\alpha$ from Eq. (2.4) for the i :th configuration and the j :th cluster.

Considering all possible clusters in Eq. (2.4) is not practically possible. However physical intuition suggests that only clusters with a limited radius and involving a limited number of sites will have a significant ECI. There are many intricate methods that involve taking out sets of trial ECI's and check their predictive power [10]. These methods is cumbersome compared to a recently developed technique, compressive sensing (CS), which provides a simple and efficient way to extract the important ECI's and compute their values in one shot [10, 14, 15]. To demonstrate compressive sensing it helps to first define the l_p norm:

$$\|u\|_p = \left(\sum_i |u_i|^p \right)^{1/p} \quad (2.9)$$

In the compressive sensing (CS) method, the problem is solved by searching for the solution with the smallest l_1 norm which still reproduces the results with a given accuracy,

$$J_{CS} = \arg \min_J \left\{ \|\vec{J}\|_1 : \|\bar{\Pi} \vec{J} - \vec{E}\|_2 < \epsilon \right\} \quad (2.10)$$

This form is a inconvenient to work with and it is common practice to work with an unconstrained approach that minimizes the l_1 norm and the least squares sum of the fitting error:

$$J = \arg \min_J \left\{ \mu \|J\|_1 + \frac{1}{2} \left\| \vec{E} - \bar{\Pi}J \right\|^2 \right\} \quad (2.11)$$

the parameter μ controls the accuracy of the fit. A high value of μ leads to sparse solution but larger fitting error and vice versa.

However it turns out that there are problems to solve mixed l_1 and l_2 minimization problems such as Eq.(2.11), Goldstein and Osher proposed the split Bregman algorithm which eliminates this problem [16]. The split Bregman iteration splits the l_1 norm of the solution from the objective function and replaces it with a variable \vec{d} which then converges towards the l_1 term, $\lim_{k \rightarrow \infty} (\vec{d} - \mu\vec{J}) = \mathbf{0}$. To this end a least-squares l_2 term is added to the objective function to ensure that $\vec{d} = \mu\vec{J}$

$$\vec{J} = \arg \min_{J,d} \left\{ \|\vec{d}\|_1 + \frac{1}{2} \left\| \bar{\Pi}\vec{J} - \vec{E} \right\|^2 + \frac{\lambda}{2} \left\| \vec{d} - \mu\vec{J} \right\|^2 \right\} \quad (2.12)$$

This formulation is advantageous because the minimization involving the quadratic form $\frac{1}{2} \left\| \bar{\Pi}\vec{J} - \vec{E} \right\|^2$ does not involve any l_1 terms and can be minimized efficiently using the Broyden–Fletcher–Goldfarb–Shanno algorithm (BFGS).

The split Bregman algorithm comprises the following steps

$$\vec{J}^{k+1} = \arg \min_J \left\{ \frac{1}{2} \left\| \bar{\Pi}\vec{J} - \vec{E} \right\|^2 + \frac{\lambda}{2} \left\| \vec{d}^k - \mu\vec{J} - \vec{b}^k \right\|^2 \right\} \quad (2.13)$$

$$\vec{d}^{k+1} = \arg \min_d \left\{ \|\vec{d}\|_1 + \frac{\lambda}{2} \left\| \vec{d} - \mu\vec{J}^{k+1} - \vec{b}^k \right\|^2 \right\} \quad (2.14)$$

$$\vec{b}^{k+1} = \vec{b}^k + \mu\vec{J}^{k+1} - \vec{d}^{k+1} \quad (2.15)$$

Where in Eq.(2.15) the residual after iteration k is added back to the residual vector \vec{b}^{k+1} for the next iteration in style with a Bregman iteration which result in a quicker convergence [17]. Starting from $\vec{d}^0 = 0$, $\vec{b}^0 = 0$ and $\vec{J}^0 = 0$. In this thesis the BFGS is used to solve the l_2 minimization in Eq.(2.13), the second step, Eq.(2.14) separates into individual vector components and is solved by shrinkage,

$$\vec{d}_n^{k+1} = \text{shrink}(\mu\vec{J}_n^{k+1} + \vec{b}_n^k, 1/\lambda) \quad (2.16)$$

which is defined by

$$\text{shrink}(y, \alpha) \equiv \text{sign}(y) \max(|y| - \alpha, 0) \quad (2.17)$$

Shrinkage decreases the absolute magnitude of the y by α and sets it to zero if $y \leq \alpha$.

2.3 Cross-Validation

Since the number of parameters to be fitted in a cluster expansion is theoretically infinite there will have to be some truncation, there is also the problem of overfitting if too many parameters are fitted which leads to lower training set error but worse predictability. Another problem with choosing too few parameters is that there might be a correlation between $\langle \Gamma_{\alpha'}(\sigma) \rangle_{\alpha}$ and $\langle \Gamma_{\beta'}(\sigma) \rangle_{\beta}$ and when J_{β} is left out, J_{α} might get accredited for the variations of cluster β which has an adverse effect on the predictability on the fit.

The cross validation (CV) score is widely accepted as the quantity for determining the accuracy of the CE. One variant is the 'leave one out' CV (LOO-CV) which is defined as

$$(CV)^2 = \frac{1}{N} \sum_{n=1}^N (\hat{E}_{(n)} - E_n)^2 \quad (2.18)$$

where E_n is the calculated energy for structure n and $\hat{E}_{(n)}$ is the predicted value of the energy of structure n as calculated with the CE fit with the $(N - 1)$ other structures. This method, however, requires making N cluster expansions, there is another method which only needs one cluster expansion [18][19],

$$(CV)^2 = \frac{1}{N} \sum_{n=1}^N \left(\frac{(\hat{E}_{(n)} - E_n)}{1 - X_i (X^T X)^{-1} X_i^T} \right)^2 \quad (2.19)$$

which includes all the structures in the data set in the fit and the matrix $X_{i\alpha}$ contain $\langle \Gamma_{\alpha'}(\sigma) \rangle_{\alpha}$ for the i :th configuration.

2.4 The Monte Carlo Method

The partition function is defined as $Z = \sum e^{-\beta U}$, where β is equal to $1/k_B T$, and the sum is over all micro states. The partition function is related to the probability P_s that micro state s is occupied,

$$P_s = \frac{1}{Z} e^{-\beta U_s} \quad (2.20)$$

The ensemble average of the energy is given by the sum of the energy weighted by their probability,

$$\langle E \rangle = \frac{1}{Z} \sum_s E_s e^{-\beta E_s} \quad (2.21)$$

Since there are 2^N micro states for a binary with N lattice sites an explicit evaluation of Eq.(2.20) is impractical, even for small systems. A numerical recipe for computing thermal averages is the Metropolis Monte Carlo method. To illustrate this approach first consider the evaluation of an integral using (unbiased) Monte Carlo,

$$I = \int_0^1 dx f(x) = \langle f(x) \rangle \quad (2.22)$$

The direct Monte Carlo method would evaluate $f(x)$ by choosing N points x_i at random with uniform probability over the interval $[0,1)$. The mean value becomes,

$$I_N = \langle f \rangle = \frac{1}{N} \sum_{i=1}^N f(x_i) = \frac{1}{N} \sum_{i=1}^N f_i \quad (2.23)$$

and the variance,

$$\sigma_f^2 = \langle f^2 \rangle - \langle f \rangle^2 \quad (2.24)$$

and the integral,

$$I = I_N \pm \frac{\sigma_f}{\sqrt{N}} \quad (2.25)$$

As $N \rightarrow \infty$ the Monte Carlo method will find the correct value of I . However, this method is not suitable for calculating averages such as in Eq.(2.21) since most of the computing will be spent on points where the Boltzmann factor is negligible which will increase the variance and the points needed to converge the sum. It is much more favourable to sample more points where the Boltzmann factor is large.

2.4.1 The Metropolis Method

Metropolis *et al* 1953 introduced the metropolis algorithm to determine the equation of state for a hard sphere liquid [20]. It is based on an idea that thermodynamic averaging only requires knowledge of relative rather than absolute probabilities like in Eq. (2.21). It uses a Markov chain to generate configurations that are more important by rejecting configurations that are unlikely, much like in the case of importance sampling and based on a distribution that is proportional to the Boltzmann factor.

A Markov chain is a stochastic process where a system undergoes transitions from one state to another. The Markov process is characterized by a lack of memory of where it has been. The future of the chain depends solely on the current state. The transition probability is given by a transition matrix,

$$T(X \rightarrow X') \equiv T_{XX'} \quad (2.26)$$

for a transition from state X to state X' . The transition probability has to satisfy

$$0 \leq T_{XX'} \leq 1 \quad (2.27)$$

and the probability has to be normalised,

$$\sum_{X'} T_{XX'} = 1 \quad (2.28)$$

The task is to generate a Markov chain of configurations such that they have a distribution proportional to the Boltzmann factor and this distribution should be independent on the position of the chain and of the initial configuration. The Markov chain can exhibit these properties under certain conditions, at least for sufficient long time so that the configuration can loose memory of its initial state. These conditions are,

- The markov chain needs to be irreducible, that is every configuration included in the ensemble should be accessible from every other configuration within a finite number of steps.
- There should be no periodicity. Periodicity means that it is not possible to revisit a configuration except after $t = nk$ steps, $n = 1,2,3 \dots$, where k is fixed.

A Markov chain that satisfies these conditions is called ergodic. If the Markov chain is ergodic it converges to a unique stationary distribution.

The transition probability need to be chosen such that the stationary distribution is the Boltzmann distribution. To assure this, consider the stationary distribution $\rho(X)$; one can also introduce a new function $\rho(X,t)$ which gives the probability of finding configuration X after t Markov steps which for an ergodic chain becomes independent of t if t is large. This function can change from one step to another by,

- Going from X at step t to X' at $t + 1$ leads to a decrease in $\rho(X)$
- Going from X' at step t to X at $t + 1$ leads to an increase in $\rho(X)$

that can be summarized with,

$$\rho(X,t+1) - \rho(X,t) = - \sum_{X'} T(X \rightarrow X')\rho(X,t) + \sum_{X'} T(X' \rightarrow X)\rho(X',t) \quad (2.29)$$

This equation is called the Master equation. The stationary solution of this equation is found by requiring $\rho(X,t+1) = \rho(X,t)$ so we have,

$$\sum_{X'} T(X \rightarrow X')\rho(X,t) = \sum_{X'} T(X' \rightarrow X)\rho(X',t) \quad (2.30)$$

Leaving out the t -dependence, which is allowed due to basic "memory loss" property of a Markov chain, yields,

$$T(X \rightarrow X')\rho(X) = T(X' \rightarrow X)\rho(X') \quad (2.31)$$

which is known as the condition of detailed balance. This means that in equilibrium the average number of moves that result in the system leaving state X must be exactly equal to the number of moves from all other states X' to X . This means that $\rho(X)$ and $\rho(X')$ doesn't change, as this is true for all pair of X and X' the probability distributions will remain stationary.

Reformulating the detailed balance condition with the transition probability in the form yields,

$$T(X \rightarrow X') = \omega_{XX'} A_{XX'} \quad (2.32)$$

where $\omega_{XX'}$, $\sum_{X'} \omega_{XX'} = 1$, is the probability for going from state X to state X' also known as trial step probability. $A_{XX'}$ which must lie between 0 and 1 is the acceptance probability of actually committing the change. The detailed balance condition can then be expressed as,

$$\frac{A_{XX'}}{A_{X'X}} = \frac{\rho(X')}{\rho(X)} \quad (2.33)$$

The Metropolis algorithm can now be formulated as follows:

- For a given state X , make a small trial move into a new state X' with a probability of $\omega_{XX'}$.
- Compare the weights of the distribution for the different states $\rho(X)$ and $\rho(X')$. $A_{XX'}$, the acceptance probability, is chosen equal to 1 if $\rho(X') > \rho(X)$ else it is chosen to be equal to $\rho(X')/\rho(X)$.
- The new state X' is accepted with a probability $A_{XX'}$, which means that X' replaces the old X system and it is rejected with probability $1 - A_{XX'}$ which means that the system remains in the state X . To decide if a state is accepted or not a random number is generated uniformly in the range $[0,1)$ and compared to the acceptance probability. If the random number is larger than the acceptance probability the trial move is accepted. Since the probabilities have a Boltzmann distribution the acceptance probability will look like,

$$A_{XX'} = e^{-\beta U(X)}/e^{-\beta U(X')} = e^{-\beta \Delta U} \quad (2.34)$$

Since each trial move is only a small change in the configuration there is an inherent correlation between state X and X' . There is a correlation step length, s , for the Markov chain that means it is necessary for s trial moves before reaching a new uncorrelated configuration. An MC step is defined as N trial steps where N is the number of particles which is commonly used to approximate s . There is also a need to initialize the initial configuration meaning that it is necessary to run the Metropolis algorithm, usually in the order of 10 – 50 MC steps depending on temperature and system.

The ensemble average of a physical quantity A is thus chosen as,

$$\langle A \rangle = \frac{1}{N} \sum_{i=1}^N A_i \quad (2.35)$$

where N is now the number of MC-steps, A_j is the value of A for configuration X_j and A_{j+1} is the value for configuration X_j after s trial moves and A_1 is the value of A

for the initial configuration X_{init} after the initial MC steps which are thrown away. The variance for A is,

$$\text{Var}(A) = \langle A^2 \rangle - \langle A \rangle^2. \quad (2.36)$$

The statistical error is

$$\frac{\sqrt{\text{Var}(A)}}{\sqrt{N}} \quad (2.37)$$

This is true if the sampling of A is uncorrelated. In this thesis the number MC-steps is of the order 10^4 making it effectively negligible.

3

Ba₈Ga₁₆Ge₃₀

The Ba₈Ga₁₆Ge₃₀ system is a type I clathrate with framework atoms consisting of Ga and Ge and Ba as host atoms. Only Ba are allowed as host atoms, and no vacancies are considered for the host atoms. This means that there can be no configurational difference with respect on the Ba atoms in the cluster expansion. Thus the Ba₈Ga₁₆Ge₃₀ system can be considered a binary mixture.

This chapter will exemplify the techniques discussed in the methods chapter. It will go more in depth of the fitting procedure. This will allow that the following chapter, with the rest of the binaries, can in higher degree to be focused on the results rather on the nuances of how the CSCE performed on each system.

3.1 Fitting the Ba₈Ga₁₆Ge₃₀ system

The clathrates of type I have 54 atoms in the unit cell, 46 of which belong to the cage structure. When considering the unit cell with boundary conditions there are 36 unique lengths, which means that for a binary mixture there are 37 different clusters if considering singlets and pairs with no cutoff radius. Compared to Si which have 2 atoms in the unit cell and 2 different clusters to consider the clathrates are complicated materials, which is one of the reasons compressive sensing is necessary.

The parameters J_α for the pair interactions can be seen in figure 3.2. The fit consider singlets as well but are not plotted since its value is much larger than that for the pairs. There is a clear decrease in parameter value for increasing radius. In figure 3.6 the clathrate is visualized along with the prominent bonds found for the parameter values. The first nearest neighbours bonds, $r \approx 2.5 \text{ \AA}$ make out a clear cage structure. The lone, noticeable peak, for the parameter value with $r \approx 4.6 \text{ \AA}$ is 4 bonds that lie on the two large sides on the larger tetrakaidecahedron cages. Finally the group of bonds just below $r = 6 \text{ \AA}$ is visualized, these are bonds that crosses a small part of the cage.

In figure 3.4 the maximum contribution for each pair interaction for the energy, volume and band gap are shown. The energy dependence on radius is much more pronounced than on the other two quantities which both decline in roughly the same manner.

The CV score can be seen in fig.3.3. Variant 1 is the method described in Eq.(2.19) and Variant 2 is similar to LOO-CV except for that 30 configurations are picked out and this is done several times to get an average estimator for the root mean square error per atom. Fits based on 55 and 160 configurations in training set, seen in more detail in figure 3.3,t were taken out visualise the spread of the estimated points. Also the CV as a function of the number of clusters considered is shown, the results show that there is an exponential dependence of the CV on the number of included clusters. The number of configurations in the training set was always set to the maximum 200 and with no radius or energy cutoff for the parameters.

The interrelationship between fitting error, norm of the parameters, number of non-zero parameters and the value of μ can be seen in fig. 3.5. The results were obtained with different training set sizes. Shown are the average RMS for the validation energy as a function of μ , the l_1 norm of the parameters as a function of μ and the number of non-zero parameters. It shows, as was explained in the method section, that you have to make a choice between low l_1 norm and non-zero parameters and low fitting errors. There are 37 parameters being fitted and it is shown that when having 20 configurations the $\mu = 0$ solution has very poor predictability. But when increasing the weight of the l_1 minimization one obtains good predictability for this undetermined system. However, when the size becomes overdetermined the l_2 minimization works but cannot decrease the number of ECT's.

3.2 Results

In many of the figures in this section there is an error bar, this is proportional to the standard deviation often it is a quarter or a fifth of the calculated std. This is to give an idea of the fluctuations in the sampling. As explained in the method section the statistical error would be too small to see as an error bar in the figure. In figure 3.7 the ensemble average of the energy during a MC simulation is shown. There is a clear flattening of the curve after 200 K, the MC simulation finds a minimum energy structure. This structure was analyzed and it was identified to belong to space group 146. A DFT calculation on this structure was made. The difference in the energy predicted by the CE and the DFT calculation is of the size 10^{-4} eV.

In figure 3.8 the ensemble average of the band gap during an MC simulation is shown. The band gap shows an increase as the disorder decreases. The calculated band gap and the CE prediction show a relatively larger error than for the energy. This can in part be explained that the DFT data is calculated with few k-samples which is enough for the

energy calculations but both gives an over-estimation of the band gap and adds noise.

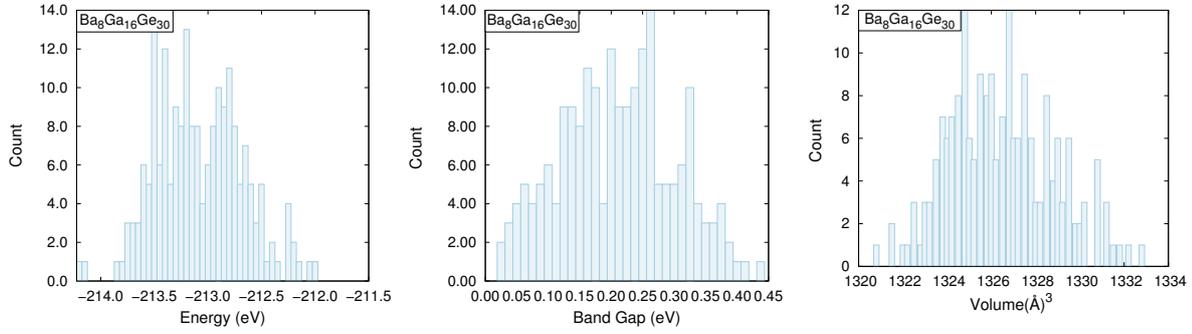


Figure 3.1: The DFT data used for the fitting of the $Ba_8Ga_{16}Ge_{30}$ is presented. In total there are 200 configurations. The distributions is for Left) Energy for the different configurations. Middle) the Band Gap. Right) The volume.

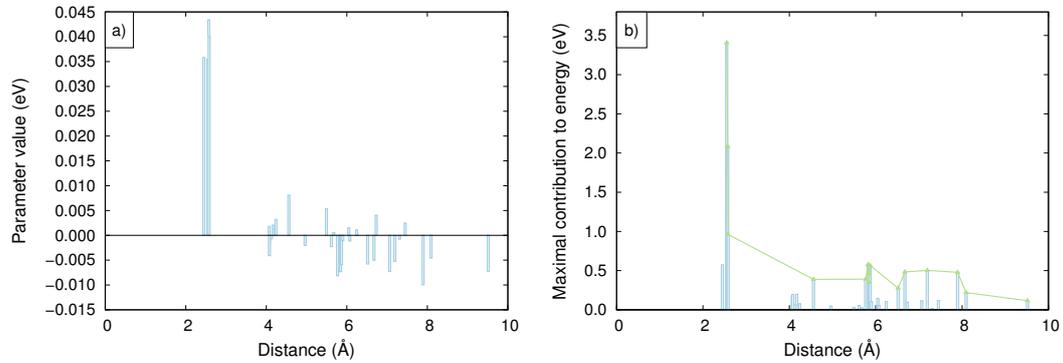


Figure 3.2: a) Parameters obtained from compressive sampling split bregrman for 170 configurations in training. b) Parameters weighted with how many pairs there are for the given distance resulting in the maximum contribution to total energy from that distance. The total energy contribution dies off, as expected, for longer distances

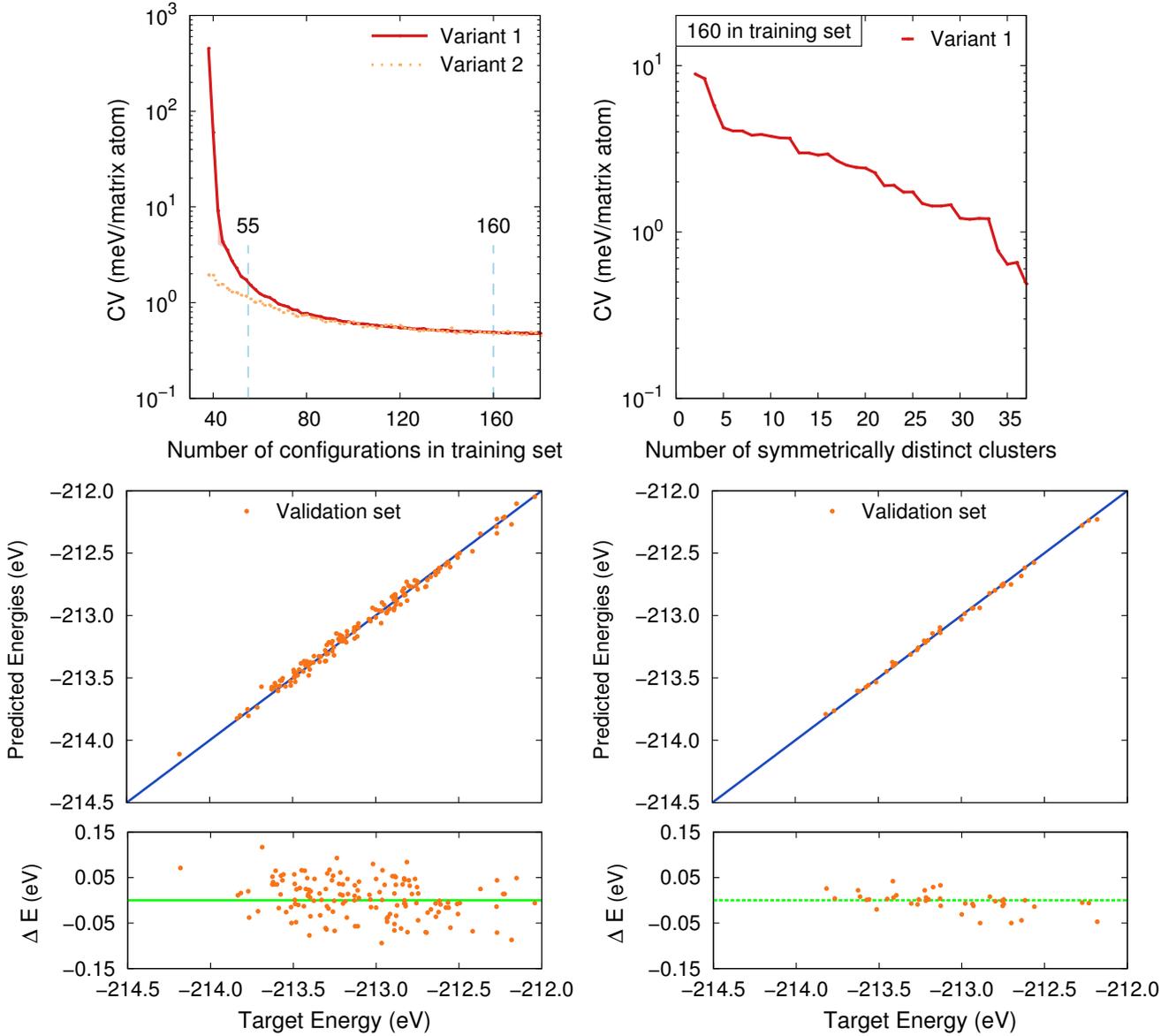


Figure 3.3: The cross validation as a function of number of configurations in the training set. Each CV value averaged over 20 different of training and validation set, where the validation set was fixed at 30 configurations. Top right figure shows the CV value as a function of number of clusters considered in the fit, the clusters was taken out in order of their pair distance. Shown at the bottom are two sets of fits that were taken out at 55 and 160 configurations in the training set.

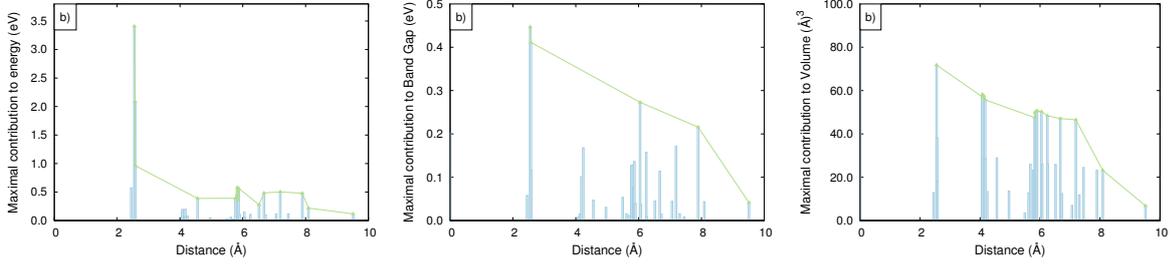


Figure 3.4: Maximum contribution for the different quantities of the $Ba_8Ga_{16}Ge_{30}$ system. The dependence on the pair radius is shown. Left) The energy parameters showing a relatively quick decrease as the distance increases. Middle) The Band Gap shows less decrease with distance compared to energy. Right) The volume parameters also show less decrease with radius compared to energy.

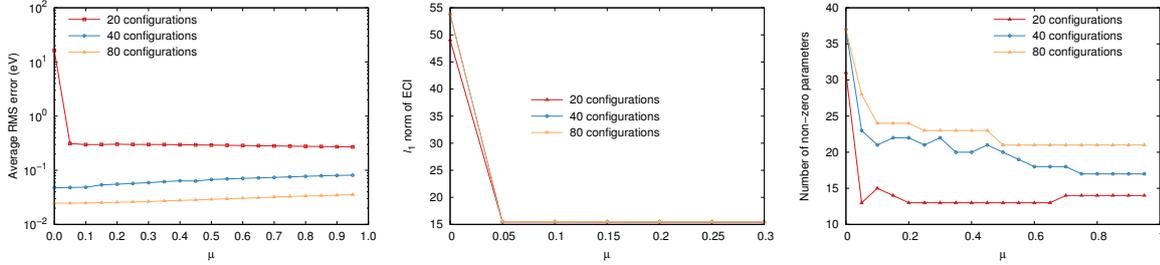


Figure 3.5: Results obtained from the $Ba_8Ga_{16}Ge_{30}$ system for different sizes of the training set. Left) Average RMS for the validation energy as a function of μ . Middle) The l_1 norm of the parameters as a function of μ . Right) The number of non-zero parameters as a function of μ . Where a non-zero parameter is defined as the absolute value should be larger than the half of the mean absolute value of the parameters. The relationship between error of the fit and the norm of the parameters as a function of μ is visualized. A value of $\mu = 0.65$ was chosen. This was considered a fair trade-off between non-zero parameters and RMS error.

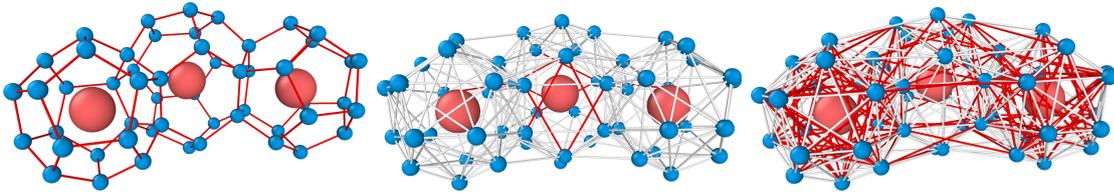


Figure 3.6: Left) The first nearest neighbour bond is shown. The cage structure is visualized. Middle) The red bonds are the noticeable peak at 4.6 Å for the energy parameters. The bonds correspond to four bonds on each of the larger side of the larger cage. Right) The bonds corresponding to the group of distances at 5.8 Å. The bonds cross a part of the cage but not entirely across.

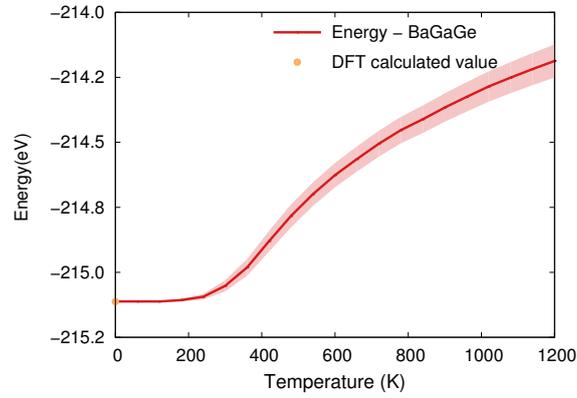


Figure 3.7: The total energy as a function of temperature.

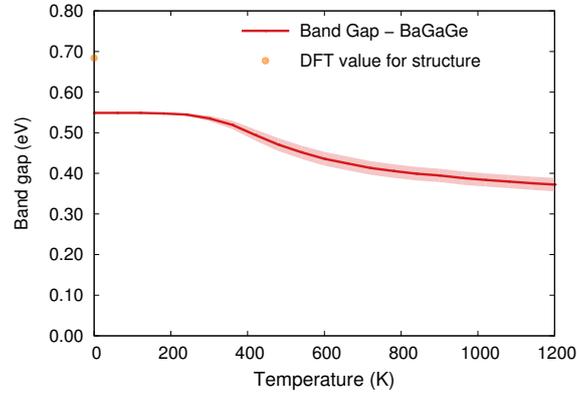


Figure 3.8: The bandgap as a function of temperature. The bandgap was recorded based of a CE during a MC simulation

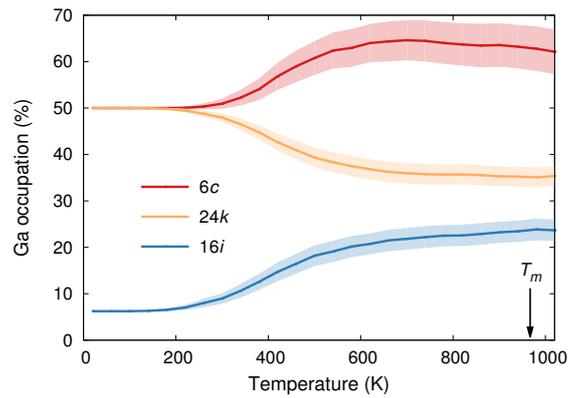


Figure 3.9: The average Ga occupation for different Wyckoff sites as a function of temperature.

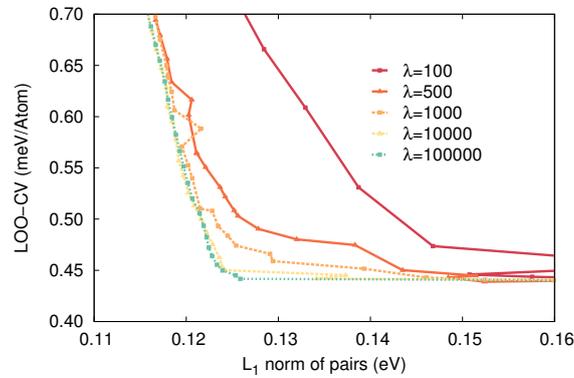


Figure 3.10: LOOCV score vs l_1 norm for pair parameters. By choosing higher λ the l_1 norm can be decreased without compromising the CV score. This leads to a sparser solution which is the goal in compressive sensing and should lead to a better CE. For each line, the λ value is fixed and CV score and the parameters are recorded for different values of μ . The final CE was with $\lambda = 10^6$ and $\mu = 0.706 \cdot 10^{-3}$. This corresponds to the third point from the right for the light green line. Just before the CV starts to increase.

4

Other binaries

With the groundwork laid out in the previous chapter this chapter can generalize the results and compare the different binaries which are, besides the $\text{Ba}_8\text{Ga}_{16}\text{Ge}_{30}$, the $\text{Ba}_8\text{Ga}_{16}\text{Si}_{30}$, $\text{Ba}_8\text{Al}_{16}\text{Si}_{30}$ and $\text{Ba}_8\text{Al}_{16}\text{Ge}_{30}$. The different distributions for the data sets used can be seen in figures 4.1, 4.2 and 4.3. The ground state structure found for the $\text{Ba}_8\text{Ga}_{16}\text{Ge}_{30}$ are added in the data sets for the new binaries. This configuration can easily be spotted in the distribution as the one that stands out. For the energy distribution in Fig.4.2 this configuration has a significantly lower energy compared to the random structures in the other data set. If this configuration is the $T = 0$ K structure for these binaries, and the random configurations represents high temperature (no order) configurations, the band gap will increase for all the binaries as the temperature decreases, the volume will increase for all binaries except for $\text{Ba}_8\text{Ga}_{16}\text{Si}_{30}$ which has an increase in volume for increasing order. In figure 4.1, for the band gap distribution all the binaries except for $\text{Ba}_8\text{Ga}_{16}\text{Ge}_{30}$ have configurations with no bandgap. That means that a fit for the band gap for these materials, with this data set, is highly suspicious.

4.1 Result

In figure 4.4 the result for the average occupation factors along with experimental results from Ref.[7] are visualised. Unfortunately it is not clear how the cooling process was performed making the temperature position of the experimental results unclear. An attempt has been made to manually fit them into the simulation results. All the different materials arrive at the same final result but all in different ways. Most noticeable is the difference whether aluminum or gallium was present. The data was taken with the same number of MC steps but for the binaries where aluminum was present the acceptance probability is up to a factor ten, the acceptance can be seen in figure 4.7 This indicates that there are larger energy barriers in the aluminum-based clathrates compared to the

gallium-based.

$\text{Ba}_8\text{Al}_{16}\text{Si}_{30}$ achieves ordering at the highest temperature at around 300 K, $\text{Ba}_8\text{Ga}_{16}\text{Si}_{30}$ has the lowest temperature required to achieve ordering at about 50 K. The temperature intervals in these simulations was 60 K.

In figure 4.5 the average energies are shown along with the DFT value for the ground state structure found for $\text{Ba}_8\text{Ga}_{16}\text{Ge}_{30}$. Structural analysis on the configurations show the same ground state structure. There is a small but noticeable difference between the final energies from the MC simulation and the DFT value of the lowest energy structure for the silicon-based clathrates. The germanium-based was showed almost exact prediction.

In figure 4.6 the change in lattice parameter as the temperature is lowered is shown, the $\text{Ba}_8\text{Ga}_{16}\text{Si}_{30}$ system behaves completely different, increasing in size for increasing order. The largest increase is between 0 K to 300 K, the temperature interval where the other materials show smallest volume differences. The cluster expansion overshoot the lattice parameter for all materials.

Figure 4.8 shows the band gap dependence on temperature. All materials show an increase in the band gap.

Figure 4.9 with data taken from Ref.[7] show the experimental Al occupation of the wyckoff sites alongside the calculated sof's. There is an apparent phase transition in the occupatiencis between 14 and 15 Al atoms in the unit cell in the experimental data. Shown also in the figure is the sof's for $\text{Ba}_8\text{Al}_{16}\text{Ge}_{30}$ with one vacancy/unit cell. The vacancy is modeled by defining that all clusters that involve a vacancy has zero energy. This is a very simplified approach to consider vacancies. A more rigorous approach would have been to create a CE based on DFT data with vacancies. From there the system can either be modeled as binary or as a ternary and including the vacancy in the cluster expansion. Still, some observations can be made. The 24k and 16i sites are largely unaffected by the vacancy. The 6c occupancy decreased with vacancy in this model. To study stoichiometry more rigorous, two DFT sets were done for The Ga occupation in $\text{Ba}_8\text{Ga}_x\text{Ge}_{46-x}$ for $x = 15$ and $x = 17$. They are charge compensated with $q = 1$ for Ge excess and $q = -1$ for Ga excess. The sof's compared with sof's from $\text{Ba}_8\text{Ga}_{16}\text{Ge}_{30}$ CE is shown for different stoichiometries in figure 4.10. The same approach with vacancies are applied to these non-stoichiometric systems. The same behaviour is seen here with little or no effect on 24k and 16i and a clear decrease in 6c occupation.

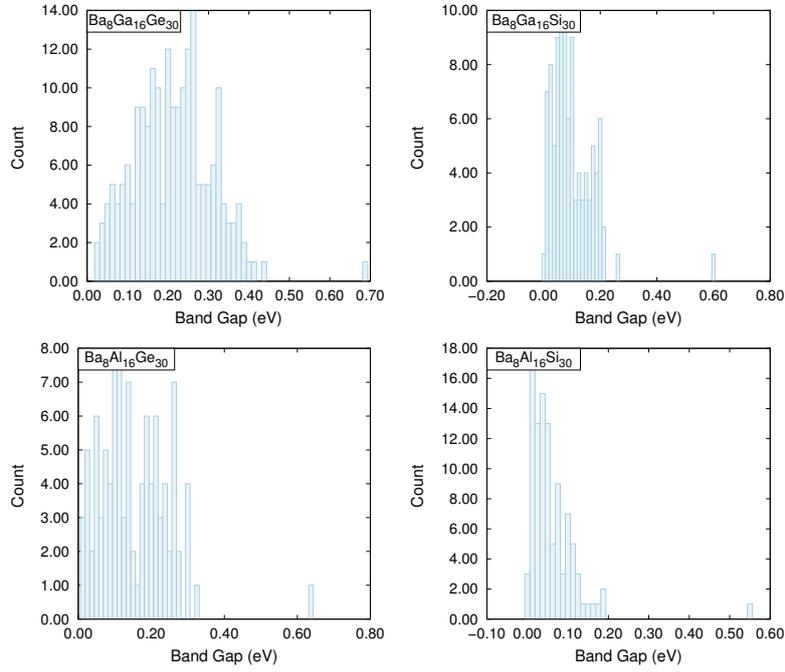


Figure 4.1: Distribution for the Band Gap over the different data sets.

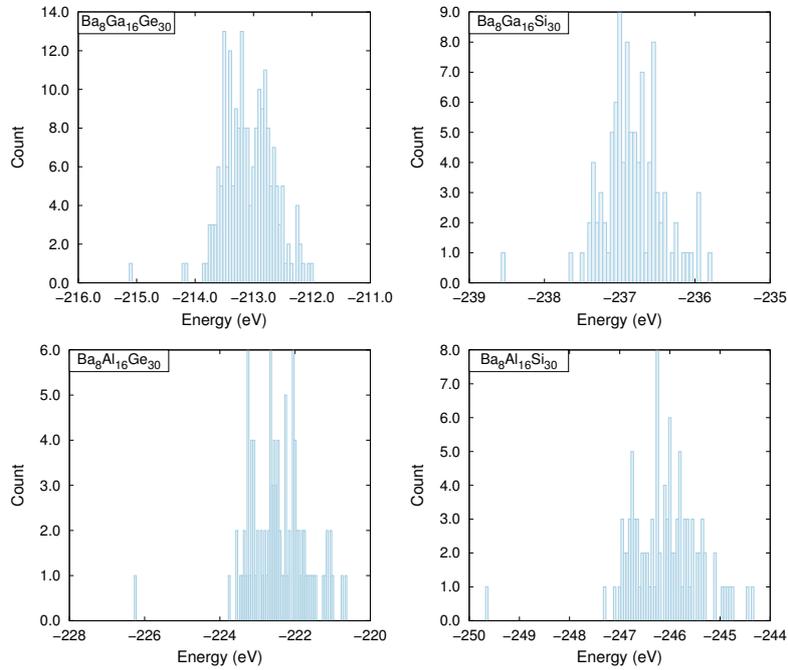


Figure 4.2: Distribution for the Energy over the different data sets.

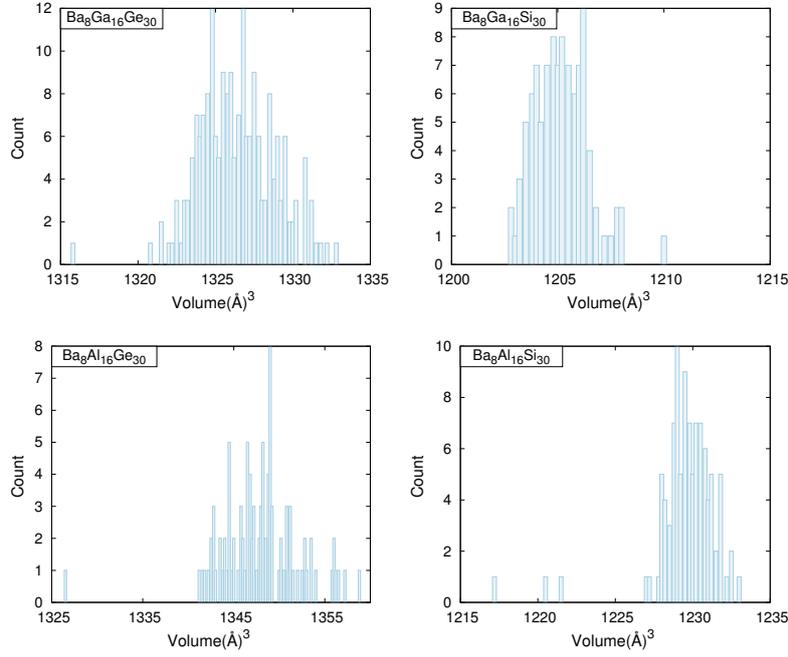


Figure 4.3: Distribution for the Volume over the different data sets.

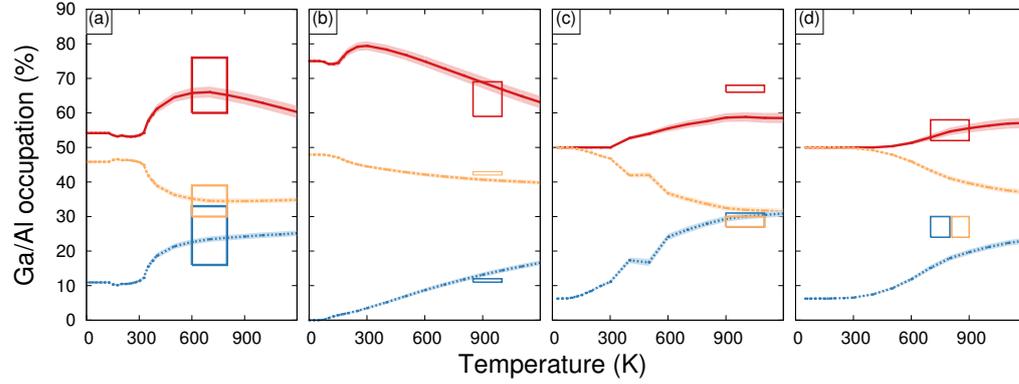


Figure 4.4: a) $\text{Ba}_8\text{Ga}_{16}\text{Ge}_{30}$, b) $\text{Ba}_8\text{Ga}_{16}\text{Si}_{30}$, c) $\text{Ba}_8\text{Al}_{16}\text{Ge}_{30}$, d) $\text{Ba}_8\text{Al}_{16}\text{Si}_{30}$ The average occupation for the different binaries. The boxes indicate the experimental results taken from the table in Ref.[7]. The height of the boxes correspond to experimental errors and the fact that different synthesis methods leads to different results. It is not clear how the cooling process looked like making the temperature position of the boxes unclear.

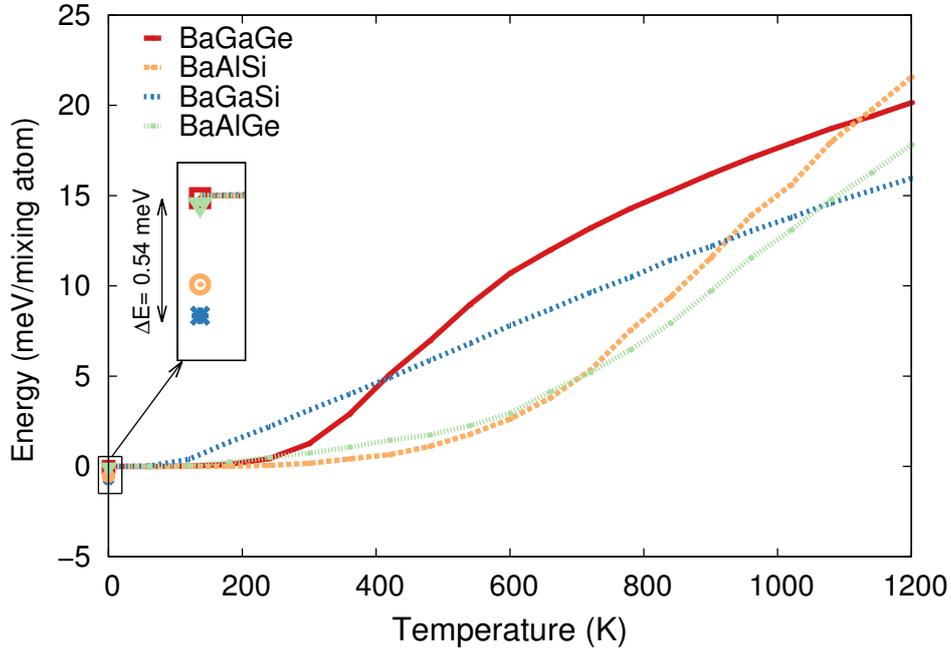


Figure 4.5: The ensemble average for the energy for the different binaries. The cluster expansion prediction for the energy of the ground state structure is very good. The energies have their respective ground state energy subtracted.

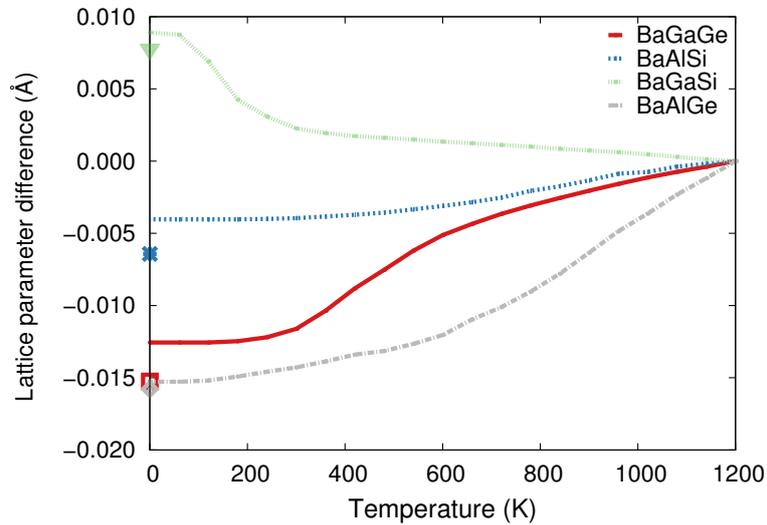


Figure 4.6: The ensemble average for the lattice parameter relative to lattice parameter average for $T = 1200$ for the different binaries. The points represent the expected value as calculated by DFT

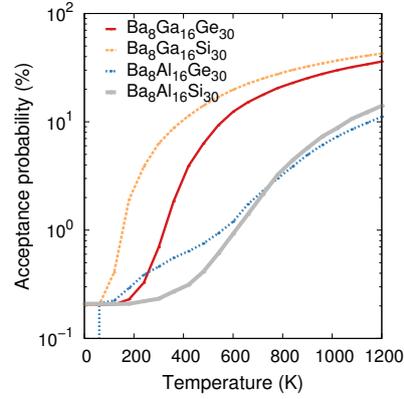


Figure 4.7: The average acceptance probability for the different binaries. The aluminum based clathrates have a much lower value, indicating larger energy barriers associated with trial moves.

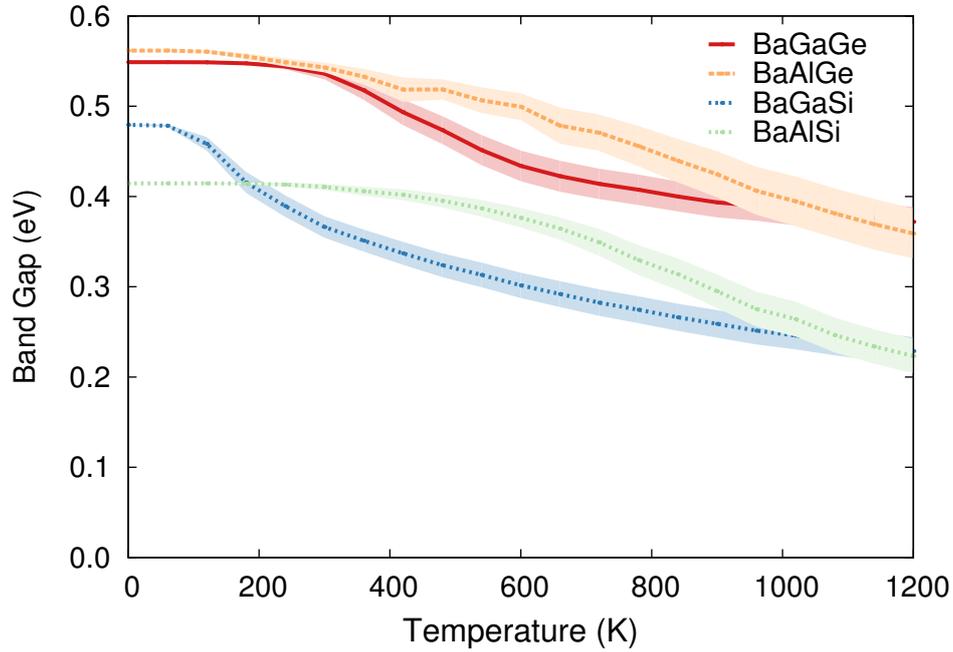


Figure 4.8: Ensemble averages for the Band Gap for different binaries as a function of temperature.

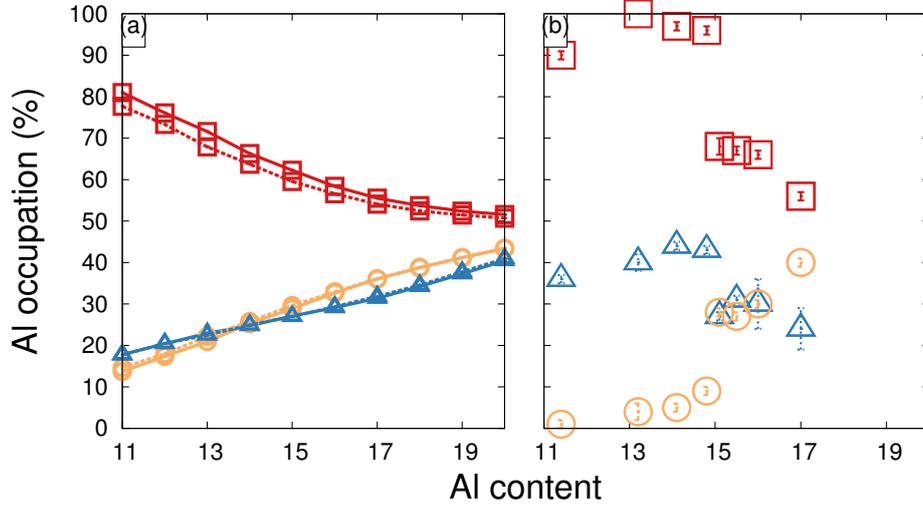


Figure 4.9: a) Al occupation as a function of Al content/unit cell. b) Data from Ref. [7]. There is a distinct phase boundary in the experimental data between 14 and 15 Al atoms in the unit cell.

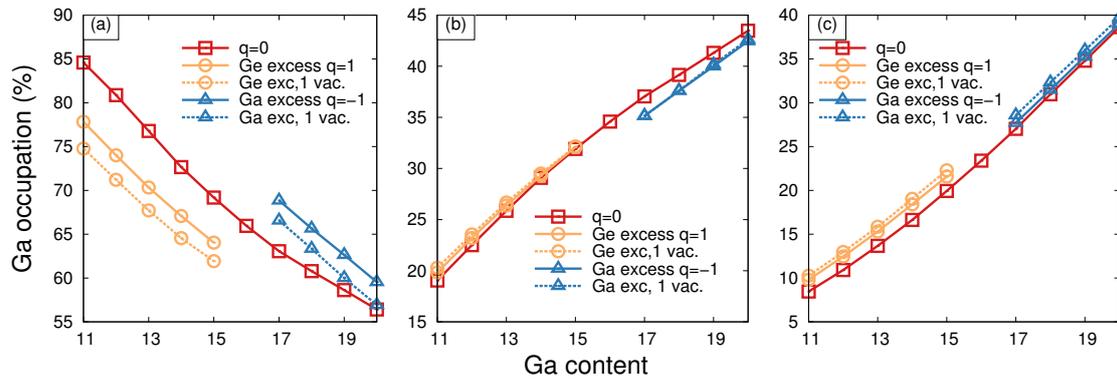


Figure 4.10: Ga occupation on a) 6c site, b) 24k and c) 16i for $\text{Ba}_8\text{Ga}_x\text{Ge}_{46-x}$ as a dependence on stoichiometry for the different cluster expansions. The different cluster expansions are explained in the text.

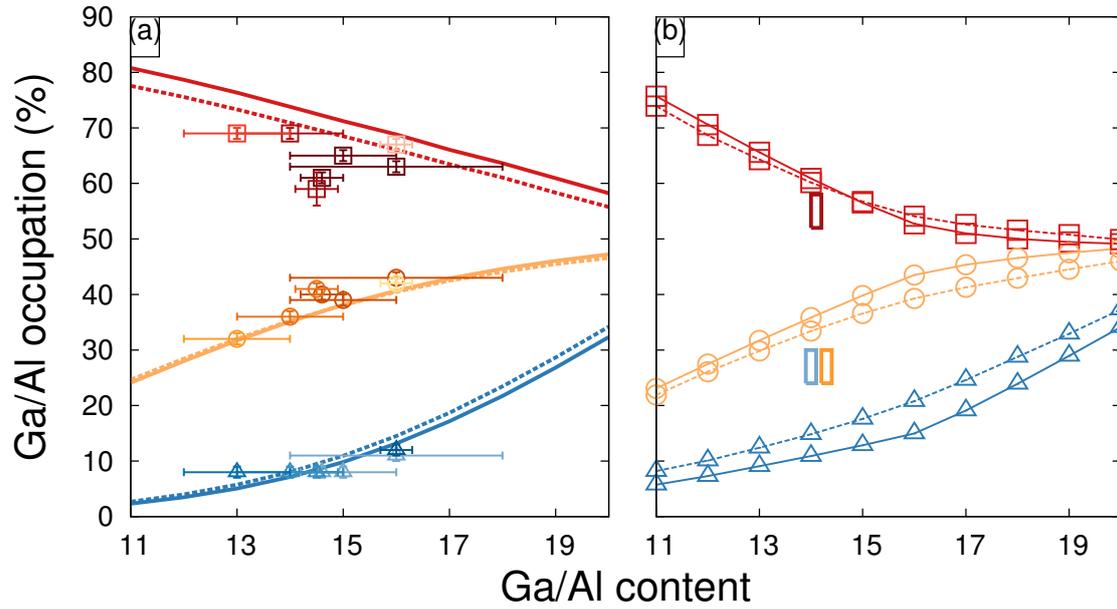


Figure 4.11: a) Ga occupation for $\text{Ba}_8\text{Ga}_{16}\text{Si}_{30}$. The lines are the occupation found from MC simulation at 900K the dotted lines are with one vacancy/unit cell. The point data are experimental values from Ref. [7]. There is an excellent agreement with experiment for 24k and 16i. 6c data is overestimated slightly. b) Al occupation for $\text{Ba}_8\text{Al}_{16}\text{Si}_{30}$. Only one experimental point available. The stoichiometry for the experimental sample is $\text{Ba}_8\text{Al}_{14.1}\text{Si}_{30.8}$ suggesting one vacancy/unit cell. The dotted lines are with one vacancy/unit cell

5

Quaternaries

Two quaternary clathrate data sets have been generated; $\text{Ba}_8\text{Al}_{16-x}\text{Ga}_x\text{Si}_{30}$ and $\text{Ba}_8\text{Al}_{16-x}\text{Ga}_x\text{Ge}_{30}$. Figure 5.1 show the $\text{Ba}_8\text{Al}_{16-x}\text{Ga}_x\text{Si}_{30}$ and $\text{Ba}_8\text{Al}_{16}\text{Ge}_{30}$ leave one out fit. The $\text{Ba}_8\text{Al}_{16-x}\text{Ga}_x\text{Si}_{30}$ cluster expansion is based on a dataset consisting of $\text{Ba}_8\text{Ga}_{16}\text{Si}_{30}$ and $\text{Ba}_8\text{Al}_{16}\text{Si}_{30}$ dataset from the quasi-binaries discussed in the previous chapter as well as their ground state structure. It also consists of a $\text{Ba}_8\text{Al}_{16-x}\text{Ga}_x\text{Ge}_{30}$ with $x = 4, 8, 12$ with 100 configurations for each stoichiometry. The $\text{Ba}_8\text{Al}_{16-x}\text{Ga}_x\text{Ge}_{30}$ cluster expansion is based on the $\text{Ba}_8\text{Ga}_{16}\text{Ge}_{30}$ and $\text{Ba}_8\text{Al}_{16}\text{Ge}_{30}$ datasets from the previous chapter as well as a set of random $\text{Ba}_8\text{Al}_{16-x}\text{Ga}_x\text{Ge}_{30}$ structures with x between 1 and 15 uniformly distributed.

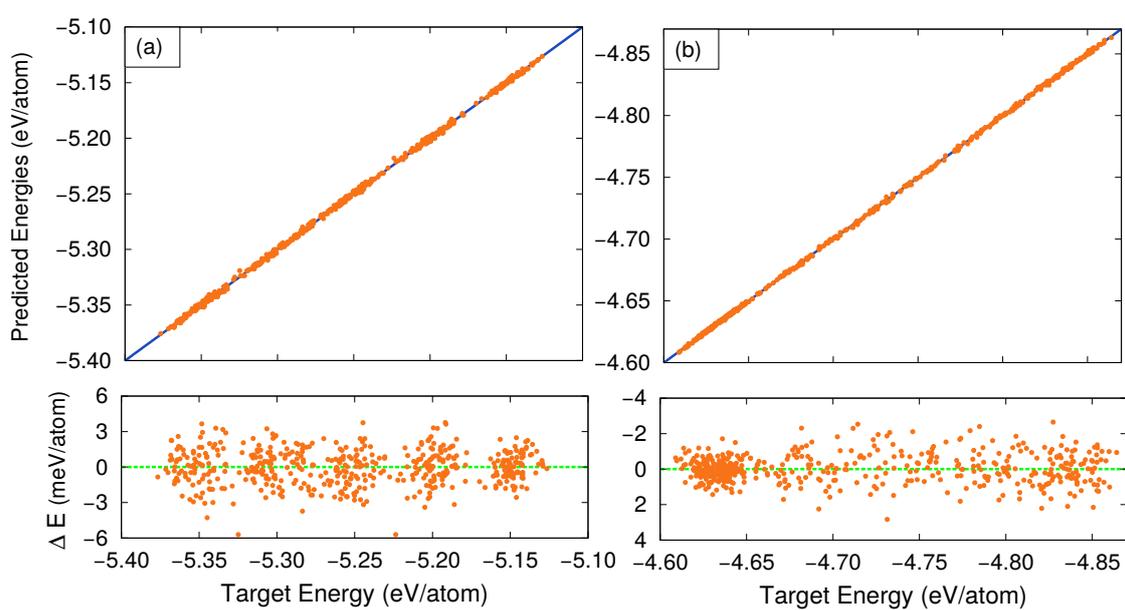


Figure 5.1: Leave one out fit. That is each point is a prediction from the cluster expansion that is found when all configurations are used except for the one configurations that is predicted. a) $\text{Ba}_8\text{Al}_{16-x}\text{Ga}_x\text{Si}_{30}$ with a LOO-CV of 1.5 meV/atom, b) $\text{Ba}_8\text{Al}_{16-x}\text{Ga}_x\text{Ge}_{30}$ with LOO-CV of 0.8 meV/atom

6

Discussion

The main results in this thesis is the site occupancy factors. They are a bridge between the real world experiments and the computational model. If the computational model emulate the experimental results it is reasonable to extrapolate the model beyond what is possible by experiments. This is, after all, the point, to be able to predict new and interesting behavior of a material by scanning through the available parameters. It is one of the most powerful aspects of computational models.

6.1 Site occupancy factors

To be able to establish whether the site occupancy factors is in agreement with experimental results or not one needs to understand how the experiments are done and why they can differ so much between experiments.

6.1.1 $\text{Ba}_8\text{Ga}_{16}\text{Ge}_{30}$

The experimental results in figure 4.4, derived from table 1 in Ref. [7], is a compilation from different works. For $\text{Ba}_8\text{Ga}_{16}\text{Ge}_{30}$ the data is a work by Iversen et al [21] where n- and p-type $\text{Ba}_8\text{Ga}_{16}\text{Ge}_{30}$ is studied. They used different models to obtain framework occupancies from neutron and resonant synchrotron data. These models differ by what constraints they set on the crystal and is what leads to the different values for the occupancies and stoichiometries for $\text{Ba}_8\text{Ga}_{16}\text{Ge}_{30}$. Only one model do not assume (16:30) ratio. For that model, the p-type ratio is (16.1:28.9) with a standard deviation of (0.14:0.9). The n-type has a similar ratio of (15.6:29.1) with standard deviation of (0.12:0.7). The 6c Ga occupation is $63\% \pm 5\%$ and $72\% \pm 4\%$ for p-type and n-type respectively. The other models also show the higher 6c occupation for n-type. In figure 4.10, the sof for varying Ga content is shown. The figure there show that, from a p-type with ratio of 17:29 to a n-type with ratio 15:31, the 6c occupation is decreasing. Nothing

is necessarily wrong with these different results. There are a few observations that can be made.

- The stoichiometry for the experimental samples suggests vacancies. The extremes of the standard deviations give a minimum of 0.4 and 0 vacancies per unit cell for n- and p-type respectively and maximum of 2.1 and 2.14 for n-type and p-type.
- The p- and n-type materials each have different synthesis methods. This, along with that each material are physically different and have different energies associated with the kinetics of the atoms can lead to that each material can be in different stages of thermal relaxation.
- The Ga content for both samples are closer to 16 than any of the p-and n-type CE's that were created. For p-type the experimental Ga concentration is about 0.36 which corresponds to about a Ga content of 16.4 with no vacancies. For n-type the concentration corresponds to Ga content of 16.
- The CE with 16:30 Zintl stoichiometry matches qualitatively well with the experimental values.
- The $\text{Ba}_8\text{Ga}_{16}\text{Si}_{30}$ system have more experimental data for sofs for different Ga contents. These show a decrease in 6c occupation as Ga decreases compared to the calculated values seen in figure 4.11 and reaches a plateau. The 16:30 CE for $\text{Ba}_8\text{Ga}_{16}\text{Si}_{30}$ and $\text{Ba}_8\text{Ga}_{16}\text{Ge}_{30}$ give similar sof dependences of Ga content.

All these observations together could explain the discrepancy in 6c Ga occupation. The experimental samples are likely to have vacancies. The samples come from different synthesis methods and possibly different thermal histories. The trend of underestimating Ga 6c occupation for decreasing Ga content is seen experimentally in the $\text{Ba}_8\text{Ga}_{16}\text{Si}_{30}$ system. Meaning that there is a decreasing 6c occupation trend for $\text{Ba}_8\text{Ga}_{16}\text{Ge}_{30}$ without defects.

The Ga occupation on 16i and 24k show no significant difference between the two samples which is also what is observed in figure 4.10. This can also be seen in the $\text{Ba}_8\text{Ga}_{16}\text{Si}_{30}$ system where the 16i and 24k occupation predictions are excellent for different stoichiometries.

An interesting observation is that all the n-type models show a combined 6c+24k occupation of about 110%. A value which, in this thesis, rarely goes below 100%. Since it is the high 6c occupation of the n-type that is hard to reproduce there can be a connection between breaking the occupancy rules (see below) and poor reproduction from the calculations.

6.1.2 $\text{Ba}_8\text{Ga}_{16}\text{Si}_{30}$

The compiled data for $\text{Ba}_8\text{Ga}_{16}\text{Si}_{30}$ are from six different publications. Two of these report a 16:30 ratio:

- Ref. [22], Nataraj *et al* 2004 used X-ray diffraction and Raman scattering measurements to study $\text{Ba}_8\text{Ga}_x\text{Si}_{46-x}$ with Ga content of 10 and 16 synthesized by melting elements in an argon atmosphere by an Ar ion gun.
- Ref. [23], Palmqvist *et al* 2002, studied $\text{Ba}_8\text{Ga}_{16}\text{Si}_{30}$ using multitemperature single crystal neutron diffraction. Synthesis method is flux growth [24]

There is no significant difference in the sof's between the different 16:30 samples. The largest difference is for the 6c site; Palmqvist *et al* reported 63(1)% where Nataraj *et al* found a 67(1)% Ga occupancy for the 6c site. The similarities in the 24k sof and difference in 6c sof can be explained by the flat temperature dependence of Ga occupation on 24k and steep dependence for 6c in figure 4.4. The other reported site occupancy factors for non stoichiometric $\text{Ba}_8\text{Ga}_{16}\text{Si}_{30}$ show a depletion of the 16i and 24k Ga occupation and no significant difference between the 6c site. The 24k and 16i occupation show excellent reproduction. The 6c occupation is overestimated. However the general trend that there are no great Ga content dependence on 6c occupation is reproduced with a relatively flat Ga occupation dependence on Ga content.

The calculated values for $\text{Ba}_8\text{Ga}_{16}\text{Si}_{30}$ with 16:30 stoichiometry match very well with the experimental 16:30 values and can even explain that the different experiments showed a discrepancy in 6c occupation and not in the other sites, a possible explanation is different cooling temperatures between the experiments and thus ending up with different sof's according to figure 4.4. Furthermore the behaviour for non stoichiometric $\text{Ba}_8\text{Ga}_{16}\text{Si}_{30}$ is also reproduced. Worth noticing is that none of the $\text{Ba}_8\text{Ga}_{16}\text{Si}_{30}$ samples reported any vacancies in the framework which would make it more difficult reproducing the results. Since 24k and 16i were spot on, it reasons that if there were vacancies in the experimental samples it is the 6c occupation that is most affected. This is also what is seen for the vacancy model. The 6c occupation decreases slightly without affecting the other sites much. Only a few of the experimental points lies outside the calculated values.

6.1.3 $\text{Ba}_8\text{Al}_{16}\text{Ge}_{30}$

The values for $\text{Ba}_8\text{Al}_{16}\text{Ge}_{30}$ is from Ref. [8] where Christensen *et al* 2007 prepare three $\text{Ba}_8\text{Al}_x\text{Ge}_{46-x}$ samples using three different synthesis methods, flux growth, Czochralski growth and conventional stoichiometric mixing. The prediction works very well close to 16:30 stoichiometry. The experimental data show a transition at 14-15 Al content leading to full 6c occupation and depletion of 24k similar to the occupation in $\text{Ba}_8\text{Ga}_{16}\text{Si}_{30}$ but with depletion of 16i. This behaviour is not reproduced in the calculations of this thesis. The simple introduction to vacancies lowers the already low prediction of 6c occupation. This shows that the $\text{Ba}_8\text{Al}_x\text{Ge}_{46-x}$ have a complex dependency on Al content. A hint of this could be by looking at the combined 6c+24k and 16i+24k occupancy. For $\text{Ba}_8\text{Ga}_x\text{Si}_{46-x}$, which reproduces it's experiments well, has a spread of 8 units of percent for 6c+24k and 18 units of percent for 16i+24k. $\text{Ba}_8\text{Al}_x\text{Ge}_{46-x}$ have 14 and 27 units of percent respectively. $\text{Ba}_8\text{Al}_x\text{Ge}_{46-x}$ is breaking the sum occupation rule of $16i+24k \leq 50\%$. Since these sum of occupation rules arise from avoiding direct III-III bonds, the

breaking of these, and also in what degree it is broken, tells of the energetics associated with the bonds. The argument is that since the $\text{Ba}_8\text{Al}_x\text{Ge}_{46-x}$ system has larger spread of the sum of occupation, and also is breaking one summation rule, the bond energies changes more for different stoichiometries than compared to $\text{Ba}_8\text{Ga}_x\text{Si}_{46-x}$. The CE for $\text{Ba}_8\text{Al}_x\text{Ge}_{46-x}$ is with 16:30 stoichiometry. From the MC run the 24k+16i occupation is 63%. This is a possible reason of why the experimental results are poorly reproduced. The general trends of increasing 6c and decreasing 24k and 16i occupation is the same. Noteworthy is also the fact that 24k occupation goes below 16i, same as for the experimental data.

6.1.4 $\text{Ba}_8\text{Al}_{16}\text{Si}_{30}$

There is only one experimental sample on sof's for $\text{Ba}_8\text{Al}_{16}\text{Si}_{30}$. One reason is of the experimental difficulties of differentiate Al and Si in X-ray diffraction [25]. The one point available is for $\text{Ba}_8\text{Al}_{14.1}\text{Si}_{30.8}$ suggesting one vacancy/unit cell. The calculated sof's for the fully occupied $\text{Ba}_8\text{Al}_x\text{Si}_{46-x}$ with 14 Al content overestimate the 6c and 24k occupation, the former just slightly however. The 16i occupation is underestimated. The sof's for the one vacancy system is much closer to the experimental values however. For $\text{Ba}_8\text{Al}_x\text{Si}_{46-x}$ the vacancy system is the opposite of the binaries. The 6c is roughly the same but the 24k has a noticable decrease while 16i occupancy increases, making the prediction of the experimental value much better.

6.1.5 Rules for occupancy

Christensen *et al* 2009 [7] established rules for the the maximum trivalent element site occupancy factor (sof):

- **Rule 1: 6c sof(III) \leq 100%**
There are no direct bonds between two 6c sites and 100% occupancy of trivalent elements can be allowed at the 6c site without causing unfavorable bonds
- **Rule 2: 16i sof(III) \leq 50%**
The 16i site binds to three 24k atoms, and one 16i atom. Therefore half of the 16i sites can be occupied by trivalent elements without forming disfavored trivalent-trivalent bonds.
- **Rule 3: 24k sof(III) \leq 50%**
The 24k site binds to one 6c site, two 16i sites and one 24k site. Therefore only up to half of the 24k sites can be simultaneously occupied by trivalent elements.
- **Rule 4: 6c+24k sof(III) \leq 100%**
The 6c site binds to four 24k sites. The sum of the trivalent occupancies at the 6c and 24k site should not exceed 100%
- **Rule 5: 16i+24k sof(III) \leq 50%**
The 16i atoms bind to three 24k atoms. Thus at a specific 16i site a maximum of one

of the four atoms ($16i + 3 \times 24k$) can be trivalent elements, that is 25% occupancy but each 24k atom binds to two 16i sites doubling the allowed combined occupancy.

6.1.6 General conclusions for all binaries

$\text{Ba}_8\text{Ga}_{16}\text{Ge}_{30}$, $\text{Ba}_8\text{Ga}_{16}\text{Si}_{30}$ and $\text{Ba}_8\text{Al}_{16}\text{Ge}_{30}$ all match very well with the 16:30 experimental results. It can also be said that by varying the stoichiometry to match the experimental values, predictions for all binaries got better than their 16:30 prediction. By introducing vacancies $\text{Ba}_8\text{Ga}_x\text{Si}_{46-x}$ and $\text{Ba}_8\text{Al}_x\text{Si}_{46-x}$ got closer to their experimental counterparts. $\text{Ba}_8\text{Al}_x\text{Ge}_{46-x}$ did not get excellent experimental agreement. The phase transition is not picked up on. However general trends were reproduced.

6.2 Outlook

This thesis has laid a groundwork on which much can be built upon. The methodology used here is in no way restricted to clathrates. Clathrates, as many other promising thermoelectric materials, have complicated structures. The fact that the cluster expansions, used in conjunction with compressive sensing, works so well with the clathrates show promise to be successful for other thermoelectric materials. Skutterudites [26], Zintl phases [27], Zn_4Sb_3 with interstitial, disordered zinc positions [28] are materials with complex structures where the order or disorder in the material can be crucial in understanding and designing thermoelectric materials. The approach taken in this thesis allows to investigate these systems and related ones. To the authors knowledge, there are no other work in the literature where the ordering of clathrates or other complex thermoelectric materials have been studied with cluster expansions and Monte Carlo simulations. These facts makes it meaningful to continue pursuing and publishing these studies. It will be interesting for experimentalists who only sample small subsets of the materials to take part of such results, allowing them to easier focus in on the interesting temperature- and stoichiometry areas and the most promising chemical constituents of the material.

The analysis can also further be developed. The Monte Carlo simulations can be used to pick out a few representattive structures for different temperatures. These structures can then be analysed to calculate the electrical conductivity, Seebeck coefficient, lattice thermal conductivity [29][30]. This effectively allows one to map out ZT for different concentrations, materials and temperatures. The full completion of such a work will be a huge accomplishment toward creating powerful thermoelectric materials.

7

Conclusion

The key results is that the order dependence of band gap and lattice parameter is shown and quantified. The findings show that Band Gap is lowered with increasing disorder this has a significant effect of electric conductivity which has a exponential dependence on the band gap.

The lattice parameter for type-I clathrates not only depends on phonon-phonon interaction but also on ordering. $\text{Ba}_8\text{Ga}_{16}\text{Si}_{30}$ show an increase in size while the other binaries show a decrease.

The site occupancy factors match experimental results very well for 16:30 stoichiometry. A good agreement is also seen in non-stoichiometric $\text{Ba}_8\text{Ga}_x\text{Si}_{46-x}$ and to some extent $\text{Ba}_8\text{Al}_x\text{Si}_{46-x}$. Simple vacancy models helps in understanding the experimental results and can improve the prediction for the sof's. The $\text{Ba}_8\text{Al}_x\text{Ge}_{46-x}$ system is reproduced quantitatively except for the phase transition which the CE based on 16:30 mixing cannot reproduce.

The cluster expansion with the compressive sensing algorithm proved to be very effective in terms of energy fitting. The ground structure found in the Monte Carlo simulation was very accurately predicted by the CE. This ground state structure correspond to space group 146 and can serve as starting points for further theoretical investigations.

Bibliography

- [1] D. Rowe, CRC Handbook of Thermoelectrics, CRC Press, 1995.
URL <http://www.crcnetbase.com/isbn/978-0-8493-0146-9>
- [2] M. S. Dresselhaus, G. Chen, M. Y. Tang, R. G. Yang, H. Lee, D. Z. Wang, Z. F. Ren, J.-P. Fleurial, P. Gogna, New Directions for Low-Dimensional Thermoelectric Materials, *Advanced Materials (FRG)* 19 (8) (2007) 1043–1053.
URL <http://search.proquest.com/docview/29897209?pq-origsite=summon>
- [3] S. B. S. Glen A. Slack, Chapter 6 Semiconductor clathrates: A phonon glass electron crystal material with potential for thermoelectric applications, *Semiconductors and Semimetals - SEMICOND SEMIMET* 69 (2001) 255–300.
- [4] J.-W. G. Bos, R. A. Downie, Half-Heusler thermoelectrics: a complex class of materials, *Journal of Physics: Condensed Matter* 26 (43) (2014) 433201.
- [5] K. Koumoto, I. Terasaki, R. Funahashi, Complex Oxide Materials for Potential Thermoelectric Applications, *MRS Bulletin* 31 (03) (2006) 206–210.
URL http://journals.cambridge.org/article_S0883769400009933
- [6] A. V. Shevelkov, K. Kovnir, Zintl Clathrates, in: T. F. Fässler (Ed.), *Zintl Phases*, no. 139 in *Structure and Bonding*, Springer Berlin Heidelberg, 2011, pp. 97–142.
URL http://link.springer.com/chapter/10.1007/430_2010_25
- [7] M. Christensen, S. Johnsen, B. B. Iversen, Thermoelectric clathrates of type I, *Dalton Trans.* 39 (4) (2010) 978–992.
URL <http://pubs.rsc.org/en/content/articlelanding/2010/dt/b916400f>
- [8] M. Christensen, B. B. Iversen, Host Structure Engineering in Thermoelectric Clathrates, *Chem. Mater.* 19 (20) (2007) 4896–4905.
URL <http://dx.doi.org/10.1021/cm071435p>
- [9] J. M. Sanchez, F. Ducastelle, D. Gratias, Generalized cluster description of multicomponent systems, *Physica A: Statistical Mechanics and its Applications*

- 128 (1–2) (1984) 334–350.
URL <http://www.sciencedirect.com/science/article/pii/S0378437184900967>
- [10] L. J. Nelson, G. L. W. Hart, F. Zhou, V. Ozoliš, Compressive sensing as a paradigm for building physics models, *Phys. Rev. B* 87 (3) (2013) 035125.
URL <http://link.aps.org/doi/10.1103/PhysRevB.87.035125>
- [11] C. Kittel, *Introduction to solid state physics*, 8th Edition, Wiley, Hoboken, N.J, 2005.
- [12] A. van de Walle, Multicomponent multisublattice alloys, nonconfigurational entropy and other additions to the Alloy Theoretic Automated Toolkit, *Calphad* 33 (2) (2009) 266–278.
URL <http://www.sciencedirect.com/science/article/pii/S0364591608001314>
- [13] P. Erhart, Dft data set - unpublished work, . . (2014) .
- [14] L. J. Nelson, V. Ozoliš, C. S. Reese, F. Zhou, G. L. W. Hart, Cluster expansion made easy with Bayesian compressive sensing, *Phys. Rev. B* 88 (15) (2013) 155105.
URL <http://link.aps.org/doi/10.1103/PhysRevB.88.155105>
- [15] E. Candes, M. Wakin, An Introduction To Compressive Sampling, *IEEE Signal Processing Magazine* 25 (2) (2008) 21–30.
- [16] T. Goldstein, S. Osher, The Split Bregman Method for L1-Regularized Problems, *SIAM Journal on Imaging Sciences* 2 (2) (2009) 21.
URL <http://search.proquest.com/docview/925335836/abstract?accountid=10041>
- [17] W. Yin, S. Osher, D. Goldfarb, J. Darbon, Bregman Iterative Algorithms for ℓ_1 -Minimization with Applications to Compressed Sensing, *SIAM Journal on Imaging Sciences* 1 (1) (2008) 26.
URL <http://search.proquest.com/docview/925542801/abstract?accountid=10041>
- [18] A. v. d. Walle, G. Ceder, Automating first-principles phase diagram calculations, *JPE* 23 (4) (2002) 348–359.
URL <http://link.springer.com/article/10.1361/105497102770331596>
- [19] M. Stone, Cross-Validatory Choice and Assessment of Statistical Predictions, *Journal of the Royal Statistical Society. Series B (Methodological)* 36 (2) (1974) 111–147.
URL <http://dx.doi.org/10.2307/2984809>
- [20] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. Teller, Equation of State Calculations by Fast Computing Machines, *The Journal of Chemical*

- Physics 21 (6) (1953) 1087–1092.
URL <http://scitation.aip.org/content/aip/journal/jcp/21/6/10.1063/1.1699114>
- [21] M. Christensen, N. Lock, J. Overgaard, B. B. Iversen, Crystal Structures of Thermoelectric n- and p-type Ba₈Ga₁₆Ge₃₀ Studied by Single Crystal, Multitemperature, Neutron Diffraction, Conventional X-ray Diffraction and Resonant Synchrotron X-ray Diffraction, *J. Am. Chem. Soc.* 128 (49) (2006) 15657–15665.
URL <http://dx.doi.org/10.1021/ja063695y>
- [22] D. Nataraj, J. Nagao, Structure and Raman scattering study on Ba₈GaxSi₄₆x (x=10 and 16) type I clathrates, *Journal of Solid State Chemistry* 177 (6) (2004) 1905–1911.
URL <http://www.sciencedirect.com/science/article/pii/S0022459604000398>
- [23] A. Bentien, B. B. Iversen, J. D. Bryan, G. D. Stucky, A. E. C. Palmqvist, A. J. Schultz, R. W. Henning, Maximum entropy method analysis of thermal motion and disorder in thermoelectric clathrate Ba₈Ga₁₆Si₃₀, *Journal of Applied Physics* 91 (9) (2002) 5694–5699.
URL <http://scitation.aip.org/content/aip/journal/jap/91/9/10.1063/1.1466531>
- [24] A. Bentien, A. Palmqvist, J. Bryan, S. Lattturner, G. Stucky, L. Furenlid, B. Iversen, Experimental Charge Densities of Semiconducting Cage Structures Containing Alkaline Earth Guest Atoms, *Angewandte Chemie International Edition* 39 (20) (2000) 3613–3616.
URL [http://onlinelibrary.wiley.com/doi/10.1002/1521-3773\(20001016\)39:20<3613::AID-ANIE3613>3.0.CO;2-D/abstract](http://onlinelibrary.wiley.com/doi/10.1002/1521-3773(20001016)39:20<3613::AID-ANIE3613>3.0.CO;2-D/abstract)
- [25] S. Johnsen, A. Bentien, G. K. H. Madsen, B. B. Iversen, M. Nygren, Crystal Structure, Band Structure, and Physical Properties of Ba₈Cu_{6-x}Ge_{40+x} (0 < x < 0.7), *Chem. Mater.* 18 (19) (2006) 4633–4642.
URL <http://dx.doi.org/10.1021/cm061195y>
- [26] J. Dyck, W. Chen, J. Yang, G. Meisner, C. Uher, Transport and magnetic properties of Co_{1-x}Ni_xSb₃ with x less than 0.01, in: XX International Conference on Thermoelectrics, 2001. Proceedings ICT 2001, 2001, pp. 65–68.
- [27] S. M. Kauzlarich, S. R. Brown, G. J. Snyder, Zintl phases for thermoelectric devices, *Dalton Trans.* (21) (2007) 2099–2107.
URL <http://pubs.rsc.org/en/content/articlelanding/2007/dt/b702266b>
- [28] G. J. Snyder, M. Christensen, E. Nishibori, T. Caillat, B. B. Iversen, Disordered zinc in Zn₄Sb₃ with phonon-glass and electron-crystal thermoelectric properties, *Nature Materials* 3 (7) (2004) 458–63.

URL <http://search.proquest.com/docview/222759382/abstract?accountid=10041>

- [29] G. K. H. Madsen, D. J. Singh, BoltzTraP. A code for calculating band-structure dependent quantities, *Computer Physics Communications* 175 (1) (2006) 67–71.
URL <http://www.sciencedirect.com/science/article/pii/S0010465506001305>

- [30] W. Li, J. Carrete, N. A. Katcho, N. Mingo, ShengBTE: A solver of the Boltzmann transport equation for phonons, *Computer Physics Communications* 185 (6) (2014) 1747–1758.
URL <http://www.shengbte.org>