

UC Berkeley

UC Berkeley Electronic Theses and Dissertations

Title

First principles calculation of material properties of group IV elements and III-V compounds

Permalink

<https://escholarship.org/uc/item/6895c20f>

Author

Malone, Brad Dean

Publication Date

2012

Peer reviewed|Thesis/dissertation

First principles calculation of material properties of group IV elements and III-V
compounds

by

Brad Dean Malone

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Physics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Marvin L. Cohen, Chair

Professor Daryl C. Chrzan

Professor Alex Zettl

Fall 2012

First principles calculation of material properties of group IV elements and III-V
compounds

Copyright 2012

by

Brad Dean Malone

Abstract

First principles calculation of material properties of group IV elements and III-V compounds

by

Brad Dean Malone

Doctor of Philosophy in Physics

University of California, Berkeley

Professor Marvin L. Cohen, Chair

This thesis presents first principles calculations on the properties of group IV elements and group III-V compounds. It includes investigations into what structure a material is likely to form in, and given that structure, what are its electronic, optical, and lattice dynamical properties as well as what are the properties of defects that might be introduced into the sample. The thesis is divided as follows:

- Chapter 1 contains some of the conceptual foundations used in the present work. These involve the major approximations which allow us to approach the problem of systems with huge numbers of interacting electrons and atomic cores. This involves a discussion of the Born-Oppenheimer approximation, Bloch's Theorem, the concept of the pseudopotential, and the empirical pseudopotential method (EPM). We end with a discussion of density functional theory (DFT), which is methodology by which the majority of solid state work is carried out today.
- Then, in Chapter 2, we discuss one of the major limitations to the DFT formalism introduced in Chapter 1, namely its inability to predict the quasiparticle spectra of materials and in particular the band gap of a semiconductor. We introduce a Green's function approach to the electron self-energy Σ known as the *GW* approximation and use it to compute the quasiparticle band structures of a number of group IV and III-V semiconductors. Our fully first principles calculations, which include spin-orbit effects, are in very good agreement with experimental values and illustrate the power of this approach to calculate quasiparticle energies.
- In Chapter 3 we present a first-principles study of a number of high-pressure metastable phases of Si with tetrahedral bonding. The phases studied include all experimentally determined phases that result from decompression from the metallic β -Sn phase, specifically the BC8 (Si-III), hexagonal diamond (Si-IV), and R8 (Si-XII). In addition to these, we also study the hypothetical ST12

structure found upon decompression from β -Sn in germanium. We find that the DFT incorrectly predicts the Si-XII phase to be semimetallic and that the true quasiparticle spectrum exhibits a band gap.

- Our attention is then turned to the first principles calculations of optical properties in Chapter 4. The optical spectrum of the Si-XII is studied, which was shown in Chapter 3 to exhibit a small, indirect band gap. The Bethe-Salpeter equation is then solved to obtain the optical spectrum of this material including electron-hole interactions. The calculated optical spectrum is compared with experimental data for other forms of silicon commonly used in photovoltaic devices, namely the cubic, polycrystalline, and amorphous forms. The calculated values of the optical functions relevant to photovoltaic absorption are found to exhibit show greater overlap with the incident solar spectrum in Si-XII than those found in these other silicon phases.
- In Chapter 5 we present first principles calculations of the quasiparticle and optical excitation spectra of recently predicted silicon and germanium polytypes in the body-centered-tetragonal (bct) structure. The quasiparticle spectra calculated within the GW approximation predict that both silicon and germanium in the bct structure are small band gap materials. The optical spectra are then evaluated by solving the Bethe-Salpeter equation taking into account.
- We examine the low-pressure phases of Ge in Chapter 6 by performing first principles calculations of the electronic structure and lattice dynamics of the R8, BC8, ST12, and hexagonal diamond structures of Ge. To aid future experimental investigation, we include predictions of the Raman-active frequencies of these phases as well as present the full phonon dispersion throughout the zone. Calculated equation of states within the LDA reveal a small pressure region where the R8 phase is energetically favored over the other metastable BC8 and ST12 structures, although the energy differences involved are relatively small and affected by the approximations used in the choice of pseudopotential. The calculated zone-center Raman-active frequencies are then used to further support experimental evidence which suggests the R8 phase of Ge to have been formed in indented samples of amorphous Ge. We also show that from an analysis of the pressure dependence of the zone-center Raman-active phonon frequencies that previous experimental results claiming to have the ST12 phase of Ge in their experimental sample were actually likely seeing the signature of the R8 phase, a phase which in Ge had until then remained unobserved.
- In Chapter 7 we demonstrate how first principles calculations can be used to predict new structures. In a study aimed at finding new useful forms of silicon, we use an *ab initio* random structure searching (AIRSS) method to identify a new phase of silicon in the *Ibam* structure. The *Ibam* phase is found to

be semimetallic within density functional theory with a small band overlap, and it is expected that quasiparticle corrections using the GW approximation would yield a semiconducting state with a small band gap. Calculation of the lattice dynamics reveals that the structure is locally stable. Enthalpy-pressure relations are calculated for the *Ibam* structure as well as all other known Si structures, including the previously predicted st12 and bct phases. These results indicate that *Ibam* silicon is metastable over the pressure range considered. Calculated coexistence pressures of the other known phase transitions are in good agreement with experimental observation. We end this section with a discussion on evolutionary algorithms and their application to the problem of structure prediction. We then use this method to determine some interesting low-energy candidate structures for the unknown phase Si-VIII.

- We present a first-principles study of boron and phosphorus substitutional defects in Si-XII in Chapter 8. Recent result from nanoindentation experiments reveal that the Si-XII phase is semiconducting and has the interesting property that it can be doped n- and p-type at room temperature without an annealing step. Using the hybrid functional of Heyd, Scuseria, and Ernzerhof (HSE), we examine the formation energies of the B and P defects at the two distinct atomic sites in Si-XII to find on which site the substitutional defects are more easily accommodated. We also estimate the thermodynamic transition levels of each defect in its relevant charge states. The hybrid calculations also give an independent prediction that Si-XII is semiconducting, in agreement with recent experimental data.

To my wife Mary -

Someday,

When my life has passed me by

I'll lay around and wonder why you were always there for me.

Contents

List of Figures	v
-----------------	---

List of Tables	viii
----------------	------

1	Introduction	1
1.1	The Problem	1
1.2	Born-Oppenheimer Approximation	2
1.3	Bloch's Theorem	3
1.4	Pseudopotentials	4
1.5	Empirical Pseudopotential Method (EPM)	7
1.6	Density functional theory (DFT)	10
1.6.1	The Hohenberg-Kohn theorems	11
1.6.2	Kohn-Sham equations	13
1.7	Conclusion	16
2	Fully <i>ab initio</i> quasiparticle spectra of semiconductors	18
2.1	Introduction	18
2.2	Method details	20
2.2.1	DFT	20
2.2.2	GW calculations	21
2.2.3	Spin-orbit interaction	27
2.3	Results	29
2.4	Conclusion	49
3	<i>Ab initio</i> survey of the electronic structure of tetrahedrally bonded phases of silicon	50
3.1	Introduction	50
3.2	Structures	51
3.3	Method	52
3.4	Results	53
3.4.1	R8	53
3.4.2	BC8	56

3.4.3	Lonsdaleite	58
3.4.4	ST12	59
3.5	Conclusion	62
4	<i>Ab initio</i> study of the optical properties of Si-XII	64
4.1	Introduction	64
4.2	Method	65
4.3	Results	66
4.3.1	Optical Spectrum	66
4.3.2	Photovoltaic application	69
4.4	Conclusion	72
5	Electronic and optical properties of body-centered tetragonal Si and Ge	74
5.1	Introduction	74
5.2	Method	75
5.3	Results and discussion	77
5.4	Conclusion	84
6	Electronic structure, equation of state, and lattice dynamics of low-pressure Ge polymorphs	86
6.1	Introduction	86
6.2	Methods	87
6.3	Results	88
6.3.1	Structural descriptions	88
6.3.2	Electronic structure	89
6.3.3	Lattice dynamics	91
6.3.4	Equation of state	96
6.4	Comparison to experiment	104
6.5	Conclusion	106
7	Prediction of a new metastable phase of silicon in the <i>Ibam</i> structure	108
7.1	Introduction	108
7.2	AIRSS method details	109
7.3	Results	109
7.4	Evolutionary methods: The search for Si-VIII	118
7.5	Conclusion	126
8	First principles study of phosphorus and boron substitutional defects in Si-XII	127
8.1	Introduction	127
8.2	Calculation of defect properties	129
8.3	Results	132

8.4	Conclusion	139
A	Convergence data for GW calculations in Chapter 2	156
A.1	Si	156
A.2	Ge	160
A.3	AlP	164
A.4	AlAs	168
A.5	AlSb	172
A.6	InP	176
A.7	InAs	180
A.8	InSb	184
A.9	GaP	188
A.10	GaAs	192
A.11	GaSb	196
B	Structural descriptions of Si-VIII candidates	200
B.1	30 atom candidate	200
B.2	32 atom candidate	202

List of Figures

2.1	Σ_{CH} empty state convergence for Si with a dielectric matrix cutoff of 40 Ry	24
2.2	Σ_{CH} empty state convergence for Si with a dielectric matrix cutoff of 10 Ry	25
2.3	Quasiparticle bandstructure of Si	32
2.4	Quasiparticle bandstructure of Ge	33
2.5	Quasiparticle bandstructure of AlP	38
2.6	Quasiparticle bandstructure of AlAs	39
2.7	Quasiparticle bandstructure of AlSb	40
2.8	Quasiparticle bandstructure of InP	41
2.9	Quasiparticle bandstructure of InAs	43
2.10	Quasiparticle bandstructure of InSb	44
2.11	Quasiparticle bandstructure of GaP	45
2.12	Quasiparticle bandstructure of GaAs	47
2.13	Quasiparticle bandstructure of GaSb	48
3.1	LDA energy overlap in R8 as a function of pressure	54
3.2	Band structure of Si in the R8 phase	55
3.3	Band structure of Si in the BC8 phase	57
3.4	Density-of-states for Si in the lonsdaleite phase	59
3.5	Band structure of the ST12 phase of Si	60
3.6	Density-of-states for Si in the ST12 phase	61
4.1	Calculated $\epsilon_2(\omega)$ for Si-XII	67
4.2	Joint density of states (JDOS) for Si-XII	68
4.3	Comparison of Si-XII $\epsilon_2(\omega)$ with other forms of Si	70
4.4	Comparison of optical absorption coefficient $\alpha(\omega)$ of Si-XII with that of other Si phases	71
4.5	Absorbed energy flux as a function of film thickness	72
5.1	Calculated $\epsilon_2(\omega)$ of the Si and Ge in the bct phase	79
5.2	JDOS and average squared matrix element for Si and Ge in the bct phase	81

5.3	Comparison of the $\epsilon_2(\omega)$ of bct Si with cubic Si	82
5.4	Optical absorption coefficient $\alpha(\omega)$ for Si in the bct and cubic phases	83
5.5	$\epsilon_2(\omega)$ for Ge in the bct and cubic phases	84
6.1	Band structure of ST12 Ge	90
6.2	Band structure of BC8 Ge	92
6.3	Band structure of R8 Ge	93
6.4	Band structure of hexagonal diamond Ge	94
6.5	Phonon dispersion of ST12 Ge	95
6.6	Phonon dispersion of BC8 Ge	97
6.7	Phonon dispersion of R8 Ge	98
6.8	Phonon dispersion of hexagonal diamond Ge	99
6.9	Energy-volume relations for low-pressure Ge phases	100
6.10	Pressure-volume relations for low-pressure Ge phases	101
6.11	Enthalpy as a function of pressure for low-pressure Ge phases	102
6.12	Pressure dependence of Raman-active modes in low-pressure Ge phases and comparison to experiment	106
7.1	Structure of <i>Ibam</i> Si.	111
7.2	Phonon dispersion of <i>Ibam</i> Si	112
7.3	Bandstructure of <i>Ibam</i> Si	113
7.4	Density of states (DOS) of <i>Ibam</i> Si	114
7.5	Energy-volume curves for observed and predicted Si phases	116
7.6	Pressure-volume curves for observed and predicted Si phases	117
7.7	Enthalpy-pressure curved for observed and predicted Si phases	118
7.8	Evolutionary algorithm data for 30 atom Si-VIII candidate	120
7.9	Energy-volume curve for the 30 atom Si-VIII candidate	121
7.10	Enthalpy as a function of pressure for the 30 atom Si-VIII candidate	122
7.11	Structure of a Si-VIII candidate obtained from an evolutionary algo- rithm run	123
7.12	Energy-volume curve for 32 atom Si-VIII candidate	124
7.13	Enthalpy as a function of pressure for the 32 atom Si-VIII candidate	125
8.1	Structure of Si-XII	129
8.2	Displacements of atoms surrounding defect in Si-XII	136
8.3	Electrostatic correction scheme data for P defect at 2(c) site in Si-XII	137
8.4	Formation energies of the P and B substitutional defects at the 2(c) and 6(f) sites of Si-XII	138
A.1	Si: Testing number of bands in Σ_{CH}	157
A.2	Si: Testing the size of the dielectric matrix	158
A.3	Si: Testing the number of bands in χ	159
A.4	Ge: Testing number of bands in Σ_{CH}	161

A.5	Ge: Testing the size of the dielectric matrix	162
A.6	Ge: Testing the number of bands in χ	163
A.7	AlP: Testing number of bands in Σ_{CH}	165
A.8	AlP: Testing the size of the dielectric matrix	166
A.9	AlP: Testing the number of bands in χ	167
A.10	AlAs: Testing number of bands in Σ_{CH}	169
A.11	AlAs: Testing the size of the dielectric matrix	170
A.12	AlAs: Testing the number of bands in χ	171
A.13	AlSb: Testing number of bands in Σ_{CH}	173
A.14	AlSb: Testing the size of the dielectric matrix	174
A.15	AlSb: Testing the number of bands in χ	175
A.16	InP: Testing number of bands in Σ_{CH}	177
A.17	InP: Testing the size of the dielectric matrix	178
A.18	InP: Testing the number of bands in χ	179
A.19	InAs: Testing number of bands in Σ_{CH}	181
A.20	InAs: Testing the size of the dielectric matrix	182
A.21	InAs: Testing the number of bands in χ	183
A.22	InSB: Testing number of bands in Σ_{CH}	185
A.23	InSb: Testing the size of the dielectric matrix	186
A.24	InSb: Testing the number of bands in χ	187
A.25	GaP: Testing number of bands in Σ_{CH}	189
A.26	GaP: Testing the size of the dielectric matrix	190
A.27	GaP: Testing the number of bands in χ	191
A.28	GaAs: Testing number of bands in Σ_{CH}	193
A.29	GaAs: Testing the size of the dielectric matrix	194
A.30	GaAs: Testing the number of bands in χ	195
A.31	GaSb: Testing number of bands in Σ_{CH}	197
A.32	GaSb: Testing the size of the dielectric matrix	198
A.33	GaSb: Testing the number of bands in χ	199

List of Tables

2.1	Lattice constant, bulk moduli, and wavefunction cutoff for group IV and III-V materials	21
2.2	Summary of convergence parameters for GW calculations	29
2.3	Quasiparticle energies for Si	31
2.4	Quasiparticle energies for Ge	34
2.5	Quasiparticle energies for III-V semiconductors AlX (X=P,As,Sb) . .	37
2.6	Quasiparticle energies for III-V semiconductors InX (X=P,As,Sb) . .	42
2.7	Quasiparticle energies for III-V semiconductors GaX (X=P,As,Sb) . .	46
5.1	LDA and GW energy gaps of bct Si and Ge	76
8.1	Details of the supercells employed in the calculation of the electrostatic corrections	133
8.2	Summary of the formation energy data in the neutral charge state and thermodynamic charge transition levels	134

Acknowledgments

There are many people whose help has made my time at Berkeley both productive and highly enjoyable, and whom I'd like to thank:

First is my thesis advisor, Professor Marvin L. Cohen, who has taught me so much, some no doubt intentionally and some that I learned from him while he wasn't looking. From him I've learned a great deal about every aspect of being a scientist, as well as some about being a good person.

When I first joined the group Jay Sau was my first collaborator, and I owe him for his availability and patience in working so well with a new graduate student that was just getting started in condensed matter, and in particular computational, physics. As such, he's probably heard me ask more "stupid questions" than almost anyone else. Here's to hoping he's forgot all those by now.

As a graduate student, having a knowledgeable postdoc around that is willing to help can be of tremendous advantage. As further proof that the world isn't fair, I got two. Georgy Samsonidze and Manish Jain have both been of tremendous help to me while I've been on the 5th floor Birge. They were always available for a good discussion, a good laugh, and made my problems their problems so many more times than I could possibly repay them for.

I've also benefited greatly from discussions and friendships among many other members of the Cohen and Louie groups. While by no means exhaustive, of particular mention are Jesse Noffsinger, Vasudha Shivamoggi (Moore group), David Strubbe, Kevin Chan, Jack Deslippe, and Manos Kioupakis.

Finally, I must thank my wife Mary. She's taught me nothing about physics, but a great deal about everything else.

Chapter 1

Introduction

1.1 The Problem

To understand the properties of a material is to understand the solution, if only approximate, of a system of interacting electrons and their nuclei. The governing equation for this system is the Schrödinger equation:

$$\hat{H} = -\frac{\hbar^2}{2m_e} \sum_i \nabla_i^2 + \sum_{i,I} \frac{Z_I e^2}{|\mathbf{r}_i - \mathbf{R}_I|} + \frac{1}{2} \sum_{i,j \neq i} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_I \frac{\hbar^2}{2M_I} \nabla_I^2 + \frac{1}{2} \sum_{I,J \neq I} \frac{Z_I Z_J e^2}{|\mathbf{R}_I - \mathbf{R}_J|} \quad (1.1)$$

where the quantities pertaining to the electrons have lower case subscripts and that of the ions upper case. At first glance the problem seems fairly simple, and can be understood (at least at a classical level) by those who have taken an introductory course on electrodynamics. The Hamiltonian only contains terms which are either kinetic in nature or which are due from the electrostatic repulsion or attraction between particles. It may surprise a great number of people that generations of the world's best physicists, chemists, and material scientists have struggled to solve this equation for even small numbers of particles without devoting a degree of computational resources that may exceed what humankind is currently capable of producing.

One of the difficulties present in the solution of the above equation is that most materials that we might be interested in studying contain a macroscopically large number of particles (e.g., $\sim 10^{23}$). Such a large number of entities means that even the solution of a classical problem may be intractable. Another problem is that the motion of the electrons and ions is correlated, i.e., the probability of finding an electron or ion in a specific position depends on the location of all the other electrons and ions of the system. This fact makes attempts to find the solution for individual particles without finding the solutions of the rest approximate at best. Finally, an additional obstacle is present because of the fact that the system is inherently quantum in nature.

Specifically, the wavefunction of the entire system $\Phi(\{\mathbf{r}_i\}; \{\mathbf{R}_I\})$ must obey the Pauli exclusion principle. This means that if two electrons of the same spin switch positions that the wavefunction Φ should change sign. Likewise, an interchange of the nuclei (for nuclei of the same type) might also introduce a sign change if the nuclear spin is half-integer. If an approximation to Eq. (1.1) is going to be realistic, it will want to be able to incorporate the physics of exchange at least to some degree.

1.2 Born-Oppenheimer Approximation

One of the first approximations to Eq. (1.1) which is often taken is known as the Born-Oppenheimer approximation. This approximation takes advantage of the fact that the nuclei, being much more massive than the electrons, move extremely slowly relative to the electrons. In this sense the electrons can be taken to *instantaneously* adjust to any ionic motion. More specifically, if the electronic system is in some state A (e.g., the ground state), the electronic system will stay in that state as the nuclei undergo their motion. In other words, the motion of the nuclei does not induce transitions between states in the electronic system. This does not mean that the electrons do not care about the positions or motion of the nuclei, simply that when solving the electronic system the positions of the nuclei only enter as parameters.

From a mathematical standpoint, the upshot of the Born-Oppenheimer approximation is that our full wavefunction $\Phi(\{\mathbf{r}_i\}; \{\mathbf{R}_I\})$, can be written as a product of a state which describes the nuclear system and a state which describes the electronic system with the nuclear coordinates entering as parameters:

$$\Phi(\{\mathbf{r}_i\}; \{\mathbf{R}_I\}) = \sum_n \Theta_n(\mathbf{R}_I) \Psi_n(\mathbf{r}, \mathbf{R}_{fixed}) \quad (1.2)$$

where $\Theta_n(\mathbf{R}_I)$ is the wavefunction describing the nuclear dynamics for each of the adiabatic eigenstates Ψ_n of the electronic system [1, 2]. The wavefunction Ψ_n is now the solution of a purely electronic Hamiltonian which can be written as

$$\hat{H}_e = -\frac{\hbar^2}{2m_e} \sum_i \nabla_i^2 + \sum_{i,I} \frac{Z_I e^2}{|\mathbf{r}_i - \mathbf{R}_I|} + \frac{1}{2} \sum_{i,j \neq i} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} \quad (1.3)$$

$$= -\frac{\hbar^2}{2m_e} \sum_i \nabla_i^2 + \sum_i V_{ion}(\mathbf{r}_i) + \frac{1}{2} \sum_{i,j \neq i} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} \quad (1.4)$$

where in the second line we have replaced the sum over the ions simply with the term $V_{ion}(\mathbf{r}) = \sum_I \frac{Z_I e^2}{|\mathbf{r} - \mathbf{R}_I|}$. Through this change of notation which hides the dependence on the parameters $\{\mathbf{R}_I\}$ we make it more readily apparent that we are concerned in the quantum problem of *only* the electrons. Additionally, as we will see later when

we examine the concept of the pseudopotential, this change of notation is convenient because the external potential acting on our electrons of interest will no longer be expressed as simply a sum over bare nuclear charges and their positions.

Before continuing with the discussion of how we solve the above electronic Hamiltonian in practice, it should be noted that it is possible (and common) to relax the assumptions made within the Born-Oppenheimer approximation. In particular, the assumption that the ionic degrees of freedom do not induce transitions between electronic eigenstates. These transitions do occur and can be treated perturbatively. The treatment of these transitions is what gives rise to the electron-phonon interaction that is central to the study of superconductivity, electrical resistivity, and other interesting physical phenomenon.

1.3 Bloch's Theorem

Here we will briefly discuss the Bloch theorem and the way in which we exploit the periodic nature of a material to simplify the solution of Eq. (1.4). After all, it is only because of the periodicity of the physical system that we can study a block of silicon containing $\sim 10^{23}$ atoms by studying instead a two-atom unit cell with a modest degree of Brillouin zone sampling. More extended descriptions, including proofs and examples, can be found in almost any introductory solid state textbook [3–5].

The study of matter is greatly simplified by the physical fact that solids are generally made up of periodically repeating units in space. This translation invariance simplifies the problem considerably. A crystal can be formed by taking the smallest repeat unit, known as a primitive cell, and repeating it in space using a set of vectors known as the lattice vectors. This translational invariance affects all of the observable properties of a material: the charge density $\rho(\mathbf{r})$, the potential that the electrons feel $V(\mathbf{r})$, etc.

$$\rho(\mathbf{r} + \mathbf{T}) = \rho(\mathbf{r}) \tag{1.5}$$

$$V(\mathbf{r} + \mathbf{T}) = V(\mathbf{r}) \tag{1.6}$$

where $\mathbf{T}$ is a translation operator which is consistent with the periodicity of the crystal.

The description of the excitations in the crystal, such as the wavefunctions of the electrons or the collective of the atoms, are not strictly periodic in the lattice. Bloch showed that the eigenstates of a Hamiltonian in a periodic potential pick up a phase difference upon a shift in position through the translation vector $\mathbf{T}$:

$$\psi_{\mathbf{k}}(\mathbf{r} + \mathbf{T}) = e^{i\mathbf{k} \cdot \mathbf{T}} \psi_{\mathbf{k}}(\mathbf{r}) \tag{1.7}$$

where we have labeled the wavefunction $\psi_{\mathbf{k}}(\mathbf{r})$ with the subscript $\mathbf{k}$ which determines the phase that this particular state takes up when being translated by the translation vector $\mathbf{T}$. An equivalent form of the Bloch theorem is that the solutions of a Hamiltonian in a periodic potential can be expressed as a function *which is* periodic in the lattice, $u_{\mathbf{k}}(\mathbf{r})$, times a plane wave with wavevector $\mathbf{k}$:

$$\psi_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{\mathbf{k}}(\mathbf{r}) \quad (1.8)$$

$$u_{\mathbf{k}}(\mathbf{r} + \mathbf{T}) = u_{\mathbf{k}}(\mathbf{r}). \quad (1.9)$$

It turns out that the $\mathbf{k}$ vectors occurring in Eqs. (1.7-1.9) cannot take on just any value. Assuming that we are considering a finite crystal with periodic boundary conditions the allowed form of the Bloch vectors is

$$\mathbf{k} = \sum_{i=1}^3 \frac{m_i}{N_i} \mathbf{b}_i \quad (1.10)$$

where m_i are integers, N_i is the number of lattice repeats in the direction of the i^{th} lattice vector, and $\mathbf{b}_i$ is a primitive vector of the reciprocal lattice (the lattice conjugate to the real-space lattice; for details see the introductory texts [3–5]). The $\mathbf{k}$ vectors can also be taken to be confined to the first primitive cell of the reciprocal lattice (also known as the Brillouin zone) because the eigenfunctions and eigenvalues of the Hamiltonian are periodic in the reciprocal lattice. It can further be shown that the number of allowed $\mathbf{k}$ vectors in Eq. 1.10 is simply equal to the number of cells in the crystal ($N_{\mathbf{k}}=N=\prod_i N_i$). Concluding this section, what the Bloch theorem allows us to do is instead of solving for the properties of an infinite (or near-infinite) crystal, we can instead solve the much smaller problem in the unit cell. While in principle the number of $\mathbf{k}$ vectors in the Brillouin zone is infinite, by considering a much smaller number of $\mathbf{k}$ vectors we will often be able to reach convergence of the properties of interest [1].

1.4 Pseudopotentials

When we are studying the properties of a material, it is common that the properties that we are interested in are intimately related to only some of the electrons that are present in the solid. For example, it is usually only the outer valence electrons that strongly participate in the bonding. Likewise, if one is interested in understanding the optical suitability of a new material for photovoltaics, it is again only the outer valence states which can even be excited by an incoming solar photon to a conduction band state. Similar statements can be made in relation to thermal excitations as well as the electrons which contribute to measured transport properties. For this fact alone it would be convenient if we could remove the electrons that do not participate *directly* in these phenomena and instead replace their effect on the valence electrons

by some other, cheaper means. The motivation for removing these inner, or core, electrons, is greater than simply reducing the number of electrons that are present in the system. These electronic wavefunctions often exhibit rapid variation as they move closer and closer to the nuclear core.¹ When we are expanding our wavefunctions in a plane-wave basis set, such rapid fluctuations in real-space correspond to extremely large cutoffs in the plane wave expansion, making the size of calculation prohibitive. The problem of removing the computationally intense nature of electrons that are essentially behaving exactly as they do in the free atom is the motivation behind the concept of the pseudopotential.

The most intuitive introduction of the ideas behind the modern pseudopotential approach is through the OPW-based derivation by Phillips and Kleinman [6, 7]. In this derivation the true valence wavefunction ψ_v is expressed as a sum of a smooth wavefunction ϕ_v plus a sum over occupied core states ψ_c ²

$$|\psi_v\rangle = |\phi_v\rangle + \sum_c A_{vc}|\psi_c\rangle. \quad (1.11)$$

These valence wavefunctions are made to be orthogonal to the core states, which implies that $A_{vc} = -\langle\psi_c|\phi_v\rangle$. Substituting Eq. (1.11) along with the relation for A_{vc} into the Schrödinger equation $H|\psi_v\rangle = E_v|\psi_v\rangle$ we obtain

$$H|\phi_v\rangle - \sum_c \langle\psi_c|\phi_v\rangle E_c|\psi_c\rangle = E_v\phi_v - E_v \sum_c \langle\psi_c|\phi_v\rangle|\psi_c\rangle \quad (1.12)$$

which can be simplified to yield

$$H|\phi_v\rangle + \sum_c (E_v - E_c)|\psi_c\rangle\langle\psi_c|\phi_v\rangle = E_v|\phi_v\rangle \quad (1.13)$$

$$(H + V_R)|\phi_v\rangle = E_v|\phi_v\rangle \quad (1.14)$$

where we have defined $V_R = \sum_c (E_v - E_c)|\psi_c\rangle\langle\psi_c|$. What we have done in the above is transformed our original equation for the full wavefunction ψ_v into an equation for the smooth part of the wavefunction ϕ_v (also known as the pseudo-wavefunction). The solution of this equation is not some pseudo-energy however, it is the actual energy E_v of the true valence wavefunction ψ_v . In doing this transformation from the full wavefunction to smooth wavefunction the terms were added in Eq. (1.11) to guarantee

¹The cost savings is actually not only because of the removal of the core wavefunctions from the problem, but also the replacement of the true valence wavefunctions with smooth, nodeless pseudo-wavefunctions that only match the true wavefunctions in the atomic calculation outside of the core radius.

²This form for the wavefunction has its origins in the early OPW-method works of Herring [8]. Herring found that by constructing the wavefunctions in such a manner it was much easier to take into account the rapid oscillatory behavior that exists near the core region. This makes sense, as the core functions that have been added in contain such rapid oscillatory behavior themselves [4].

the correct orthogonalization to the core states become a new potential V_R felt by the smooth wavefunction ϕ_v . Because the energy of the valence states, E_v , is greater than that of the core states, E_c , V_R is a positive (and hence repulsive) potential. Furthermore, because it only acts in the region where the core wavefunctions are appreciable, is relatively short-ranged.

The Hamiltonian for our pseudo-wavefunction can also be expressed as

$$(H + V_R)\phi_v = (T + V_c + V_R)\phi_v = E_v\phi_v \quad (1.15)$$

$$(T + V_{pseudo})\phi_v = E_v\phi_v \quad (1.16)$$

where V_c is the regular ionic potential generated from the nuclear cores and V_R is the repulsive part obtained above. The sum of these two terms is what is called the pseudopotential V_{pseudo} . This pseudopotential will result in a weaker potential due to the partial cancellation between the attractive ionic potential and the repulsive part V_R . This allows for the resulting potential to more easily be treated within perturbation theory starting from a plane wave basis set³.

The construction of the pseudopotential that we have outlined above is not how pseudopotentials are constructed in modern calculations, although they do exhibit some of the main general features of modern pseudopotentials. Like V_R , modern pseudopotentials are often both non-local operators and behave differently towards states of different momentum. These characteristics are present in a pseudopotential of the form [1]

$$\hat{V}_{pseudo}(\mathbf{r}) = \sum_{l=0}^{\infty} \sum_{m=-l}^l v_{ps}^l(r) |lm\rangle \langle lm| = \sum_{l=0}^{\infty} v_{ps}^l(r) \hat{P}_l \quad (1.17)$$

where $\hat{P}_l$ is the projector operator on the angular momentum channel l and m is the quantum number corresponding to the projection on a chosen axis. Typically the expression in Eq. (1.17), which is an explicit sum over l -dependent pseudopotential contributions, is rewritten as the sum of a local channel which affects all angular momenta plus l -dependent corrections for each channel. The l -dependent corrections are zero outside the wavefunction cutoff where the pseudopotential then becomes exclusively the local part. The local part of the pseudopotential incorporates the long-range effects of the Coulomb potential (such as approaching the limit of $-\frac{Z_{ion}}{r}$ for large r) [2].

There are a number of details regarding modern pseudopotential generation that we will not go into here (the reader is referred to a number of the good references on the subject [1, 2, 5, 9]), but we will conclude this discussion on pseudopotentials by giving a quick outline of how modern pseudopotentials *are* constructed:

³Plane waves are of course the form of the true solution when the potential is zero, i.e. an empty lattice.

- For each valence wavefunction, a cutoff in the radial coordinate r is chosen such that outside the cutoff the pseudo-wavefunction should be equal to the true all-electron valence wavefunction. Furthermore, the pseudo-wavefunction should be an eigenstate of the pseudo-Hamiltonian with the same eigenvalue as the all-electron wavefunction (although sometimes this energy, known as the *reference energy*, is taken to be something else).
- Given these conditions (and possibly others, such as the norm-conservation of the pseudo-wavefunction within the cutoff radius [10]), the radial Schrödinger equation is inverted to find what potential gives rise to such a solution. This potential is what is known as the *screened* pseudopotential.
- The reason why the “screened” pseudopotential is not the desired quantity is that it contains the effects of the valence electrons themselves. These effects do not describe what we are intending to model - purely the interaction between the valence electrons and the electrons of the core. Thus we must subtract out the valence contributions, or “descreen”, this pseudopotential. This is simply done by subtracting out the Hartree and exchange-correlation contributions coming from the valence electrons.⁴

1.5 Empirical Pseudopotential Method (EPM)

Before the advent of the modern *ab initio* pseudopotential and methods which allowed for these pseudopotentials to be used efficiently within the framework of density functional theory (DFT) [11], understanding the potential that the electrons felt inside a solid was no small task. However, because a crystal is periodic it was known that whatever form the potential took it should be expressible as a sum of Fourier components in the reciprocal lattice vectors of the crystal. Furthermore, as we saw in section (1.4), the effective potential felt by the valence electrons in a solid could be made weaker than the true potential without losing a good description of the valence electrons in the solid. This weaker potential implies that its Fourier representation can be constructed with a small number of reciprocal lattice vectors. If it were possible to obtain these parameters from experiment then the full bandstructure (as well as other properties) could be calculated without a detailed microscopic understanding of how that effective potential was generated. This idea is the motivation behind the empirical pseudopotential method [7, 12, 13]

If we take the full crystal potential to be of the form of a sum of potentials around

⁴The Hartree potential is simply the regular electrostatic potential due to the valence electrons. The exchange-correlation contribution, which will be discussed in more detail later, is a best attempt to take into account the exchange property of the many-body wavefunction and correlated nature of the electrons that was discussed in Section 1.1.

each ion (below assumed to be the same type of ion for simplicity) [4, 5, 7]

$$V_{xtal}(\mathbf{r}) = \sum_{\mathbf{R}, \tau} V_{atom}(\mathbf{r} - \mathbf{R} - \tau) \quad (1.18)$$

we can expand this potential in the reciprocal lattice

$$V_{xtal}(\mathbf{r}) = \sum_{\mathbf{G}} V_{xtal}(\mathbf{G}) e^{i\mathbf{G} \cdot \mathbf{r}}. \quad (1.19)$$

The Fourier components of this potential can be expressed as

$$V_{xtal}(\mathbf{G}) = \int_{xtal} \frac{d^3\mathbf{r}}{\Omega_{xtal}} V_{xtal}(\mathbf{r}) e^{-i\mathbf{G} \cdot \mathbf{r}} = \sum_{\mathbf{R}, \tau} \int_{xtal} \frac{d^3\mathbf{r}}{\Omega_{xtal}} V_{atom}(\mathbf{r} - \mathbf{R} - \tau) e^{-i\mathbf{G} \cdot \mathbf{r}} \quad (1.20)$$

$$= \frac{1}{N_{cell}} \left(\sum_{\mathbf{R}} e^{i\mathbf{G} \cdot \mathbf{R}} \right) \left[\int_{cell} \frac{d^3\mathbf{r}}{\Omega_{cell}} V_{atom}(\mathbf{r}) e^{-i\mathbf{G} \cdot \mathbf{r}} \right] \sum_i e^{i\mathbf{G} \cdot \tau} \quad (1.21)$$

$$= \frac{1}{N_{cell}} N_{cell} \left[\int_{cell} \frac{d^3\mathbf{r}}{\Omega_{cell}} V_{atom}(\mathbf{r}) e^{-i\mathbf{G} \cdot \mathbf{r}} \right] \sum_i e^{i\mathbf{G} \cdot \tau} \quad (1.22)$$

$$= V_{atom}(\mathbf{G}) S(\mathbf{G}) \quad (1.23)$$

where we have introduced the structure factor $S(\mathbf{G}) = \sum_i e^{i\mathbf{G} \cdot \tau}$. The term in brackets in Eq. (1.22) is the atomic form factor $V_{atom}(\mathbf{G})$. Putting all of this together we can see that the crystal potential can be given as

$$V_{xtal}(\mathbf{r}) = \sum_{\mathbf{G}} V_{atom}(\mathbf{G}) S(\mathbf{G}) e^{i\mathbf{G} \cdot \mathbf{r}}. \quad (1.24)$$

These atomic form factors $V_{atom}(\mathbf{G})$ are the factors which we believe to only need a few Fourier components to express. Additionally for some values of $\mathbf{G}$ the structure factor will be identically zero, and so regardless of what the value of that particular $V_{atom}(\mathbf{G})$ might be, we will not need it in our computations. Additionally, if we assume that $V_{atom}(\mathbf{r}) = V_{atom}(|\mathbf{r}|)$ (i.e., the atomic potentials are spherical), then the atomic form factors only depend on the magnitude of $\mathbf{G}$, which reduces the number that are needed to be determined experimentally by even more [13].

Now if we specialize to zinc-blende type materials in which we have two types of atoms, A and B, at the atomic positions $\tau = \pm(\frac{1}{8}, \frac{1}{8}, \frac{1}{8})$ we can write

$$V_{atom}(\mathbf{G}) S(\mathbf{G}) \rightarrow \sum_{\alpha} V_{\alpha} S_{\alpha}(\mathbf{G}) = V_S(\mathbf{G}) \cos(\mathbf{G} \cdot \tau_+) + i V_A(\mathbf{G}) \sin(\mathbf{G} \cdot \tau_+) \quad (1.25)$$

where we have explicitly shown the generalization back to a system where the atoms are not identical and $\tau_+ = (\frac{1}{8}, \frac{1}{8}, \frac{1}{8})$. V_S and V_A are known as the symmetric and antisymmetric form factors, respectively, and are defined as

$$2V_S(\mathbf{G}) = V_A(\mathbf{G}) + V_B(\mathbf{G}) \quad (1.26)$$

$$2V_A(\mathbf{G}) = V_A(\mathbf{G}) - V_B(\mathbf{G}) \quad (1.27)$$

Clearly in the case of a diamond structure $A = B$ and the antisymmetric form factor is zero. Because a number of the structure factors going to zero, a good description of the zinc-blende materials can be obtained by only knowing the form factors $V(\mathbf{G})$ for $|\mathbf{G}|^2=0, 3, 4, 8$, and 11 (in units of $(2\pi/a)^2$ [12]⁵. These form factors can be obtained from the use of experimental data, in particular from the results of reflectivity and photoemission measurements [12].

Once an initial estimate for the form factors have been obtained, thus determining the potential, the next step is to solve the Schrödinger equation:

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V_{xtal}(\mathbf{r}) \right) \psi_{n\mathbf{k}}(\mathbf{r}) = E_{n,\mathbf{k}} \psi_{n\mathbf{k}}(\mathbf{r}). \quad (1.28)$$

But wait! Equation (1.28) is a single-particle Schrödinger equation. The Hamiltonian we were originally intending to solve, Eq. (1.4), clearly involves *many* electrons, and our wavefunction in Eq. (1.2) depends on the positions of all the electrons in the system. Why should we be considering the solution of an equation like Eq. (1.28) at all? What would such a solution mean, and how would we relate it back to the many-body wavefunction that we expected to be solving for? By allowing our wavefunctions to be the solutions of single-particle Schrödinger equations, we are indeed making an approximation to the full many-body wavefunction. However, it is not as if in doing so we are neglecting the effects of the other electrons. We are simply assuming that such effects can be represented in an average way by a suitable construction of the potential that the electrons feels. This is known as the *independent particle*, or *mean field*, approximation. Despite the fact that this approximation may at first glance seem quite severe, a great deal of progress can be made by considering solutions of such single-particle equations. The methods known as the Hartree and Hartree-Fock approximations begin with different forms of trial many-body wavefunctions, the components of which reduce to a set of single-particle equations similar to that considered in Eq. (1.28). Furthermore, as we'll discuss in the next section, the Kohn-Sham equations of density functional theory are also a set of single-particle equations. However, in principle this set of single-particle equations can describe the physics of the many-body system *exactly*. To conclude this sidebar, while it is important to understand that in setting up our single-particle Schrödinger equations that we are making an approximation, the severity of such an approximation need not be large depending on the surrounding formalism and the quality of our effective potential.

⁵Specifically, $|\mathbf{G}|^2=3,4,11$ for the antisymmetric part and $|\mathbf{G}|^2=0,3,8,11$ for the symmetric part.

Returning now to the discussion of the empirical pseudopotential method, it can now more easily be understood what we are doing when we are determining the form factors, and thus the potential, from experimental input. We are determining the effective single-particle potentials that give best reproduce the experimental input. Once these effective potentials are known, we are able to solve for the wavefunctions and energies that correspond to such an effective potential. These can then be used to calculate other quantities of interest, such as the density-of-states, the valence charge density, etc [13].

1.6 Density functional theory (DFT)

The many-body wavefunction which is a solution to Eq. (1.4) is certainly an object of great complexity. If we are considering a system of 10 electrons, the many-body wavefunction depends explicitly on all 10 electron positions, i.e. $\Psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4, \mathbf{r}_5, \mathbf{r}_6, \mathbf{r}_7, \mathbf{r}_8, \mathbf{r}_9, \mathbf{r}_{10})$. The correlated nature of the electrons, plus the fact that this wavefunction must satisfy the Pauli exchange principle, means that a great deal of effort would need to be expended to get this solution. But the situation is even worse than that. Let's assume we had a trade agreement with an alien civilization. This agreement allows us to obtain the many-body wavefunction for any system that we are interested in. Would we even be able to store the solution that they sent us? Take the example of 10 electrons above. If we represent the 10 electrons on a real-space grid of size $10 \times 10 \times 10$ (which is quite coarse), the many-body wavefunction would be specified by $10^{3 \cdot 10} = 10^{30}$ numbers. If we assume only single-precision floats at 4 bytes each, we are talking about a storage requirement of 4×10^{30} bytes, or 4×10^{18} terabytes (TB)! Currently the cost of a TB of storage goes for about \$75 retail. Even with technology that would drop that cost down to \$1/TB then the cost to store the many-body wavefunction for just those 10 electrons on a coarse grid would amount to 4 million *trillion* dollars ⁶. This is about 4 million times more money then currently is circulating around the world in USD. Clearly, we have strong reasons to consider approaches which allow us to get around the full determination of the many-body wavefunction, "especially in this economy".

One such method is to reformulate the problem not in terms of the many-body wavefunction, but simply in terms of the electronic density. If we knew the density on a grid of $10 \times 10 \times 10$, we would only need to store 1000 numbers, or 4 kilobytes (corresponding to the size of about 28 text messages on your cell phone). Early attempts at reformulating the many-body problem in terms of the charge density were made by Thomas and Fermi [14, 15], with later contributions made by Dirac [16]. The idea behind their approach was that they would use known results from the homogeneous electron gas at a certain density and apply it to an inhomogeneous systems by treating the energy contributions *locally* like that of a homogeneous gas of

⁶Plus tax. We would likely qualify for free shipping.

that particular density. They did this by attempting to rewrite all of the contributions to the energy (kinetic, exchange, and correlation) in the form [1, 2]

$$E_\alpha[\rho] = \int \rho(\mathbf{r}) \varepsilon_\alpha[\rho(\mathbf{r})] d^3\mathbf{r} \quad (1.29)$$

where $\varepsilon_\alpha[\rho(\mathbf{r})]$ is the energy density of energy contribution α . In this manner, they were successful in constructing an energy functional that depended on the electronic density alone [2]

$$E_{TFD}[\rho] = C_1 \int \rho(\mathbf{r})^{5/3} d^3\mathbf{r} + \int \rho(\mathbf{r}) V_{ext} d^3\mathbf{r} + C_2 \int \rho(\mathbf{r})^{4/3} d^3\mathbf{r} + \frac{1}{2} \int \int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3\mathbf{r} d^3\mathbf{r}' \quad (1.30)$$

where C_1 and C_2 are simply numerical constants that are not of interest to us here. Minimization of this functional with respect to the density $\rho(\mathbf{r})$ leads to an equation in which $\rho(\mathbf{r})$ can be written as a unique function of the external potential. Thus if you specify the external potential, you will know the interacting charge density of your system. However, this theory did not meet much success in its applications to real materials [1, 2]. One of the major flaws of the theory was its inaccuracy in treating the kinetic contribution to the energy. It turns out that the kinetic energy cannot be very adequately represented as a functional of a local density. As we will soon discuss, progress in the treatment of this term was later made by a shrewd mapping of the interacting problem to a non-interacting problem in which the kinetic energy contribution could be readily calculated.

The above Thomas-Fermi-Dirac approach was, for the most part, unsuccessful in the modeling of real materials. However, the idea behind it was the predecessor of the immensely successful density functional theory (DFT), the theory that is currently the workhorse for electronic structure calculations throughout the world and for which W. Kohn⁷ later shared the Nobel Prize in chemistry with John Pople in 1998 [5]. The main idea behind DFT is the replacement of many-body problem by an associated independent particle problem, both formulated in terms of the density. This mapping can be made *in principle* without the introduction of any approximations. However, a discussion on how this is done will make little sense if we do not understand a couple of facts about the many-body problem and its reformulation in terms of the density. We must first understand two important theorems introduced by Hohenberg and Kohn [17].

1.6.1 The Hohenberg-Kohn theorems

One of the results of the Thomas-Fermi-Dirac formulation was that the interacting density could be related uniquely to the external potential. However, this was only

⁷Walter Kohn was one of the primary developers of DFT.

shown within the approximations that they used (and which we already have said are inadequate for real calculations). One question that could be asked is, is it true that true (without approximation) interacting density should be uniquely related to the external potential? This is the content of the first theorem of Hohenberg and Kohn. It can be proven by *reductio ad absurdum*⁸. We assume instead that the statement is not true, i.e., we assume that the external potential is not uniquely determined by the density. Thus we assume that two different potentials $V_{ext}^1(\mathbf{r})$ and $V_{ext}^2(\mathbf{r})$, which differ by more than a trivial constant, give rise to the same ground state density $\rho_0(\mathbf{r})$. Associated with these two external potentials are two different Hamiltonians $\hat{H}^1$ and $\hat{H}^2$, which differ only in form that they have the two different external potentials. The lowest energy solution of $\hat{H}^1$ and $\hat{H}^2$ (i.e., the ground state) are the many-body wavefunctions Ψ^1 and Ψ^2 , respectively, which are both assumed to give rise to the same ground state density $\rho_0(\mathbf{r})$. The ground state energy of $\hat{H}^1$ is simply $E_0^1 = \langle \Psi^1 | \hat{H}^1 | \Psi^1 \rangle$ while that of $\hat{H}^2$ is given by $E_0^2 = \langle \Psi^2 | \hat{H}^2 | \Psi^2 \rangle$.

By the variational principle it must be true that

$$E_0^1 < \langle \Psi^2 | \hat{H}^1 | \Psi^2 \rangle = \langle \Psi^2 | \hat{H}^2 + \hat{H}^1 - \hat{H}^2 | \Psi^2 \rangle \quad (1.31)$$

$$= E_0^2 + \int \rho_0(\mathbf{r}) (V_{ext}^1(\mathbf{r}) - V_{ext}^2(\mathbf{r})) d^3\mathbf{r}. \quad (1.32)$$

We can make the same argument for E_0^2 , obtaining

$$E_0^2 < E_0^1 + \int \rho_0(\mathbf{r}) (V_{ext}^2(\mathbf{r}) - V_{ext}^1(\mathbf{r})) d^3\mathbf{r} \quad (1.33)$$

Now if we add Eq. (1.32) and Eq. (1.33) we get the following result

$$E_0^1 + E_0^2 < E_0^1 + E_0^2 \quad (1.34)$$

which under no circumstance can be correct. Thus we conclude that our initial assumption must be incorrect. Therefore there is a direct one-to-one correspondence between the interacting ground state charge density and the external potential. Since the integral of the charge density gives the number of electrons in our system (thus determining the number of electrons we should sum over in the kinetic and interacting electron terms in the Hamiltonian), the ground state charge density determines the *full* Hamiltonian [18]. If it determines the full Hamiltonian, then it determines the solutions of that Hamiltonian, i.e. all many-body wavefunctions including the excited states (as well as *any* other property which is determined by the Hamiltonian).

The second Hohenberg-Kohn theorem regards the expression of the total energy not as an expectation value of the Hamiltonian of the many-body wavefunction, but as an explicit functional of the density. We know that this must be true, since we just discussed that the external potential is uniquely specified by the density, and

⁸Also known as proof by contradiction.

thus the full Hamiltonian is (including all of its properties, one of which is the total energy of the system). Thus we may write the energy as an explicit functional of the density [2]

$$E_{HK}[\rho] = T[\rho] + E_{int}[\rho] + \int V_{ext}(\mathbf{r})\rho(\mathbf{r})d^3\mathbf{r} \quad (1.35)$$

$$\equiv F_{HK}[\rho] + \int V_{ext}(\mathbf{r})\rho(\mathbf{r})d^3\mathbf{r} \quad (1.36)$$

where F_{HK} is a universal (i.e., the same for all systems of electrons) since the kinetic energy and interacting terms of the Hamiltonian operator are the same regardless of the external potential. The content of the second Hohenberg-Kohn theorem is that by minimum of the energy functional in Eq. (1.36) with respect to the density ρ is the true, ground state energy of the interacting system. Furthermore, the density which minimizes this functional is the true, interacting ground state electron density. The above statement can be easily proven by again using the variational principle. We know that the ground state of a system with Hamiltonian $\hat{H}$ and ground-state many-body wavefunction Ψ_0 (with associated ground state density ρ_0) can be written in two ways:

$$E_0 = E_{HK}[\rho_0] = \langle \Psi_0 | \hat{H} | \Psi_0 \rangle \quad (1.37)$$

If we now consider a different density, ρ' , which is constructed from a different many-body wavefunction Ψ' , the variational principle tells us that

$$E' = E_{HK}[\rho'] = \langle \Psi' | \hat{H} | \Psi' \rangle > \langle \Psi | \hat{H} | \Psi \rangle = E_{HK}[\rho_0] = E_0 \quad (1.38)$$

which concludes the proof. If we knew the form of the universal functional F_{HK} (and thus the full energy functional via Eq. (1.36)) then by finding the minimum of that functional with respect to the density we would be able to obtain the ground state density and the ground state total energy.

We have shown, and proven, the two results that are present in the Hohenberg-Kohn theorems. However, notice that we have only learned some general truths about transforming the many-body problem from a picture in which the many-body wavefunction is the fundamental quantity of interest to one in which the electronic density is. We still have no idea what this functional F_{HK} looks like, and thus we cannot begin to attempt to minimize it with respect to the density.

1.6.2 Kohn-Sham equations

Walter Kohn and his postdoc student Lu Jeu Sham introduced a set of equations in 1965, known as the Kohn-Sham equations, which suggested a way by which we might proceed in solving for the interacting ground state density and energy [19, 20].

As Kohn said in his 1998 Nobel Prize lecture, the aim was to extract Hartree-like equations from the Hohenberg-Kohn functional for the total energy expressed in terms of the density [20]. The reasoning for being interested in a Hartree-like theory is that he wanted to work in an independent particle system. If this were the case the kinetic energy, which was difficult to approximate in terms of the density in the Thomas-Fermi-Dirac theory discussed above, could simply be constructed *exactly* via the expression

$$T = -\frac{\hbar^2}{2m} \sum_i^\infty f_i \langle \phi_i | \nabla^2 | \phi_i \rangle \quad (1.39)$$

where f_i is the occupation number of the wavefunction ϕ_i [1, 2]. Of course, the system is *not* interacting. But Kohn and Sham hoped to construct an auxiliary independent particle system which had the same density as the true, interacting system. If we could do this, then we could also obtain the true total energy of the system as well by use of the F_{HK} functional given above by Eq. (1.35) and Eq. (1.36).⁹

What Kohn-Sham did is rewrite the expression for $F_{HK}[\rho]$ as being equal to

$$F_{HK}[\rho] = F_{KS}[\rho] = T_s[\rho] + \frac{1}{2} \int \int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3\mathbf{r} d^3\mathbf{r}' + E_{xc}[\rho] \quad (1.40)$$

$$= T_s[\rho] + E_{Hartree}[\rho] + E_{xc}[\rho] \quad (1.41)$$

where $T_s[\rho]$ is the kinetic energy of the particles in the auxiliary, non-interacting Kohn-Sham system [18]. The second contribution is the standard electrostatic expression for a charge density interacting with itself, the Hartree energy $E_{Hartree}$. The exchange-correlation energy, E_{xc} , is *defined* by Eq. (1.40).

Using Eq. (1.40) along with Eqs. (1.35) and (1.36), we can rewrite the exchange-correlation energy in an illuminating way

$$E_{xc}[\rho] = \langle T \rangle - T_s[\rho] + \langle E_{int} \rangle - E_{Hartree}[\rho]. \quad (1.42)$$

Thus E_{xc} is simply the difference between the true kinetic energy of the many-body system and the kinetic energy of the non-interacting system plus the difference between the true interacting energy of the many-body system and the Hartree energy. This is a useful strategy because typically the difference between the interacting kinetic energy and the non-interacting kinetic energy is small. Similarly, the dominant

⁹“But we don’t know the functional form of F_{HK} ” you might object. Yes, that is true. Our ignorance of this functional prevents us from obtaining the correct total energy if we were simply given, say from an alien civilization, the interacting density. However, if we were actually able to construct the auxiliary system with the same density as the interacting system then we would have had to already solve for what the functional F_{HK} was. In practice, the auxiliary Kohn-Sham system *approximates* the full, interacting charge density. Using the same approximations, we then can then obtain our approximate total energy. This should become more clear shortly when we introduce the exchange-correlation energy E_{xc} .

contribution to the electron-electron interaction energy is the Hartree energy [1]. The introduction of the auxiliary Kohn-Sham system sets up a new problem in place of the old where we can calculate exactly the dominant contributions to the kinetic energy and the electron-electron interaction energy and then make our best attempt to approximate the remainder (i.e., E_{xc}). If good approximations to E_{xc} can be obtained¹⁰, then this strategy becomes very promising.

If we have the exact E_{xc} then the equality of Eq. (1.40) holds, equating the Kohn-Sham universal functional F_{KS} with that of the Hohenberg-Kohn F_{HK} . By the Hohenberg-Kohn theorems we know that we can find the ground state density and energy by minimizing the functional E_{HK} with respect to ρ . Since F_{KS} is equivalent to F_{HK} , then we know that if we minimize the energy $E_{KS} = F_{KS} + \int V_{ext}\rho(\mathbf{r})d^3\mathbf{r}$ then we will also obtain the correct ground state density and energy. Taking the functional derivative of E_{KS} with respect to the density, subject to the constraint that the number of electrons remains fixed, we obtain the Kohn-Sham equations:

$$\left[-\frac{\hbar^2}{2m} + V_{ks}[\rho](\mathbf{r}) \right] \phi_i(\mathbf{r}) = \varepsilon_i \phi_i(\mathbf{r}) \quad (1.43)$$

where ϕ_i are the Kohn-Sham wavefunctions and V_{ks} is the Kohn-Sham potential, given as

$$V_{ks}[\rho](\mathbf{r}) = V_{ext}(\mathbf{r}) + V_{Hartree}[\rho](\mathbf{r}) + V_{xc}[\rho](\mathbf{r}) \quad (1.44)$$

$$V_{Hartree}[\rho](\mathbf{r}) = \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3\mathbf{r}' \quad (1.45)$$

$$V_{xc}[\rho](\mathbf{r}) = \frac{\delta E_{xc}[\rho]}{\delta \rho(\mathbf{r})} \quad (1.46)$$

where in the last line it can be seen that the exchange-correlation potential is simply the functional derivative of the exchange-correlation energy with respect to the density.

In the Kohn-Sham system the particles are non-interacting and thus the many-body wavefunction (which we will never explicitly care about) could be constructed in the form of a Slater determinant of one-particle orbitals. From this it follows that the density of the system can be calculated as

$$\rho(\mathbf{r}) = \sum_i^{occ} |\phi_i(\mathbf{r})|^2 \quad (1.47)$$

where the sum extends over all of the occupied orbitals. The Kohn-Sham wavefunctions are the solutions of the Kohn-Sham equations, which depend on the density.

¹⁰In practice there are a number of generally good approximations available, e.g. the local density approximation (LDA), the generalized gradient approximation (GGA), etc.

The density, in turn, depends on the Kohn-Sham wavefunctions as in Eq. (1.47). These equations are then solved by self-consistent iteration, the cycling ending when some criterion has been met (typically when the total energy changes by less than some specified small amount).

Once the self-consistent cycle has converged, the total energy is evaluated. In the Kohn-Sham system this is written as [9, 21]¹¹

$$E = \sum_i^{occ} \int \phi_i(\mathbf{r})^* \frac{\hbar^2}{2m} \phi_i(\mathbf{r}) d^3\mathbf{r} + \int V_{ext}(\mathbf{r}) \rho(\mathbf{r}) d^3\mathbf{r} + \frac{1}{2} \int \int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3\mathbf{r} d^3\mathbf{r}' + E_{xc}[\rho] \quad (1.48)$$

$$= \sum_i^{occ} \varepsilon_i - \frac{1}{2} \int \int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3\mathbf{r} d^3\mathbf{r}' - \int V_{xc}(\mathbf{r}) \rho(\mathbf{r}) d^3\mathbf{r} + E_{xc}[\rho] \quad (1.49)$$

where the second expression can be obtained using the Kohn-Sham equation and the relationship [21]

$$\sum_i^{occ} \varepsilon_i = \sum_i^{occ} \langle \phi_i | -\frac{\hbar^2}{2m} \nabla^2 + V_{ks}[\rho](\mathbf{r}) | \phi_i \rangle. \quad (1.50)$$

1.7 Conclusion

In the above sections we have laid out some of the important foundations for the work that is to follow. Density functional theory, with the aid of *ab initio* pseudopotentials, is the basis for the majority of the work. In addition to the calculation of the total energy as described in Eq. (1.48) and Eq. (1.49) and the electronic density through Eq. (1.47), the results of DFT can be further used to calculate other properties that depend on the solutions to the Kohn-Sham equations. For example, by understanding how the electronic density changes with respect to perturbations of the nuclear coordinates, we can calculate the Hessian of the Born-Oppenheimer surface [22]¹² This Hessian can be used to calculate the dynamical matrix which can be diagonalized to find the vibrational frequencies of a material. We will see examples of these types of extensions to DFT in Chapter 6, where the lattice dynamical properties of a number of low-pressure polytypes of Ge are calculated. Other extensions to DFT that will be discussed in this work include Green's function methods which can be used to get accurate quasiparticle and optical absorption spectra. This type of information is not available within pure DFT because, as we have discussed, DFT is a ground state theory which gives us information about the ground state charge

¹¹Typically the repulsive ion-ion term E_{II} is included as well as this is clearly required if one is interested in geometry optimization or any other calculation in which the ions are taken to move.

¹²The Born-Oppenheimer energy surface is simply another name for the total-energy of the system with fixed ionic positions, e.g. Eq. (1.49) with the addition of the ion-ion E_{II} term.

density and the ground state energy, but *not* information on excited states. We will begin our discussion of the quasiparticle energy spectra in the next chapter, and will show applications of these methods to optical properties starting in Chapter 4.

Chapter 2

Fully *ab initio* quasiparticle spectra of semiconductors

2.1 Introduction

In the past few decades, there has been tremendous progress in the first-principles calculation of material properties. The availability of highly parallelized codes within the framework of density functional theory (DFT) [17, 19, 20] as well as efficient approximations to the exchange-correlation energy E_{xc} [23–26] allow for the reliable prediction of many properties of materials, especially those which are related to ground-state properties. For a large class of materials cohesive energies, equilibrium crystal structures, lattice dynamical properties, elastic constants, etc. can be calculated from first principles with an accuracy high enough to be of great predictive value [22, 27–29].

The “band gap” problem of DFT [30, 31] is perhaps the most obvious and permeating example of when such approaches applicable to the ground-state fail. Band gaps evaluated as a difference of DFT eigenvalues are routinely substantially lower than those measured experimentally [32]. This is because the Kohn-Sham eigenvalues, entering into the variational equations only as Lagrange multipliers in the fictitious Kohn-Sham auxiliary system, have no rigorous meaning as quasiparticle energies. Correcting such deficiencies has understandably been a major goal of theoretical work because an accurate representation of the quasiparticle spectrum is often necessary to understand and reliably predict a host of other properties that are of immense scientific and technological interest. This includes the calculation of optical absorption [33], band alignments in semiconductor heterojunctions [34–36], the placement of defect levels [37, 38], quantum transport [39], as well as many others.

Fortunately, accurate quasiparticle spectra can be calculated with knowledge of the electron self-energy operator Σ . While the complete determination of the electron self-energy is equivalent to the complete solution of the many-body problem

and thus is impractical, there exist computationally efficient approximations to the self-energy which result in energies that remedy much of the deficiency found in the use of the DFT eigenvalues. By far the most well-known approach is that which approximates the self-energy as the first-order term in an expansion of the screened Coulomb interaction W as $\Sigma \approx iGW$, where G is the single-particle Green's function [32, 40–42]. This evaluation of this approximation, which is already considerably more simple than the full self-consistent solution of Hedin's equations including vertex corrections, still presents a formidable challenge. The calculation of the Σ operator includes large sums over empty states and reciprocal lattice vectors, q -point summations, a frequency integration to account for the dynamical effects, and in principle a self-consistent evaluation of the self-energy. In addition to this, in implementations of the GW approximation one might also find a number of other factors that could possibly affect the results, such as the use of the pseudopotential approximation, incomplete basis sets, the separation between core and valence electrons, etc.

The difficulty of the calculations often necessitate taking additional approximations beyond that taken by setting the vertex function in Hedin's equations to unity. As a result, there is often much ambiguity as to whether the results are indicative of the method generally or whether it is a result of the other approximations used, their implementation, or even of the convergence. The resolution of these issues is important not only to determine which approximations are most in need to relaxation in order to improve agreement with experiment (and thus where efforts should be focused as the availability of computational resources continues to grow). But in addition, from a practical perspective of predicting materials properties, it is also of interest to understand which methodological recipe will best reproduce the experimental properties regardless of its theoretical purity (i.e., even if such agreement is only due to a cancellation of errors).

The motivation behind this work is to present highly precise first-principles calculations of the quasiparticle spectra of group-IV semiconductors and III-V compounds within a chosen set of approximations. Specifically, we consider the Si and Ge in the diamond structure, as well as the zincblende structures of AlP, AlAs, AlSb, GaP, GaAs, GaSb, InP, InAs, and InSb. In relation to the above discussion, our calculation is done utilizing the pseudopotential method within a plane wave basis set [11]. This involves approximations relating the partitioning between core and valence electrons as well as the use of smooth, nodeless pseudo-wavefunctions rather than that of their all-electron counterparts. In situations where there are low-lying semicore electrons that are known to be of importance, we include these explicitly into the valence to take into account their effect on the bandstructure in the vicinity of the band gap. We do this for all GaX and InX systems (where X=P,As,Sb). Additionally, in order to take into account the dynamical screening we employ a generalized plasmon-pole model which extends the first principles calculated static dielectric function to finite frequencies [42]. Finally, we evaluate the quasiparticle energies within the “one-shot” GW approach, also known as G_0W_0 . The effects of self-consistency on the band gaps

has been seen to result in band gaps that are too large in comparison to experiment, the rationale for this assumed to be the neglect of electron-hole correlations which would result in stronger screening, thus reducing the obtained quasiparticle gap [43]. In fact recently it has been shown that the inclusion of vertex corrections, limited in the construction of W only, results in improved agreement with experiment for band gaps within self-consistent approaches [44]. The use of the plasmon pole model is believed to be a good approximation, especially in the vicinity of the band gap, with differences from the full frequency results of only a few tens of meV [32]. The use of the pseudo-wavefunctions, rather than those of the all-electron calculations, has recently been questioned [45] as being possibly problematic. However, the evaluation of these claims is outside the scope of the present work. The intent of this work is to provide demonstrably converged results for the quasiparticle spectra of group IV and III-V semiconductors within the PP- G_0W_0 methodology, which will not only be useful in the assessment of their accuracy in reproducing experimental gaps, but also as benchmark calculations that can be compared with other methods. In addition to providing the quasiparticle gaps and energies at the high-symmetry points, we will also present full bandstructures including the calculated spin-orbit splitting for all studied materials.

2.2 Method details

2.2.1 DFT

The calculation of the self-energy Σ proceeds as a perturbation to an appropriate mean-field starting point. In this work we use wavefunctions, energies, and exchange-correlation potentials calculated with the local density approximation (LDA) to density functional theory (DFT). The DFT calculations are obtained with the **Quantum-ESPRESSO** package [46], which utilizes a planewave basis set [11]. Pseudopotentials for all of the elements are generated with the **APE** pseudopotential generator [47]. For the majority of the elements considered in this study, there is a clear and substantial separation between the electrons that are typically treated as valence with those treated as core states. This is not the case for Ga or In, however, which are known to have semicore d states which strongly affect the electronic structure [48–50]. Moreover, it has been previously found that the inclusion of a semicore state of one angular-momentum quantum number into the valence while keeping the rest of the electrons in the same shell in the core leads to poor quasiparticle energies. This has to do with the spatial localization between semicore d states and their same-shell s and p neighbors, which lead to large exchange energies [51]. In order to avoid this potential problem, we have included the full semicore shell into the valence when treating systems with either the cation Ga or In. These semicore pseudopotentials are generated using the multireference pseudopotential (MRPP) scheme as outlined by Reis *et al.*

Table 2.1: Shown is the energy cutoff for the wavefunctions, E_{cut} , along with the theoretically predicted lattice constant and bulk modulus in comparison to experiment. The In and Ga pseudopotentials incorporate the full semicore shell as valence. The subsequent GW calculations are always done at the experimental lattice constant.

	E_{cut} (Ry)	Theory		Expt. ^a	
		a_0 (Å)	B_0 (GPa)	a_0 (Å)	B_0 (GPa)
Si	60	5.39	94.7	5.43	97.8
Ge	40	5.54	80.2	5.66	75.8
AlP	70	5.43	86.7	5.46	86
AlAs	80	5.63	73.1	5.66	78.1
AlSb	70	6.09	55.2	6.14	55.1
InP	350	5.82	67.9	5.87	71
InAs	350	6.01	59.7	6.06	60
InSb	350	6.40	48.0	6.48	48.3 ^b
GaP	700	5.39	88.0	5.45	89
GaAs	700	5.60	71.6	5.65	74.8
GaSb	700	6.01	55.3	6.10	57

^a Ref. 53 ^b Ref. 54

[52] and as implemented in the **APE** code [47]. All pseudopotentials used have been tested carefully for convergence in terms of the planewave energy cutoff as well as in reproducing the experimental lattice constants and bulk moduli. This information is given in Table 2.1 and indicates the quality of the pseudopotentials employed. The introduction of semicore states dramatically increases the computational cost associated with the calculations, not only in the generation of the wavefunctions, but also in the subsequent GW calculation.

2.2.2 GW calculations

Our GW calculations are done using the publicly-available **BerkeleyGW** [55] package which interfaces with a variety of DFT codes. As mentioned in the introduction, the calculation of the self-energy involves summations over empty states, q -point summations, as well as summations over reciprocal lattice vectors. These components can be seen in the expressions for $\Sigma = \Sigma_{SX} + \Sigma_{CH}$, where the self-energy is broken up into the usual screened-exchange and Coulomb-hole contributions [32, 55], as

$$\begin{aligned}
 \langle n\mathbf{k} | \Sigma_{SX}(E) | n'\mathbf{k} \rangle = & - \sum_{n''}^{occ} \sum_{\mathbf{q}\mathbf{G}\mathbf{G}'} M_{n''n}^*(\mathbf{k}, -\mathbf{q}, -\mathbf{G}) M_{n''n'}(\mathbf{k}, -\mathbf{q}, -\mathbf{G}') \\
 & \times \left[\delta_{\mathbf{G}\mathbf{G}'} + \frac{\Omega_{\mathbf{G}\mathbf{G}'}^2(\mathbf{q})(1 - i \tan \phi_{\mathbf{G}\mathbf{G}'}(\mathbf{q}))}{(E - E_{n''\mathbf{k}-\mathbf{q}})^2 - \tilde{\omega}_{\mathbf{G}\mathbf{G}'}^2(\mathbf{q})} \right] v(\mathbf{q} + \mathbf{G}')
 \end{aligned} \tag{2.1}$$

and

$$\begin{aligned} \langle n\mathbf{k} | \Sigma_{CH}(E) | n'\mathbf{k} \rangle &= \frac{1}{2} \sum_{n''}^{\infty} \sum_{\mathbf{q}\mathbf{G}\mathbf{G}'} M_{n''n}^*(\mathbf{k}, -\mathbf{q}, -\mathbf{G}) M_{n''n'}(\mathbf{k}, -\mathbf{q}, -\mathbf{G}') \\ &\times \frac{\Omega_{\mathbf{G}\mathbf{G}'}^2(\mathbf{q})(1 - i \tan \phi_{\mathbf{G}\mathbf{G}'}(\mathbf{q}))}{\tilde{\omega}_{\mathbf{G}\mathbf{G}'}(\mathbf{q})(E - E_{n''\mathbf{k}-\mathbf{q}})^2 - \tilde{\omega}_{\mathbf{G}\mathbf{G}'}^2(\mathbf{q})} v(\mathbf{q} + \mathbf{G}') \end{aligned} \quad (2.2)$$

where v is the bare Coulomb interaction, and

$$M_{nn'}(\mathbf{k}, \mathbf{q}, \mathbf{G}) = \langle n\mathbf{k} + \mathbf{q} | e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}} | n'\mathbf{k} \rangle \quad (2.3)$$

$$\Omega_{\mathbf{G}\mathbf{G}'}^2 = \omega_p^2 \frac{(\mathbf{q} + \mathbf{G}) \cdot (\mathbf{q} + \mathbf{G}')}{|\mathbf{q} + \mathbf{G}|^2} \frac{\rho(\mathbf{G} - \mathbf{G}')}{\rho(\mathbf{0})} \quad (2.4)$$

$$\tilde{\omega}_{\mathbf{G}\mathbf{G}'}^2(\mathbf{q}) = \frac{|\lambda_{\mathbf{G}\mathbf{G}'}(\mathbf{q})|}{\cos \phi_{\mathbf{G}\mathbf{G}'}(\mathbf{q})} \quad (2.5)$$

$$|\lambda_{\mathbf{G}\mathbf{G}'}(\mathbf{q})| e^{i\phi_{\mathbf{G}\mathbf{G}'}(\mathbf{q})} = \frac{\Omega_{\mathbf{G}\mathbf{G}'}^2(\mathbf{q})}{\delta_{\mathbf{G}\mathbf{G}'} - \epsilon_{\mathbf{G}\mathbf{G}'}^{-1}(\mathbf{q}; 0)}. \quad (2.6)$$

In the above equations ρ is the reciprocal space charge density and ω_p is the usual plasma frequency, $\sqrt{4\pi\rho(\mathbf{0})e^2/m}$. In these equations $\epsilon(\mathbf{q}; 0)$ is the static dielectric matrix, which is itself related to a summation over wavevectors and empty states via its relationship with the static polarizability χ $\epsilon_{\mathbf{G}\mathbf{G}'}(\mathbf{q}; 0) = \delta_{\mathbf{G}\mathbf{G}'} - v(\mathbf{q} + \mathbf{G})\chi_{\mathbf{G}\mathbf{G}'}(\mathbf{q}; 0)$.

where χ is given as

$$\begin{aligned} \chi_{\mathbf{G}\mathbf{G}'}(\mathbf{q}; 0) &= \\ \sum_n^{occ} \sum_{n' \notin occ}^{\infty} \sum_{\mathbf{k}} M_{nn'}(\mathbf{k}, \mathbf{q}, \mathbf{G}) M_{nn'}^*(\mathbf{k}, \mathbf{q}, \mathbf{G}') \frac{1}{E_{n\mathbf{k}+\mathbf{q}} - E_{n'\mathbf{k}}} \end{aligned} \quad (2.7)$$

There are a number of convergence parameters in the equations above, and they cannot all be tested independently of one another. For example, in Eq. 2.2 the summation over reciprocal lattice vectors and empty states is not independent because the contributions from higher energy conduction band states have a large part of their contribution resulting from their matrix element corresponding to reciprocal vectors of higher energy. In practice, what this means is that convergence tests with respect to the number of empty states in the Coulomb-Hole summation cannot be done independently of the sum over $\mathbf{G}$ vectors (and hence independently of the size of the calculated dielectric matrix $\epsilon_{\mathbf{G}\mathbf{G}'}$). This can be seen from examining Fig. 2.1 and Fig. 2.2 where we have shown the convergence of the matrix element of Σ_{CH}

in Si for a $\mathbf{G}$ vector cutoff of 40 Ry and 10 Ry, respectively, in Eq. 2.2. Shown are the convergence of the first 8 bands in Si (4 valence and 4 conduction) for a grid of $6 \times 6 \times 6$. It is clearly seen that with different sizes of $\mathbf{G}$ vector summations, hereafter interchangeably referred to as the energy cutoff for the dielectric matrix, we come to different conclusions about the number of bands needed for convergence. In the case of a 40 Ry cutoff, the number of bands required is slightly over 600 whereas in the case of the 10 Ry cutoff the number of bands to reach the same level of convergence is less than 300. The difference in the predicted direct band gap in these two cases is about 50 meV which is not negligible. In the literature it has become more common to show convergence with respect to the number of empty states in Σ_{CH} which is positive development. However, there is frequently an absence of discussion or evidence that the dielectric matrix cutoff is converged. As Fig. 2.1 and Fig. 2.2 illustrate, without a consideration of the dielectric matrix cutoff this convergence with respect to number of bands is meaningless and may mislead the community to grant a higher degree of precision to the calculations than is warranted.

In this work we are very explicit in the convergence of the sums present in Eqs. 2.1, 2.2, and 2.7. For all materials studied we have included figures which demonstrate the convergence of the summations over bands in the Eq. 2.2 as well as those in the polarizability as seen in Eq. 2.7. As per the above discussion, we have also included the convergence tests with regard to the dielectric matrix cutoff. The procedure that we follow in these convergence tests is the following:

1. We construct the polarizability (and thus the dielectric matrix) up to a high energy cutoff. This is chosen such that lowering the cutoff should will not affect the quasiparticle eigenvalues in the subsequent self-energy calculation in any meaningful way (in other words, it is converged). This number is not known in advance, and whether or not a high enough energy cutoff was chosen initially will be determined later. The band summation in Eq. 2.7 is chosen to be the number of empty states equivalent to the number of $\mathbf{G}$ vectors implied by that dielectric energy cutoff. Thus in this step we hope to construct a dielectric matrix of size which we expect to be larger than needed to converge subsequent Σ calculations with individual matrix entries that are themselves very close to the values that would be obtained with an infinite band summation.
2. With the dielectric matrix constructed as in Step 1, which because it is expected to be highly converged, should be equivalent, to any meaningful precision in the quasiparticle eigenvalues, to a dielectric matrix of infinite size constructed with infinite bands. We can now test the band summation in Eq. 2.2 without worrying that the results of these tests are dependent upon incomplete convergence with respect to the other parameters. These tests, carried out until all quasiparticle eigenvalues (taken relative to the value at the top of the valence band maximum (VBM) at Γ), are converged to within 5 meV gives us the value of bands in Σ_{CH} that we will use in subsequent production runs.

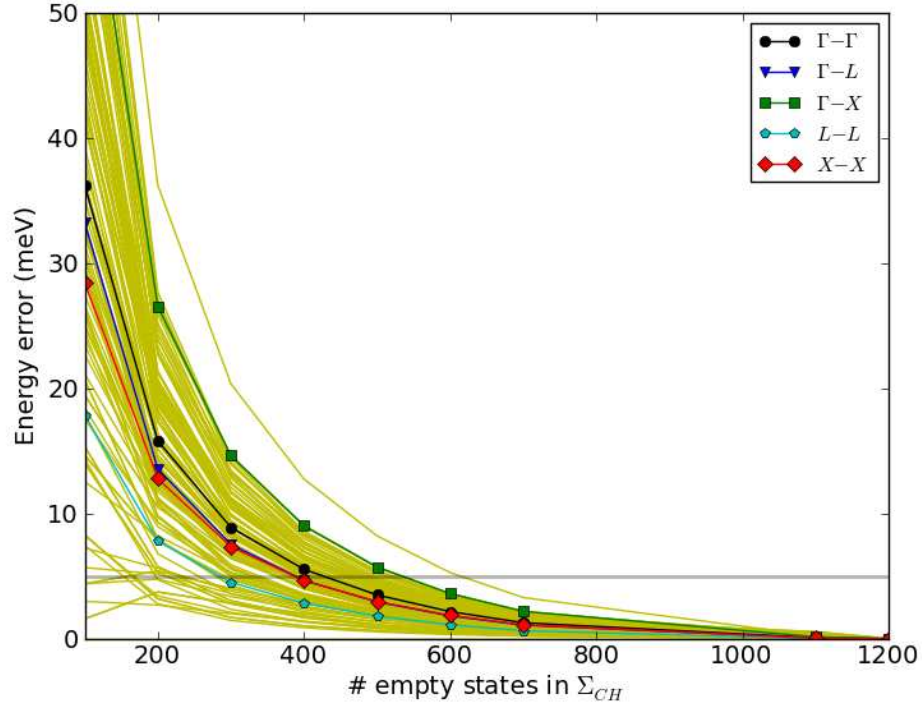


Figure 2.1: Convergence in Si with respect to the number of bands included in the Σ_{CH} summation. The summation over $\mathbf{G}$ vectors in Eq. 2.2 is cut off at 40 Ry. Shown are the calculated deviations from the converged result for 8 bands in Si (4 valence and 4 conduction) on a $6 \times 6 \times 6$ grid. Direct and indirect band gaps of particular interest are singled out in different colors as shown in the legend. As can be seen, it takes slightly more than 600 unoccupied states to converge all eigenvalues to less than 5 meV.

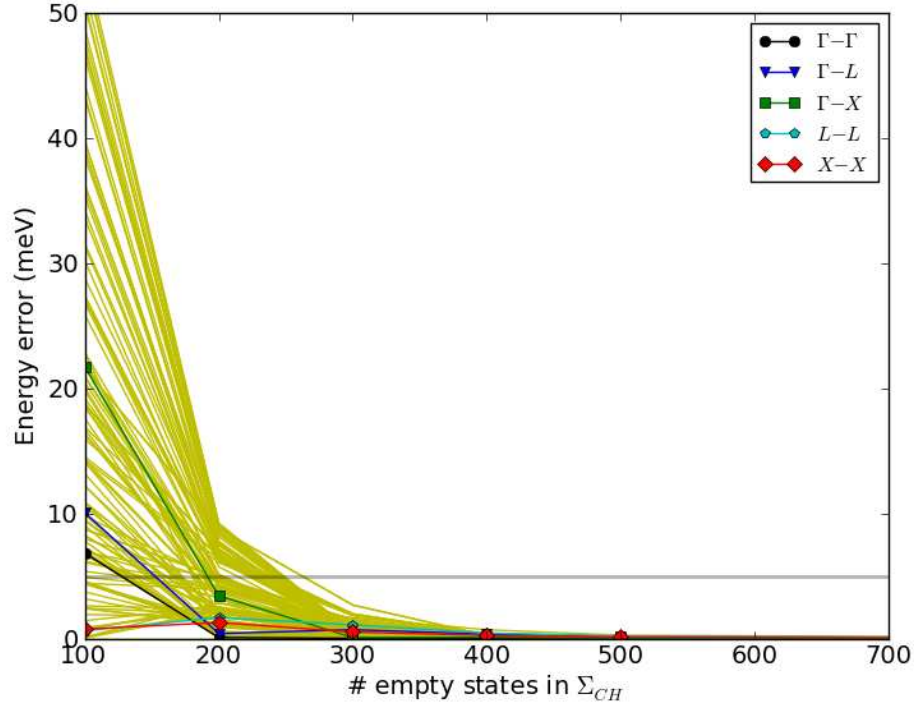


Figure 2.2: Convergence in Si with respect to the number of bands included in the Σ_{CH} summation. The summation over $\mathbf{G}$ vectors in Eq. 2.2 is cut off at 10 Ry. Shown are the calculated deviations from the converged result for 8 bands in Si (4 valence and 4 conduction) on a $6 \times 6 \times 6$ grid. Direct and indirect band gaps of particular interest are singled out in different colors as shown in the legend. As can be seen, it takes less than 300 unoccupied states to converge all eigenvalues to less than 5 meV.

3. Having determined the number of bands needed needed in the construction of Σ_{CH} in Step 2, we can now ask how many $\mathbf{G}$ vectors are needed in the sums in Σ . We expect that the quasiparticle energies are not sensitive to some fraction of the higher energy $\mathbf{G}$ vector contributions that we have available from our calculation in Step 1. So here we carry out another set of Σ calculations, similar to the ones in Step 2, but now with varying the dielectric matrix cutoff. Again, when the quasiparticle eigenvalues relative to the top of the VBM are less than 5 meV, we take the resulting dielectric matrix cutoff as our converged value for subsequent production runs.
4. Finally, now that we understand the convergence of the quasiparticle eigenvalues with respect to the number of bands in Σ , as well as the size of the needed dielectric matrix, we determine how many bands are actually needed to construct a dielectric matrix of that size without introducing extra error in the quasiparticle energies. Again, we expect that the number of bands needed will be less than those used in the Step 1 in which we were attempting to accurately represent the dielectric matrix components of a larger dielectric matrix. Here we construct a number of dielectric matrices, obtained by varying the number of bands included in the summation in Eq. 2.7, and then use these different dielectric matrices to compute the quasiparticle energies. We examine the quasiparticle energies rather than simply look at the convergence of the components of the dielectric matrix because we are only interested in the convergence of components that strongly affect the final energies. From these calculations, again with a 5 meV criterion for specifying the convergence, we obtain the number of bands needed in the summation in Eq. 2.7. This is a useful step to do, as it turns out that for all systems studied that the number of bands needed in Eq. 2.7 is substantially less than that needed in Eq. 2.2. This observation has been made before [56] in relation to approximate methods to converge the Coulomb-hole summations in Σ which appear to be the dominant computational expense in terms of empty state summations.

We have yet to discuss the convergence of the quasiparticle energies with respect to the q -point summations in Eq. 2.1 and Eq. 2.2, nor the k -point summation in Eq. 2.7. In principle the above procedure should be carried out at converged values with respect to these summations. To do this would entail a substantial convergence testing overhead for any calculation. However, it might be expected that the converged parameters with regards to summations over bands and $\mathbf{G}$ vectors might not depend strongly on the $\mathbf{k}/\mathbf{q}$ point. We have explicitly tested this expectation on Si, where the above procedure was carried out on $\mathbf{k}/\mathbf{q}$ grids of size $6\times 6\times 6$, $4\times 4\times 4$, and $1\times 1\times 1$. In all three cases the decisions with respect to the number of bands and the $\mathbf{G}$ summations were nearly identical. This motivates a use of the above procedure on smaller $\mathbf{k}/\mathbf{q}$ grids for the convergence testing of unoccupied states and dielectric matrix cutoff and then to use these parameters in the final calculation. In practice,

for the materials with lower wavefunction cutoff (Si, Ge, AlP, AlAs, and AlSb) we carry out these tests on a $2 \times 2 \times 2$ grid¹ while the calculations involving semicore states (InX and GaX compounds) are done on a $1 \times 1 \times 1$ grid to save computational resources. The claim that a $1 \times 1 \times 1$ (i.e., 1^3) grid gives similar convergence parameters as would be obtained on larger grids may at first seem suspect, since it is known that different gaps show different convergence characteristics, such as the indirect gaps versus the direct gaps in GaAs [48]. However, even in the case of the 1^3 grid we converge the quasiparticle energies for 8 bands around the gap (4 valence and 4 conduction) relative to the VBM. In this sense we are already converging gaps between states of different character and thus the tests represent a more stringent set of tests than if we simply converged the direct $\Gamma - \Gamma$ gap. And lastly, the small energy difference criterion of 5 meV is a tight tolerance for convergence. That the relative gaps of 8 bands at Γ should all be converged within 5 meV makes it unlikely that there exists a gap between different k -points with a much different degree of convergence.

Finally, with respect to the $\mathbf{k}/\mathbf{q}$ summations themselves, for the materials with lower wavefunction cutoff we use a grid of size 8^3 for both the summation over $\mathbf{k}$ and over $\mathbf{q}$. In Si a grid of 8^3 results in band gap errors for the Γ - Γ , Γ -X, and Γ -L gap of only 2, 2, and 4 meV, in comparison to that obtained with a grid of 12^3 . For the other materials without semicore states we use an 8^3 grid in expectation that the convergence is not substantially different from that obtained in Si. For materials with semicore states an 8^3 calculation is extremely computationally demanding and so for these materials we use a smaller 4^3 grid². Finally, we note that in number of the materials (Ge, InAs, InSb, and GaSb) DFT incorrectly predicts a band overlap at Γ . Upon performing a GW calculation, these gaps are lifted and a semiconductor is obtained. However, the resulting quasiparticle energies reflect inaccuracies because of the incorrect occupation of the states in the sum of Eq. 2.1. Thus, for states where DFT predicts a metal, we first calculate the quasiparticle energy for states with incorrect occupation (i.e., at Γ) and then use this information to correct the eigenvalues only at Γ such that the occupations of the considered states are now correct. With the exception of this eigenvalue update to correct for the overlap within DFT, we never iterate on the the quasiparticle eigenvalues and do only a “single-shot” GW calculation, G_0W_0 .

2.2.3 Spin-orbit interaction

In this work many of the materials under study are relatively heavy and thus have non-negligible degrees of spin-orbit interaction. If we were simply interested in the gap values between high-symmetry points, we could simply use experimental numbers to adjust the GW results before comparison. However, we would like to calculate

¹For Si we have already mentioned tests carried out at $6 \times 6 \times 6$, so in this chapter we will present those figures instead of those on the $2 \times 2 \times 2$ grid.

²In Si, a 4^3 grid results in band gap errors of about 50 meV with respect to that of a 12^3 grid.

the full bandstructure throughout the zone and thus we need much more detailed information on the effects of spin-orbit. To accomplish this, we use the vector-part of the pseudopotential to treat spin-orbit as a perturbation on top of the GW results [57], using the LDA wavefunctions but the GW determined quasiparticle energies. This is done via a Wannier-interpolation of the spin-orbit matrix. First, we utilize the **Wannier90** code [58] to generate the relevant unitary matrices $\mathbf{U}^{(\mathbf{k})}$ and $\mathbf{U}^{dis(\mathbf{k})}$ which relate the Bloch wavefunctions with the associated Wannier functions. For all materials studied we use a uniform 8^3 grid to generate the Wannier functions, thus requiring that we calculate the quasiparticle eigenvalues and the spin-orbit matrix on 8^3 grids as well ³. First we rotate the Hamiltonian into the basis of rotated Bloch states as [58]

$$H^W(\mathbf{k}) = (\mathbf{U}^{(\mathbf{k})})^\dagger (\mathbf{U}^{dis(\mathbf{k})})^\dagger H(\mathbf{k}) \mathbf{U}^{dis(\mathbf{k})} \mathbf{U}^{(\mathbf{k})} \quad (2.8)$$

where we have neglected the band and spin indices. In the equation above

$$H_{n\sigma, n'\sigma'}(\mathbf{k}) = H_{nn'}^{GW}(\mathbf{k}) \delta_{\sigma, \sigma'} + H_{n\sigma, n'\sigma'}^{SO}(\mathbf{k}) \quad (2.9)$$

$$= E_{n\mathbf{k}}^{QP} \delta_{nn'} \delta_{\sigma, \sigma'} + H_{n\sigma, n'\sigma'}^{SO}(\mathbf{k}). \quad (2.10)$$

With the generalization of the rotation matrices U to [59, 60]

$$U_{n\sigma, n'\sigma'}^{\mathbf{k}} = U_{nn'}^{\mathbf{k}} \delta_{\sigma\sigma'} \quad (2.11)$$

$$U_{n\sigma, n'\sigma'}^{dis(\mathbf{k})} = U_{nn'}^{dis(\mathbf{k})} \delta_{\sigma\sigma'} \quad (2.12)$$

we can continue with the standard Wannier interpolation procedure and construct the Fourier transform as

$$H(\mathbf{R}) = \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}} e^{-i\mathbf{k} \cdot \mathbf{R}} H^W(\mathbf{k}) \quad (2.13)$$

Because the matrix $H(\mathbf{R})$ corresponds to matrix elements between Wannier functions, these elements decay rapidly with increasing $\mathbf{R}$. This Fourier sum can then be truncated at finite $\mathbf{R}$ but is equivalent in value if we had considered a sum over an infinitely large mesh of $\mathbf{R}$ vectors. This allows us to Fourier interpolate back onto an arbitrary $\mathbf{k}$ -point as

$$H(\mathbf{k}') = \sum_{\mathbf{R}} e^{i\mathbf{k}' \cdot \mathbf{R}} H(\mathbf{R}) \quad (2.14)$$

Diagonalizing the matrix $H(\mathbf{k}')$ then results in our interpolated quasiparticle band-structure including spin-orbit interactions.

³The grid on which we compute quasiparticle eigenvalues does not have to be the same as that used in the $\mathbf{k}/\mathbf{q}$ summations in Σ and χ . The $\mathbf{q}$ grid simply needs to be commensurate with the grid on which we calculate the GW eigenvalues.

Table 2.2: A summary of some of the convergence parameters used in GW calculations, as determined from the tests shown in Appendix A. Shown are the number of bands included in Eq. 2.2 as well as the number included in Eq. 2.7. The size of the dielectric matrix, corresponding to the sum over $\mathbf{G}$ vectors in Eq. 2.1 and Eq. 2.2 is also given. The value of other convergence parameters is discussed in the text.

	$\epsilon_{cutoff}(\text{Ry})$	bands in χ	bands in Σ_{CH}
Si	25	300	600
Ge	25	400	700
AlP	28	500	700
AlAs	30	500	800
AlSb	23	500	800
InP	35	400	1400
InAs	33	400	1300
InSb	33	400	1300
GaP	64	500	1500
GaAs	65	600	1600
GaSb	67	600	1500

2.3 Results

We now present the results of the GW quasiparticle calculations, starting with the group IV materials Si and Ge. For all materials we have included the convergence studies with respect to empty states and dielectric matrix cutoff in Appendix A. A summary of these convergence parameters is shown in Table 2.2.

The calculated quasiparticle energies of Si are shown in Table 2.3. In Si, all of the direct and indirect gaps are well described. The greatest error occurs for the state with lowest energy, Γ_{1v} . In Fig. 2.3 we have shown the band structure along the high-symmetry lines, including the effects of spin-orbit, which because of its maximal value of 0.05 eV, is difficult to discern in the plot. In Si, the fundamental gap is between the valence band maximum at Γ and the conduction band minimum along the Γ -X line. We find this indirect gap to be of value 1.27 eV, comparing well to the experimental value of 1.17 eV [61]. The value of 1.27 eV is also in good agreement with the pioneering work of Hybertsen and Louie [42], who obtained 1.29 eV.

In Ge, we also find remarkably good agreement for the majority of the quasiparticle energies with available experimental values (see Table 2.4). One clear exception, however, is that of the direct gap at Γ , which we underestimate by about 0.5 eV. Previous theoretical works using the same formalism have also underestimated this gap, although typically by less than in our results [42]. Another study on Ge by Rohlffing *et al.* [62], who used a Gaussian-orbital basis, also underestimated this transition, but only by about half the amount seen here. Closer to the value reported here, results by Shirley *et al.* indicate a gap of 0.53 eV for this transition [63]. As

Hybertsen and Louie have discussed [42], the position of the Γ'_{2c} is expected to be underestimated to some extent because this state has the majority of its weight on the atomic sites and thus is most sensitive to the approximate treatment of the valence-core interactions in the self-energy. This underestimation can also be seen in the calculation of the Γ'_{2c} state in Si, the other homopolar material considered in this study. Hybertsen and Louie gave an estimate of this approximate treatment of the valence-core interactions in Ge at a maximal value of 0.3 eV, with a sign which would increase the direct band gap at Γ . In fact, GW calculations performed on Ge with the entire semicore shell included by Tiago *et al.* [48] find a larger direct band gap of 0.85 eV, in excellent agreement with the experimental value. The same effect was found by Rohlfing in an earlier study [49]. Additionally, work done by Shirley *et al.* have also found that a core-polarization potential (CPP) treatment of the core-valence interactions in Ge results in a direct band gap which is increased by over 0.3 eV in comparison to the treatment of the core-valence interactions at the DFT level only [63].

The cause for the additional underestimation in the present work compared to earlier results is not completely clear. Within our tests, we find that the direct gap at Γ is relatively insensitive to the number of bands included in the calculation of the polarizability, Eq. 2.7, or the size of the resulting dielectric matrix, and so we doubt that incomplete convergence in the former calculations with respect to these parameters is likely to be a factor. Even in the summation over empty states for Σ_{CH} , we find that our calculated gap for Γ'_{2c} would only be 20 meV larger if we only included 64 bands as was done in Ref. 42, although we do find other gaps that are more sensitive to this smaller summation over empty states. Part of the discrepancy may be due to the fact that some of early calculations in Ge were done using coarser q -grid meshes. We find that if we reduce the size of our q -grid from 8^3 to 4^3 , as was used in Ref. 42, the location of the Γ'_{2c} rises by 80 meV relative to the VBM. Any additional discrepancy may originate in the sensitivity of the Γ'_{2c} state to the pseudopotential at the LDA level. For example, our Γ'_{2c} state is over 0.12 eV lower in our mean-field calculation than in that of Ref. 62. If these energy biases remain to some extent in the final quasiparticle energies, we would expect to obtain a smaller value. Thus our calculated value of the Γ'_{2c} state seems to be comparable to previous values when we take into account the tighter convergence in the present work. As found in prior studies [48, 49, 63], taking the semicore electrons into account in Ge may be necessary to obtain better agreement with experiment for this direct transition at Γ , which needs to be larger in order for the calculation to predict the indirect band gap of Ge that is seen experimentally. The quasiparticle band structure of Ge is shown in Figure 2.4. The spin-orbit splitting, being much larger in Ge than in Si, can clearly be seen in the figure.

In Table 2.5 we have summarized the quasiparticle energies for AlP, AlAs, and AlSb. Unfortunately, we were unable to find many experimental results with which to compare. For all three compounds, we reproduce the fundamental indirect gap

Table 2.3: Quasiparticle energies for Si. Spin-orbit split values are shown on the same line separated by commas. All energies are in eV. Experimental data is taken from Ref. 61 (and references therein) unless otherwise noted.

	LDA	GW	GW+SO	Expt.
Γ_{1v}	-11.96	-11.61	-11.63	-12.5
Γ'_{25v}	0.0	0.0	-0.05,0.0	-0.044 ,0.0
Γ_{15c}	2.56	3.32	3.28,3.32	3.34 ^a
Γ'_{2c}	3.15	3.83	3.81	4.18
X_{1v}	-7.82	-7.76	-7.78	
X_{4v}	-2.84	-2.84	-2.86	-2.9 ^a
X_{1c}	0.64	1.45	1.43	1.3 ^b
L'_{2v}	-9.62	-9.48	-9.49	-9.3 ^a
L_{1v}	-6.99	-6.92	-6.94	-6.8 ^a
L'_{3v}	-1.19	-1.20	-1.23,-1.20	-1.2 ^a
L_{1c}	1.43	2.15	2.13	2.04 ^a
L_{3c}	3.33	4.19	4.16,4.18	3.9 ^a

^a Ref. 64

^b Estimate from the indirect gap and longitudinal mass as given in Ref. 42.

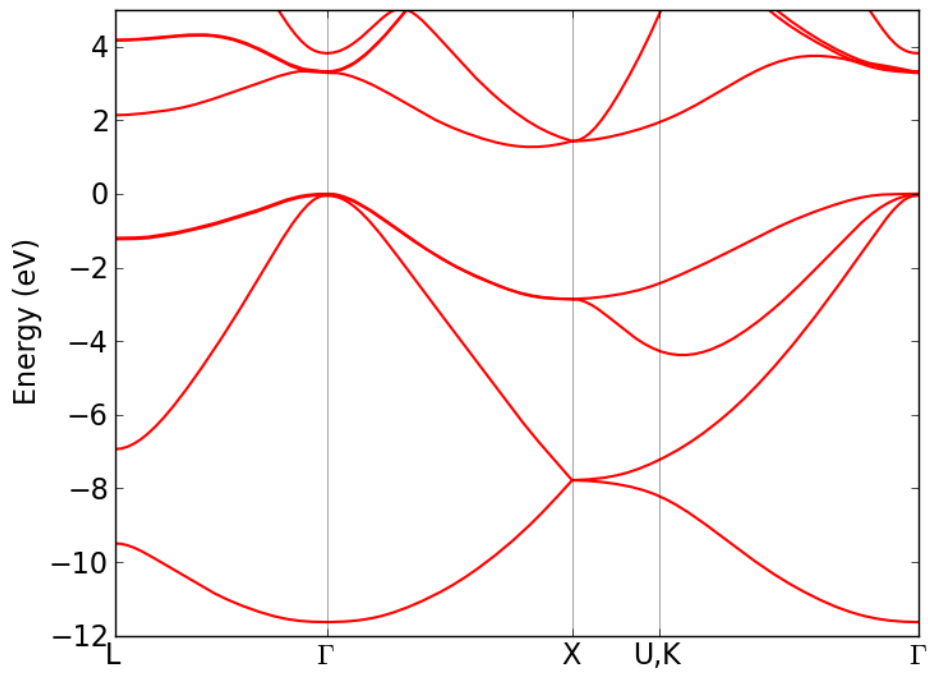


Figure 2.3: Quasiparticle bandstructure of Si, including the effects of spin-orbit. The valence band maximum is set at zero energy. The VBM is located at Γ and the CBM is located along the Γ -X line with magnitude 1.27 eV.

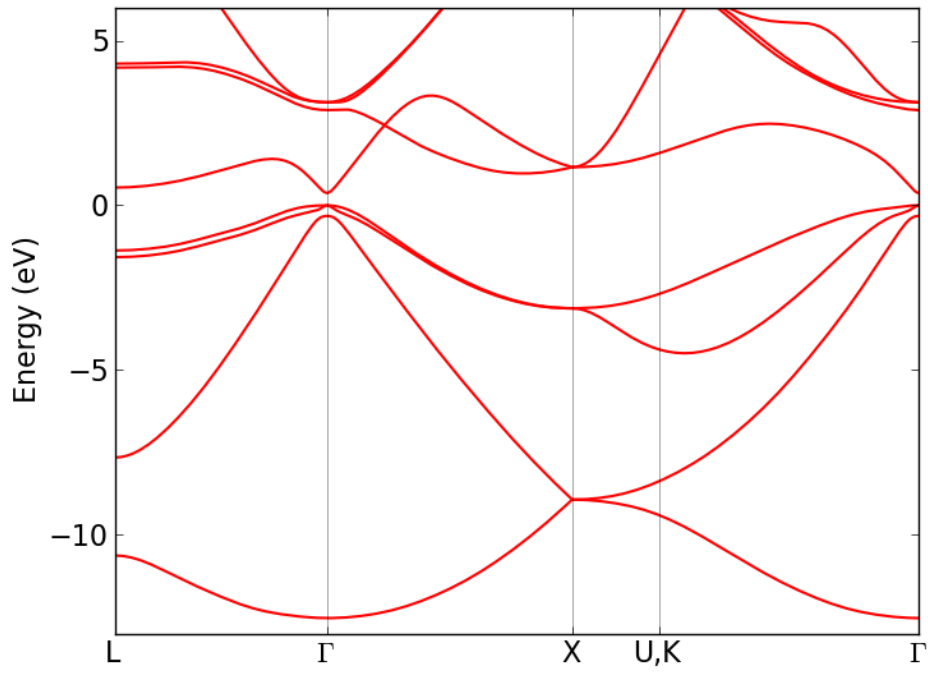


Figure 2.4: Quasiparticle bandstructure of Ge, including the effects of spin-orbit. The valence band maximum is set at zero energy.

Table 2.4: Quasiparticle energies for Ge. Spin-orbit split values are shown on the same line separated by commas. In instances where differing experimental data was reported we have included them separated by a slash (/). All energies are in eV. Experimental data is taken from Ref. 61 (and references therein) unless otherwise noted.

	LDA	GW	GW+SO	Expt.
Γ_{1v}	-12.73	-12.42	-12.53	-12.6
Γ'_{25v}	0.0	0.0	-0.32,0.0	-0.3,0.0
Γ'_{2c}	-0.12	0.48	0.38	0.89 ^a
Γ_{15c}	2.58	3.16	2.89,3.13	3.25
X_{1v}	-8.89	-8.82	-8.93	-9
X_{4v}	-3.00	-3.02	-3.13	-3.2
X_{1c}	0.72	1.27	1.16	1.3
L'_{2v}	-10.67	-10.52	-10.63	-10.6
L_{1v}	-7.58	-7.55	-7.75	-7.7
L'_{3v}	-1.36	-1.37	-1.57,-1.37	-1.5
L_{1c}	0.09	0.65	0.54	0.7
L_{3c}	3.75	4.35	4.18,4.30	4.2/4.3

^a Ref. 65

between Γ and X to a very satisfying degree, with a maximum deviation of only 0.13 eV. We also produce the direct transition at L in AlAs extremely well. However, the direct gap at Γ as well as the location of the L_{1c} state relative to the VBM all show relatively disappointing agreement with the available experimental data. Calculations by Zhu and Louie [34], using a similar methodology except with the use of model dielectric functions instead of those computed *ab initio*, get much of the same disagreement with the available values as shown here. The experimental value of the direct gap at Γ in AlP was called into question in Ref. 34, where it was pointed out that the feature attributed to the direct gap in the photoluminescence spectrum of AlP is not very distinct, likely as a result of the large defect concentration in the sample. Additionally, they suggested that the feature claimed to be the direct gap at 3.63 eV might actually be an indirect transition allowed by phonon or defect scattering from Γ_{15v} to X_{3c} , which they found to be at 3.56 eV, precisely the same value found in this study. The results for AlAs also show discrepancy larger than normal for the direct transition at Γ . The work of Shirley *et al.* suggests that half of this underestimation may be the result of core-valence exchange and correlation effects relating to the As core states. Using their CPP approach, they find that the direct gap at Γ in AlAs is increased from 2.74 eV to 2.93 eV, bringing the result closer to agreement with the experimental value of 3.11 eV [63]. Other transitions in AlAs were more weakly affected by the inclusion of these effects. The indirect transition $\Gamma_{15v} \rightarrow L_{1c}$ is strongly overestimated, which is similar to that reported in a number of other calculations [34, 63, 66]. Godby and co-workers pointed out that this value is simply an extrapolation to $x=1$ for an alloy of $\text{Ga}_{1-x}\text{Al}_x\text{As}$ [67]. However, the experimental range only goes to $x=0.6$ and thus the accuracy of such an extrapolation is questionable. Therefore this value, as in previous works, has been presumed to not accurately represent the energy of this transition.

The calculated direct gap at Γ in AlSb is, like in AlAs, underestimated somewhat relative to experiment. We find a larger underestimation of this direct gap than found in Ref. 34, a difference of 0.36 eV between the calculations. The difference likely has multiple causes. One of the dominant reasons may be the update of the energies in the Green's function, which is done in Ref. 34 and is said to increase the quasiparticle band gaps by approximately 0.1 eV for the majority of the materials studied. We strictly perform G_0W_0 in this work without update to any eigenvalues, with the exception of correcting the occupation at Γ in situations where DFT incorrectly predicts a metal as described earlier. The use of a model dielectric function in Ref. 34 might also be partly responsible for the differences. For example, in Ref. 34 it was found that the use of a model dielectric function increased the Γ_{1c} gap in AlAs by 0.11 eV in comparison to their calculations with a dielectric function computed from first principles. A similar overestimation due to the model dielectric function may be present in AlSb. Finally, various levels of convergence could be present as well, e.g. reducing the q -grid in this work from 8^3 to the 4^3 used in Ref. 34 results in a direct band gap which is larger by ~ 60 meV. We note that the underestimation found in the

present work in AlSb with respect to experiment might be related to the same effects just mentioned in regards to the Γ_{1c} gap in AlAs, although an investigation of this possibility is outside the scope of the current work. Finally, in Table 2.5 the position of L_{1c} is strongly underestimated with respect to the available experimental value. This underestimation is seen in Ref. 34 as well, where the experimental evidence behind this value was questioned.

The quasiparticle band structures for AlP, AlAs, and AlSb are shown in Figures 2.5, 2.6, and 2.7. The spin-orbit splitting can be clearly seen in the plots for AlAs and AlSb. Our spin-orbit splitting values are generally always in good agreement with experiment. For AlP, we are unaware of experimental Δ_0 splitting parameters, although our calculated value of 63 meV is in fantastic agreement with values of 62 and 68 meV found elsewhere in the literature [68, 69]. In Table 2.5 it is clear that the splitting for the Γ_{15c} state (Δ_{15}) in AlSb is dramatically underestimated compared to the experimental value. We obtain a value of 59 meV in comparison to the 300 meV found experimentally [70]. However, we note that our result is in very good agreement with other theoretical predictions, such as 61 meV from a calculation performed with the Augmented Spherical Wave (ASW) method [69] and 90 meV with a relativistic orthogonalized plane wave (ROPW) model [71]. Thus we agree with the view brought forth in Ref. 69 calling for the experimental value for this splitting to be reconsidered.

We have also included in our study the In-based III-V semiconductors InP, InAs, and InSb. Like the $3d$ states of Ga, the In $4d$ states are shallow in energy and are known to be important for structural and electronic properties [73, 74]. In our calculations we therefore include the entire $n=4$ shell as valence in the calculations. As mentioned previously, in this work we use a generalized plasmon pole (GPP) model for the frequency dependence of the screening. In this model the imaginary part of the dielectric matrix is related to the charge density via a generalized f -sum rule [42]

$$\int_0^\infty \text{Im} [\epsilon_{\mathbf{G}\mathbf{G}'}^{-1}(\mathbf{q}, \omega)] \omega d\omega = -\frac{\pi}{2} \omega_p^2 \frac{(\mathbf{q} + \mathbf{G}) \cdot (\mathbf{q} + \mathbf{G}')}{|\mathbf{q} + \mathbf{G}|^2} \frac{\rho(\mathbf{G} - \mathbf{G}')}{\rho(\mathbf{0})}. \quad (2.15)$$

However, in the present calculation using the full charge density means including the semicore shell into the above sum rule. As Tiago *et al.* have noted, this implies that the semicore electrons are able to screen as effectively as the standard valence electrons [48]. To avoid this unphysical assumption, they only included the charge density due to the valence electrons in the above sum rule in their study of Si, Ge, and GaAs. It was found that this choice made less than a 0.1 eV difference in the resulting quasiparticle eigenvalues than if the full semicore + valence charge density was taken. We follow this approach and reconstruct the charge density from only the typical valence electrons as input to the GPP model. The results from our calculations are shown in Table 2.6. The computed values are almost uniformly in very good agreement with the experimental values. The quasiparticle corrections fix the band overlap present for InAs and InSb, which LDA incorrectly predicts to be metallic. In

Table 2.5: Quasiparticle energies for group III-V semiconductors AlX (X=P,As,Sb). Spin-orbit split values are shown on the same line separated by commas. In instances where differing experimental data was reported we have included them separated by a slash (/). All energies are taken from Ref. 61 unless otherwise noted. All energies are in eV.

	AlP				AlAs				AlSb			
	LDA	GW	GW+SO	Expt.	LDA	GW	GW+SO	Expt.	LDA	GW	GW+SO	Expt.
Γ_{1v}	-11.51	-11.19	-11.21		-11.91	-11.63	-11.73		-10.77	-10.39	-10.62	
Γ_{15v}	0.0	0.0	-0.06,0.0	0.0	0.0	0.0	-0.31,0.0	-0.31,0.0	0.0	0.0	-0.68,0.0	-0.67,0.0
Γ_{1c}	3.11	4.16	4.14	3.63 ^a	1.90	2.83	2.73	3.11	1.39	2.09	1.87	2.38
Γ_{15c}	4.49	5.70	5.66,5.69		4.28	5.18	5.05,5.10		3.17	3.79	3.53,3.58	3.7,4.0 ^{b,c}
X_{1v}	-9.16	-8.97	-8.99		-9.97	-9.78	-9.88		-8.82	-8.57	-8.79	
X_{3v}	-5.35	-5.36	-5.38		-5.44	-5.54	-5.64		-5.49	-5.58	-5.81	
X_{5v}	-2.12	-2.14	-2.17,-2.15		-2.17	-2.23	-2.40,-2.27	-2.45,-2.30	-2.17	-2.22	-2.60,-2.30	
X_{1c}	1.47	2.65	2.63	2.5	1.38	2.29	2.19	2.1/2.23 ^b	1.18	1.84	1.62	1.6
X_{3c}	2.35	3.58	3.56		2.22	3.17	3.06		1.41	2.07	1.85	
L_{1v}	-9.82	-9.60	-9.62		-11.51	-10.30	-10.40		-9.40	-9.12	-9.35	
L_{1v}	-5.58	-5.56	-5.58		-5.58	-5.67	-5.77		-5.42	-5.47	-5.70	
L_{3v}	-0.77	-0.78	-0.82,-0.79		-0.82	-0.86	$E_{L_{1c}} - 3.9$ $E_{L_{1c}} - 3.72$	$E_{L_{1c}} - 3.9^b$	-0.89	-0.93	$E_{L_{1c}} - 3.03$ $E_{L_{1c}} - 2.64$	$E_{L_{1c}} - 3.31^b$ $E_{L_{1c}} - 2.89^b$
L_{1c}	2.67	3.81	3.79		2.04	2.95	2.85	2.36 ^b /2.49/2.54	1.26	1.91	1.68	2.33
L_{3c}	4.76	6.03	6.01,6.02		4.68	5.65	5.54,5.56		3.78	4.46	4.22,4.26	

^a Ref. 72.

^b Taken from experimental results compiled in Ref. 34.

^c See discussion in text.

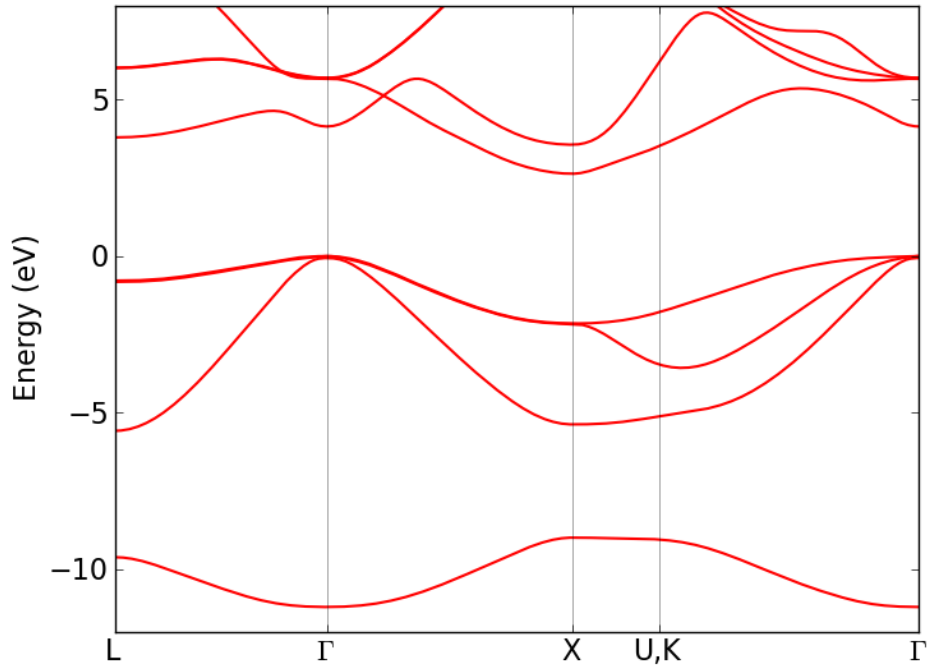


Figure 2.5: Quasiparticle bandstructure of AlP, including the effects of spin-orbit. The valence band maximum is set at zero energy.

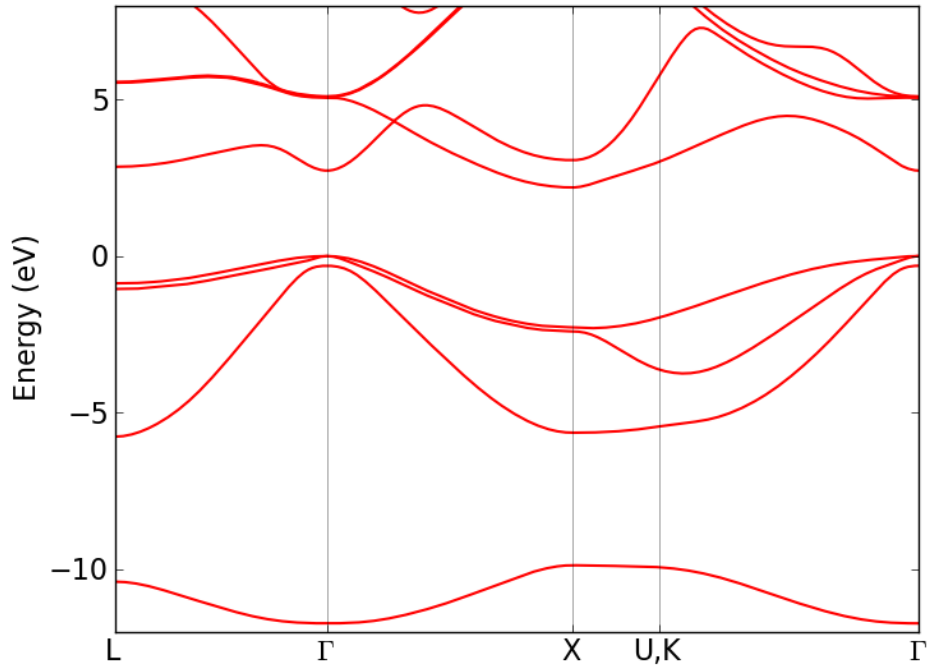


Figure 2.6: Quasiparticle bandstructure of AlAs, including the effects of spin-orbit. The valence band maximum is set at zero energy.

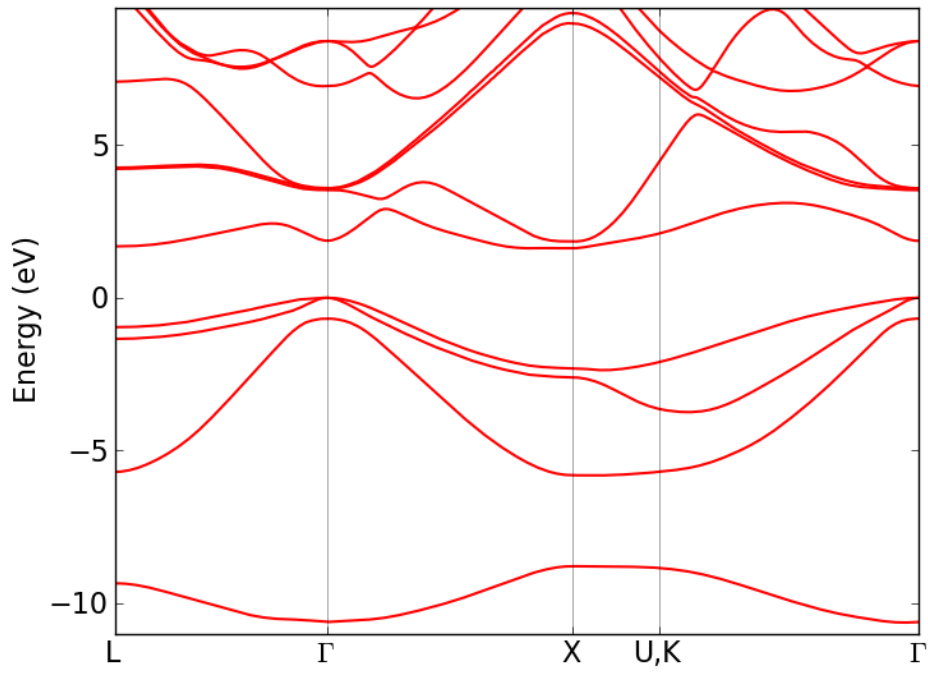


Figure 2.7: Quasiparticle bandstructure of AlSb, including the effects of spin-orbit. The valence band maximum is set at zero energy.

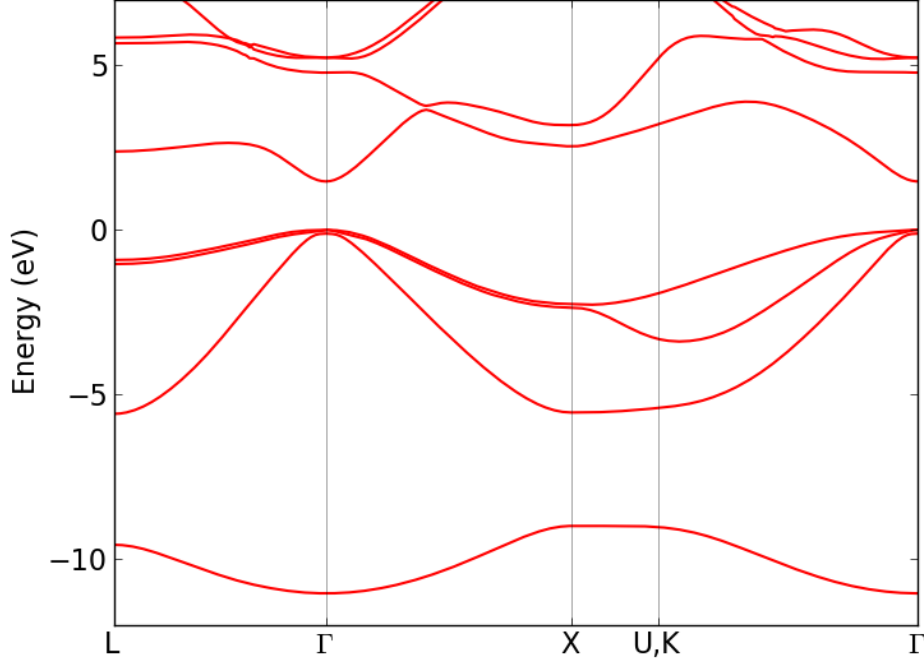


Figure 2.8: Quasiparticle bandstructure of InP, including the effects of spin-orbit. The valence band maximum is set at zero energy.

In Appendix A we have shown the explicit convergence tests for the band summations and size of the dielectric matrix, which leave only ~ 15 meV remaining uncertainty in the calculation due to these parameters. However, due to the computational cost of the calculation we have performed the q -grid summation on a 4^3 grid. In our discussion for Ge and AlSb, we have noted a lack of convergence in gaps between high symmetry points of about 80 and 60 meV due to this reduced q sampling, and uncertainties of similar size should be expected to be present in the results of Table 2.6. The band structures of InP, InAs, and InSb along high-symmetry lines are shown in Figures 2.8, 2.9, and 2.10.

Finally, we discuss our results for the Ga-containing compounds GaP, GaAs, and GaSb. It has been shown previously that in order to get results of high accuracy, semicore states must be taken into account in GaAs due to the shallow $3d$ electrons in Ga [48]. Accordingly, we have included the full $n=3$ shell of Ga in the calculation of the quasiparticle energies of GaP, GaAs, and GaSb. The results of the calculations are shown in Table 2.7. Again, the calculated quasiparticle energies show very good agreement with the experimental values. While still small, the largest deviations in gaps between band edge states occur for the direct transition at Γ in GaAs and GaSb.

Table 2.6: Quasiparticle energies for group III-V semiconductors InX (X=P,As,Sb). Spin-orbit split values are shown on the same line separated by commas. In instances where differing experimental data was reported we have included them separated by a slash (/). Experimental data is from Ref. 61 unless otherwise specified. All energies are in eV.

	InP				InAs				InSb			
	LDA	GW	GW+SO	Expt.	LDA	GW	GW+SO	Expt.	LDA	GW	GW+SO	Expt.
Γ_{1v}	-11.53	-11.01	-11.04	-11.4/-11.0	-11.89	-11.40	-11.53	-12.3	-10.79	-10.38	-10.54	-11.7/-10.8
Γ_{15v}	0.0	0.0	-0.11,0.0	-0.11,0.0	0.0	0.0	-0.38,0.0	-0.37,0.0	0.0	0.0	-0.78,0.0	-0.80 ^a /-0.85,0.0
Γ_{1c}	0.44	1.51	1.47	1.42	-0.46	0.55	0.42	0.36	-0.45	0.54	0.28	0.24
Γ_{15c}	3.96	5.12	4.78,5.23	4.78,5.14	3.84	4.70	4.25,4.73	4.5	2.96	3.61	3.06,3.50	3.14,3.53
X_{1v}	-9.12	-8.96	-9.00	-9.6/-8.9	-9.91	-9.71	-9.84	-9.8	-8.83	-8.58	-8.84	-9.5
X_{3v}	-5.85	-5.51	-5.55	-6.0/-5.9	-5.90	-5.65	-5.78	-6.3/-6.0	-5.97	-5.69	-5.96	-6.4/-6.1
X_{5v}	-2.34	-2.27	-2.36,-2.25	-2.7/-2.2/-2.0	-2.32	-2.29	-2.42,-2.42	-2.7/-2.4	-2.26	-2.25	-2.60,-2.43	-2.4
X_{1c}	1.62	2.58	2.54	2.8	1.45	2.17	2.04	1.9	1.27	1.81	1.55	1.8
X_{3c}	2.12	3.22	3.18		1.96	2.80	2.67		1.27	1.90	1.65	
L_{1v}	-9.79	-9.53	-9.57	-10.0	-10.44	-10.18	-10.31	-10.6	-9.43	-9.12	-9.38	-10.5/-10.0
L_{1v}	-5.81	-5.55	-5.59		-5.81	-5.63	-5.75		-5.67	-5.47	-5.74	
L_{3v}	-0.98	-0.94	-1.04,-0.92	-1.0	-0.99	-0.97	-1.23,-0.96	-0.9	-1.00	-0.99	-1.50,-1.01	-1.4/-0.9
L_{1c}	1.29	2.42	2.38	2.38	0.79	1.70	1.57	1.55 ^a	0.40	1.16	0.89	
L_{3c}	4.71	5.79	5.67,5.84		4.63	5.43	5.20,5.40		3.82	4.41	4.05,4.25	4.32,4.47

^a Taken from experimental results compiled in Ref. 34.

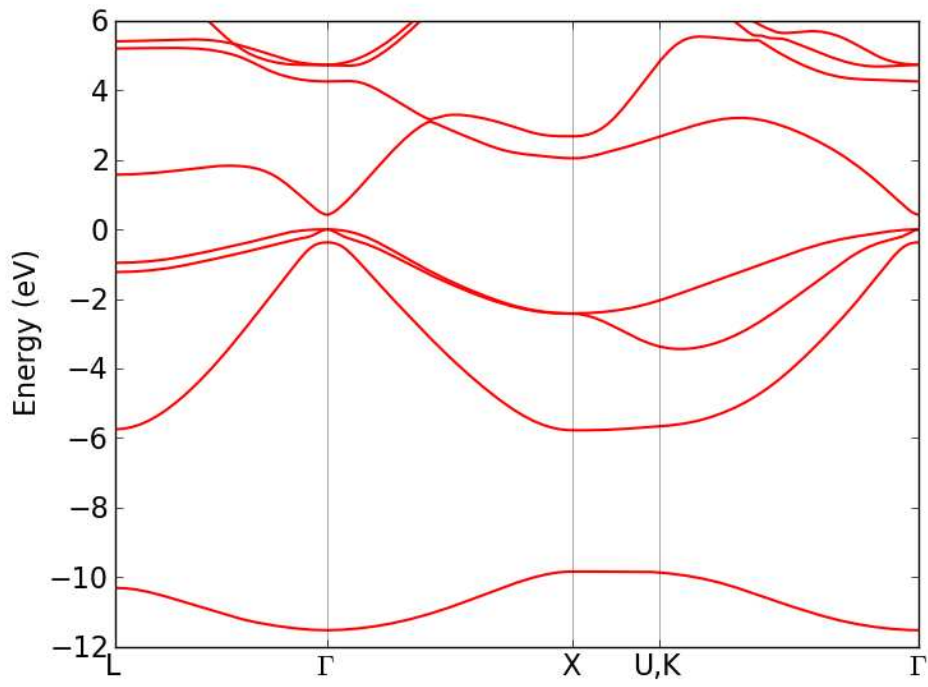


Figure 2.9: Quasiparticle bandstructure of InAs, including the effects of spin-orbit. The valence band maximum is set at zero energy.

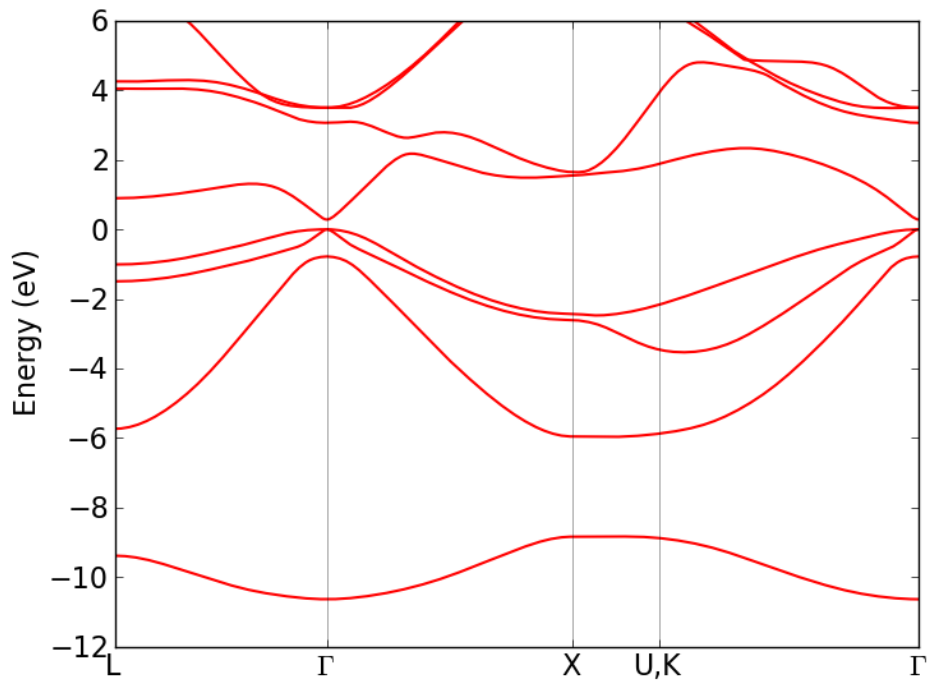


Figure 2.10: Quasiparticle bandstructure of InSb, including the effects of spin-orbit. The valence band maximum is set at zero energy.

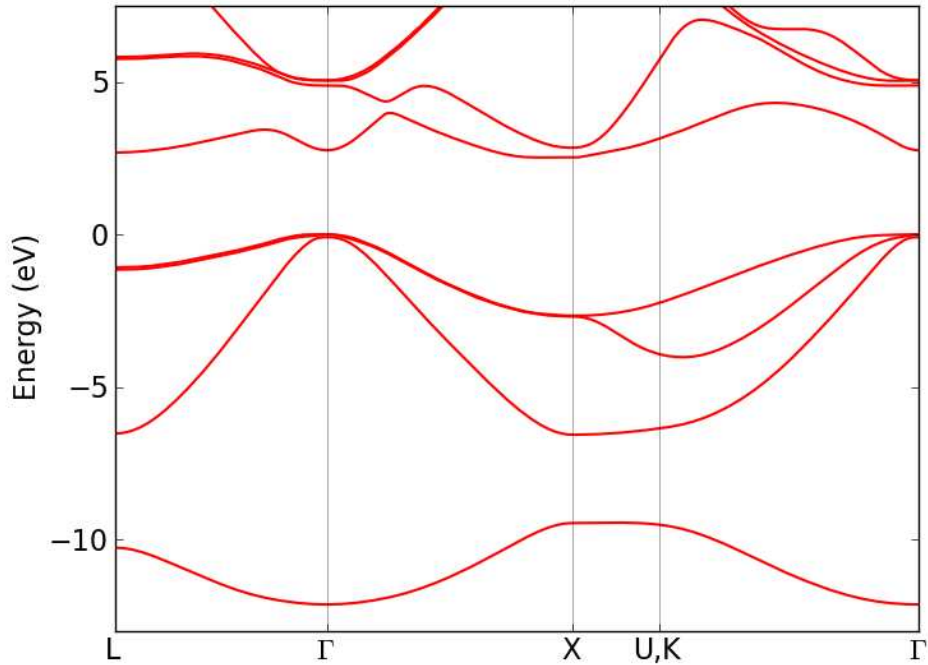


Figure 2.11: Quasiparticle bandstructure of GaP, including the effects of spin-orbit. The valence band maximum is set at zero energy.

Shirley discussed the difficulty of calculating this gap when post-transition elements are involved due to LDA's strong overbinding of s and p states, with the s being more strongly bound. Since the transition at Γ has a strong p to s character, the calculated band gaps are expected to have a bias towards being smaller [63]. Additionally, core-polarization effects in As and Sb might slightly increase the direct transition at Γ , as we discussed in relation to AlAs. Our calculated band structures for GaP, GaAs, and GaSb are shown in Figures 2.11, 2.12, and 2.13.

Table 2.7: Quasiparticle energies for group III-V semiconductors GaX (X=P,As,Sb). Spin-orbit split values are shown on the same line separated by commas. In instances where differing experimental data was reported we have included them separated by a slash (/). Experimental data is from Ref. 61 unless otherwise specified. All energies are in eV.

	GaP				GaAs				GaSb			
	LDA	GW	GW+SO	Expt.	LDA	GW	GW+SO	Expt.	LDA	GW	GW+SO	Expt.
Γ_{1v}	-12.52	-12.11	-12.14	-12.5	-12.79	-12.42	-12.53	-13.8/-13.1	-11.60	-11.23	-11.47	-11.6
Γ_{15v}	0.0	0.0	-0.09,0.0	-0.08,0.0	0.0	0.0	-0.35,0.0	-0.34,0.0	0.0	0.0	-0.73,0.0	-0.76/0.0
Γ_{1c}	1.59	2.79	2.75	2.88	0.32	1.42	1.31	1.52 ^a /1.63	-0.13	0.94	0.70	0.81
Γ_{15c}	3.81	5.02	4.88,5.05	4.7/4.87 ^a	3.71	4.63	4.38,4.58	4.61 ^a /4.71/4.8	2.82	3.47	3.09,3.31	3.2 ^a ,3.4 ^a
X_{1v}	-9.52	-9.44	-9.47	-9.9	-10.30	-10.18	-10.29	-10.7	-9.14	-8.97	-9.21	-9.6/-9.4
X_{3v}	-6.85	-6.54	-6.57	-6.8	-6.88	-6.67	-6.79	-7.1/-6.81 ^a /-6.7	-6.88	-6.62	-6.86	-6.9
X_{5v}	-2.70	-2.65	-2.69,-2.67	-3.0/-2.9	-2.68	-2.68	-2.84,-2.76	-2.91 ^a /-2.8/-2.5	-2.58	-2.58	-2.96,-2.70	-3.10,-2.86
X_{1c}	1.51	2.55	2.52	2.39 ^a	1.37	2.15	2.03	1.9 ^a /2.10/2.18	0.85	1.50	1.26	1.13
X_{3c}	1.69	2.87	2.84	2.75 ^a	1.57	2.47	2.35	2.47 ^a /2.58	1.17	1.73	1.48	
L_{1v}	-10.44	-10.24	-10.27	-10.7	-11.03	-10.84	-10.96	-12.0/-11.24	-9.96	-9.72	-9.97	-10.3/-10.1
L_{1v}	-6.71	-6.50	-6.53	-6.6	-6.70	-6.56	-6.68	-7.1/-6.7	-6.44	-6.27	-6.51	-6.9/-6.6
L_{3v}	-1.23	-1.09	-1.16,-1.09	-0.9	-1.14	-1.13	-1.36,-1.14	-1.4/-1.3	-1.15	-1.14	-1.60,-1.18	-1.55,-1.1
L_{1c}	1.50	2.71	2.68	2.64 ^a	0.87	1.87	1.75	1.74 ^a /1.84	0.31	1.09	0.85	0.92 ^a /1.09
L_{3c}	4.64	5.81	5.75,5.81		4.61	5.49	5.33,5.42		3.78	4.39	4.09,4.21	4.36,4.49

^a Taken from experimental results compiled in Ref. 34.

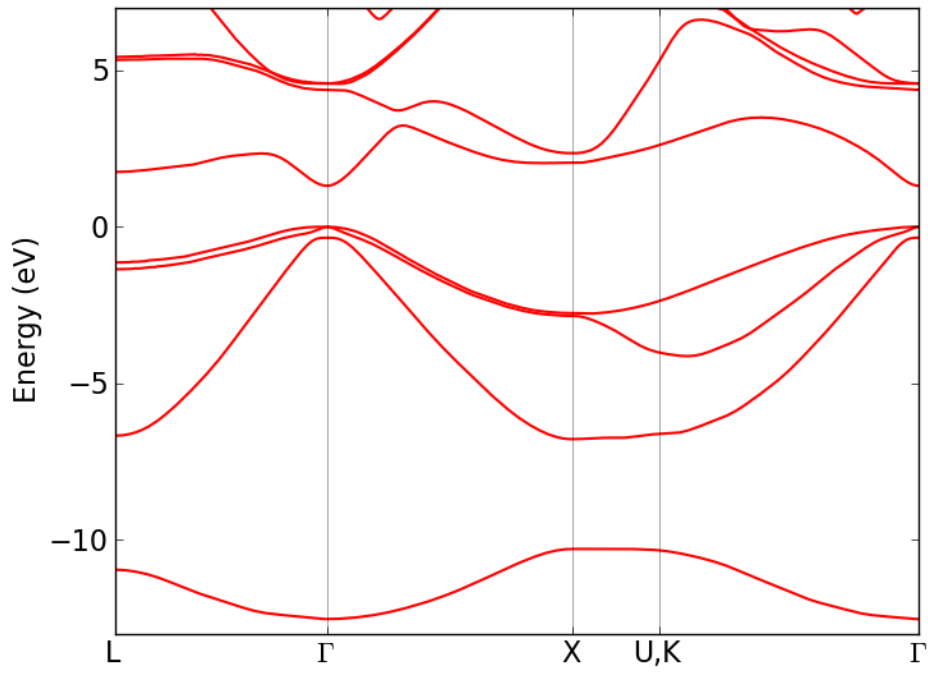


Figure 2.12: Quasiparticle bandstructure of GaAs, including the effects of spin-orbit. The valence band maximum is set at zero energy.

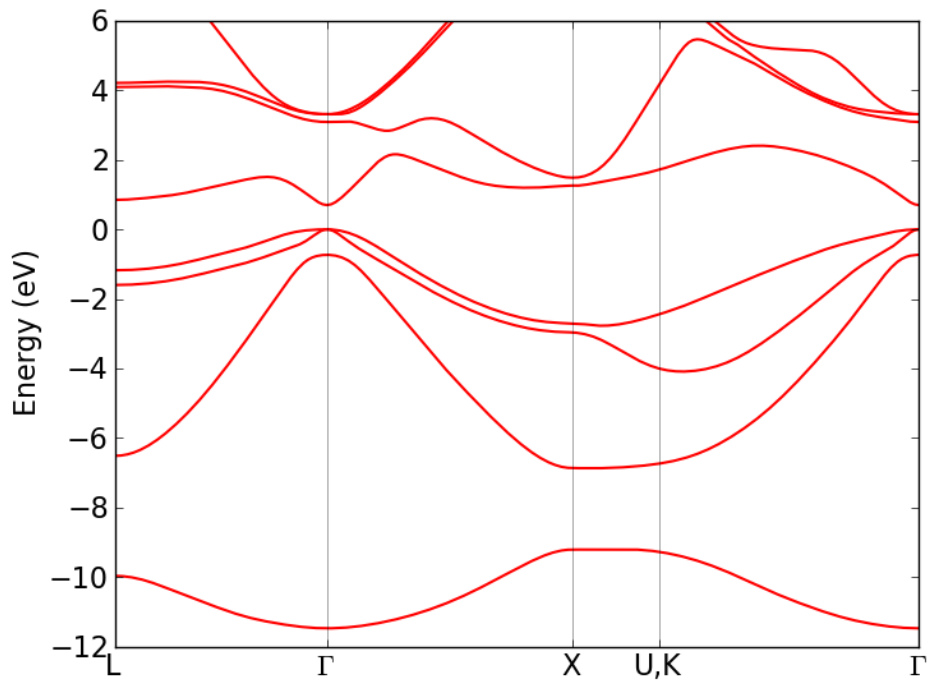


Figure 2.13: Quasiparticle bandstructure of GaSb, including the effects of spin-orbit. The valence band maximum is set at zero energy.

2.4 Conclusion

In summary, we have presented fully first principles quasiparticle band structures for the group IV and group III-V materials Si, Ge, AlP, AlAs, AlSb, InP, InAs, InSb, GaP, GaAs, GaSb. Our calculations are done using the planewave pseudopotential method and the “one-shot” GW, or G_0W_0 , method. We have paid particular attention to the convergence aspects of the calculation which we have demonstrated in Appendix A. Semicore electrons were included in the calculation of the In- and Ga-containing compounds because of their shallow d electrons which are typically left in the core. The spin-orbit interaction is calculated from first-principles as a perturbation and included in the results. Finally, using a Wannier interpolation of the GW eigenvalues and the spin-orbit matrix, we have quasiparticle band dispersions throughout the Brillouin zone. Our obtained results are generally in very good agreement with experiment and demonstrate the ability of the method to reliably predict quasiparticle energies in these materials.

Chapter 3

Ab initio survey of the electronic structure of tetrahedrally bonded phases of silicon

3.1 Introduction

Crystalline silicon in the cubic diamond (cd) phase has been the subject of a large number of theoretical and experimental studies because of its importance in electronic and photovoltaic devices. Less studied have been the other members of silicon's rich phase diagram. This includes a large number of metastable silicon polytypes which are found upon decompression from the metallic β -Sn phase. Some of these phases have only been known for a little more than a decade. These phases, the BC8, hexagonal diamond, R8, and yet unobserved ST12 phase, are in many ways similar to their cubic counterpart. They are all characterized by four-fold coordination, distorted tetrahedral bonding, and are of semiconducting (or semi-metallic) nature. These similarities lead to the hope that the electronic features of the cubic phase found useful in applications might also be found in these phases, perhaps resulting in one of these phases being better suited for a particular application than its cubic counterpart. In particular we will be looking at the features in the electronic structure which would be relevant for application in high-mobility semiconductor devices and comparing these to those found in the cubic diamond phase.

The majority of previous theoretical work on the high-pressure phases of silicon have concentrated on their structural properties and characteristics of the phase transitions between them with results generally in good agreement with experiment [27, 75–80]. These calculations were done using density-functional theory (DFT) within the local-density approximation (LDA), which is known to give band gaps in disagreement with experiment [42]. This has been shown to arise from the discontinuity in the exchange-correlation potential upon the addition of an electron [30, 31].

As discussed in Chapter 2, strictly speaking the Kohn-Sham eigenvalues have no justification as the quasiparticle energies in a solid [32]. The true quasiparticle energies can be calculated from a Green's function approach, allowing the full nonlocal, energy-dependent electron self-energy to be calculated. A successful approximation within this framework is the so-called GW approximation (GWA). In this chapter we present our findings from calculations of the quasiparticle spectrum within the GWA for materials whose electronic nature is in question with regards to LDA.

To evaluate the suitability of these materials for high-mobility applications we examine the carrier effective masses which allow for an estimate of the relative magnitude of the low-field mobility at room temperature. The mobility μ of the carriers in a semiconductor depend strongly on their effective masses, as can be seen by simple scaling arguments for the scattering amplitudes which give [81]

$$\mu \propto T^{-3/2} m_{DOS}^{*-3/2} m_t^{*-1} \quad (3.1)$$

where m_{DOS}^* is the density-of-states effective mass and m_t^* is the transport effective mass. Missing in the above simple expression is a detailed account of the scattering matrix elements. However, in this work the phases under study have similar bonding characteristics and atomic densities and so we assume that the electron-phonon scattering matrix elements are comparable. In order to get an estimate of the mobility of the carriers relative to those in cubic silicon we compare a thermally averaged density-of-states effective mass, given by

$$\left(\frac{m^*}{m_0}\right)_{DOS}^{3/2} = \frac{\int f(E)N(E)dE}{\int f(E)N_0(E)dE} \quad (3.2)$$

where $f(E)$ is the Boltzmann distribution given by $f(E) = \exp[-E/(k_b T)]$ and $N(E)$ and $N_0(E)$ being the calculated density-of-states and the free electron density-of-states, respectively. The effective masses are determined to a very high precision by using a Wannier interpolation scheme which approximates 100x100x100 grids in the Brillouin zone.

The chapter is organized as follows. In Sec. 3.2 we briefly review the structures of the phases discussed in this paper and in Sec. 3.3. we detail the methods used in the calculations. The results of the electronic structure calculations are presented in Sec. 3.4. Finally, in Sec. 3.5 we conclude with a short summary of the results and some final remarks.

3.2 Structures

It is well known that cubic silicon (Si I) transforms upon compression at ~ 11.7 GPa [82] to the metallic β -Sn phase (Si II) and that upon release of pressure does not return to the cubic phase but instead goes into a series of high-density metastable

phases with distorted tetrahedral bonding. Calculation of the energetics of these phases confirm their metastable character and also correctly predict the ordering of the phases upon decompression [76, 83]. Since the structural characteristics of these phases have been studied previously [75, 76, 78] we will mention below only the features relevant to this study.

The first phase formed under decompression is the R8 phase (Si XII). This structure has a rhombohedral unit cell containing 8 atoms with the space group $R\bar{3}$ [84]. The R8 structure has the interesting property of containing five-membered rings, which in addition to affecting the electronic properties [85] have also been used to explain why the BC8 analog has not been seen in III-V semiconductors. If these compound semiconductors were to follow the same transition path as silicon the existence of odd-membered rings would require the existence of energetically unfavorable same-species bonds [86].

Upon further decompression the R8 phase transforms into the BC8 phase (Si III) with body-centered cubic symmetry and 8 atoms in the primitive cell. Its space group classification is $T_h^7(Ia\bar{3})$. This structure is closely related to that of R8 and in fact can be formed by changing only one pair of bonds [76]. The internal structure is specified by a single parameter x , which experimentally has the value 0.1025 [84]. The relations connecting the internal parameter to bond lengths and angles can be found in the work of Crain *et al.* [75]

Upon moderate heating of the BC8 phase the hexagonal diamond structure, also referred to as lonsdaleite, is formed. The 4 atom primitive cell of lonsdaleite is classified with space group $P6_3/mmc$. Further details of this structure can be found in the literature [87, 88].

Another possible structure examined in this work is the tetragonal ST12 structure, which has a 12 atom unit cell given by space group $P4_32_12(D_8^4)$. The ST12 phase of silicon has been the study of a number of theoretical calculations although has yet to be experimentally observed in silicon. It is, however, seen in germanium upon slow decompression from the germanium β -Sn phase [89]. Work done by Clark *et al.* [77] using an empirical potential model suggest that the reason for this is that in silicon bond-bending is more costly than in germanium because of the additional electrons in the Ge core. The ST12 phase has bond angles further from the ideal diamond value than BC8, and it is this fact which results in ST12 being favored in Ge and BC8 in Si. From their model they also suggest that it might be possible to form ST12 silicon by decompressing from β -Sn at high temperature.

3.3 Method

As a starting point, we have performed *ab initio* pseudopotential calculations within density-functional theory using the parameterization by Perdew and Zunger [24] of the Ceperley-Alder [23] exchange-correlation functional in the local-density

approximation (LDA). The interaction of the valence electrons with the ion cores is described by a norm-conserving pseudopotential generated by the scheme proposed by von Barth and Car. The wave functions are expanded in plane waves [11] up to a cutoff of 40 Ry. Structural relaxation of the lattice parameters and internal degrees of freedom were performed unless otherwise specified using a quasi-Newton method [90].

As noted in the introduction, density-functional theory is unable to correctly predict the quasiparticle energies, and hence the band gap, in solids. In this chapter we thus employ the GW approximation to the self-energy, which results from setting the vertex function to unity in Hedin’s equations [40]. The self-energy then becomes a simple product of the Green’s function G , and the screened interaction W [42]

$$\Sigma(1, 2) = iG(1, 2)W(1^+, 2) \quad (3.3)$$

from which the name follows. The Green’s function and the screened interaction must in principle be calculated self-consistently, although this is rarely done in practice. We follow the technique of Hybertsen and Louie [42] in which the Green’s function is approximated by the wavefunctions and energies that result from a DFT calculation and the frequency-dependent dielectric matrix needed to compute the screened interaction is obtained by extending the static dielectric matrix to finite frequencies by using a generalized plasmon-pole model. The GW calculations are done using the PARATEC plane wave code, which implements this approach to the GWA [91].

Density-of-states and effective mass calculations were performed on a 100x100x100 fine mesh of Wannier-interpolated eigenvalues [92] calculated from the WANNIER90 code [58] used in conjunction with the Bloch states obtained from the Quantum ESPRESSO package [46].

3.4 Results

3.4.1 R8

Experimentally the R8 phase is formed upon decompression from the high-pressure metallic β -Sn phase at ~ 10 GPa and measurements taken at 8.2 GPa find β -Sn to be only a small fraction of the experimental sample [86]. As the pressure is reduced R8 remains the dominant phase until approximately 2 GPa, when BC8 is formed. The R8 phase persists, albeit as a small minority, at ambient pressure. Upon increase of pressure the R8 phase is found to begin and complete its transition to the β -Sn phase at 10 and 12 GPa, respectively [84]. We have performed *ab initio* calculations over the range of pressures that R8 has been observed experimentally, from 0-12 GPa.

A central question left unresolved by previous theoretical studies of R8 is whether R8 should be semiconducting or metallic, the uncertainty being a result of the “band gap” problem of LDA. Our calculations showing the LDA overlap as a function of

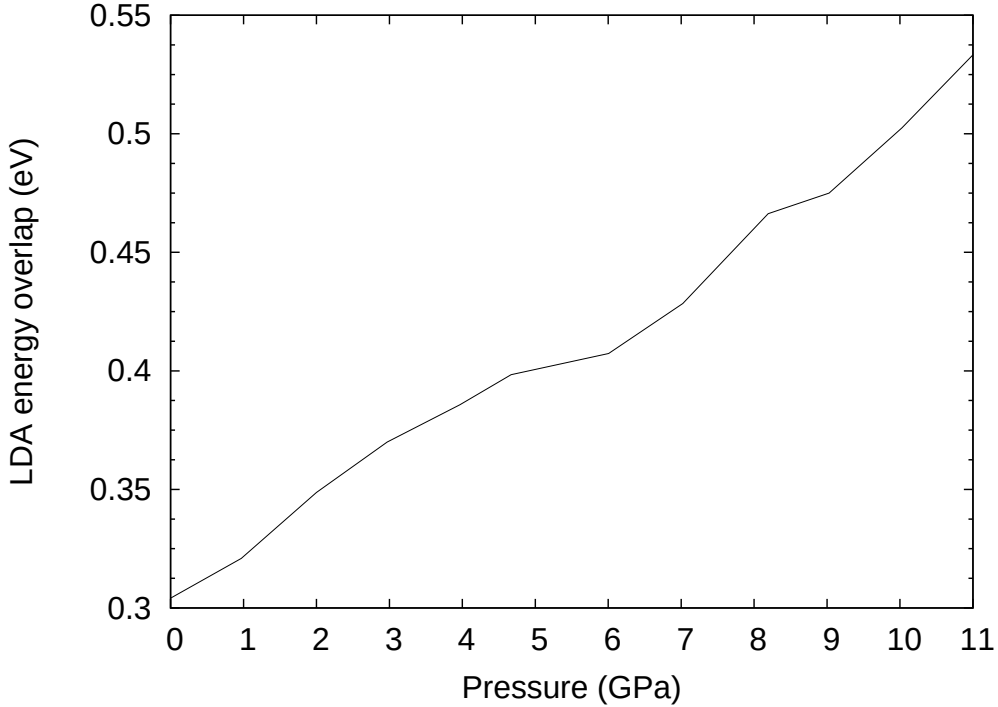


Figure 3.1: The LDA energy overlap in R8 as a function of pressure. The overlap occurs between the T and the L points of the Brillouin zone.

external pressure are shown in Fig. 3.1. At ambient pressure we find an indirect overlap of 0.3 eV between the T and L points in agreement with Pfrommer *et al.* [76] This is somewhat lower than the 0.5 eV found by Piltz *et al.* [86] This value compares more favorably to our value of 0.47 eV at 8.2 GPa, the pressure of the x-ray diffraction profile. The LDA band structure for R8 at ambient pressure is shown in Fig. 3.2.

An overlap of 0.3 eV is small enough that it could conceivably be lifted by quasiparticle corrections to the LDA. The GW approximation for the electron self-energy generally yields gaps much closer to experiment than what one would get if using the Kohn-Sham DFT eigenvalues. Cubic silicon is a well-known example of this. Here LDA predicts an indirect gap of 0.52 eV while the gap computed using GWA is 1.21 eV, which compares much more favorably to the experimental value of 1.17 eV [42]. With this motivation we have performed calculations of the electron self-energy within the GWA on the R8 structure at ambient pressure. The usual application of the GWA is to ignore the fact that self-consistency is in principle required. This “one-shot GWA”, or G^0W^0A , was used in calculating the quasiparticle corrections and we find that it results in an indirect gap of 0.1 eV, resulting in R8 being semiconducting. However, since the screening is qualitatively different in a semiconductor than it is

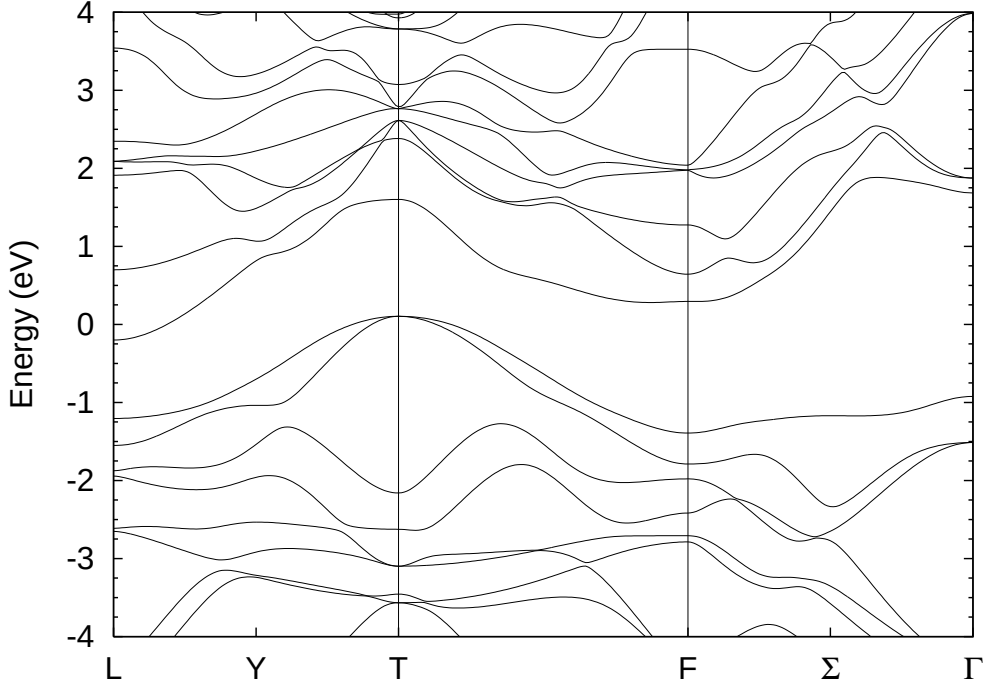


Figure 3.2: Band structure of silicon in the R8 phase.

in a metal, a one-shot GWA approach may be insufficient and it may be better to go to self-consistency in G and W . We approximate the self-consistency requirement by using updated quasiparticle energies from one iteration to construct G and W for the next iteration. The energies are updated by using a “scissor-shift” by which the quasiparticle energy shifts are approximated by a linear relation of the LDA energy [93]. By following this procedure until self-consistency is reached we find the indirect gap increasing to the value of 0.24 eV. If this value of the quasiparticle energy correction changes little with pressure [94], R8 should exhibit a gap over a substantial range of pressures for which it occurs. Noting the increase in the LDA overlap with increasing pressure, we find that the quasiparticle gap should be eliminated at ~ 11 GPa. This pressure lies between the onset and completion of the R8 to β -Sn transition upon recompression of the R8 phase. Inadequacies in applying the zero-pressure quasiparticle corrections to finite pressure or in the GW approximation itself could result in a lowering of this semiconductor-to-metal transition pressure. It may also be possible to experimentally adjust the gap by applying a strain. Applying a uniaxial strain equivalent to a 1% decrease in the c/a ratio of the equivalent hexagonal lattice we find that the overlap within LDA is reduced by 0.07 eV, a reduction of $\sim 23\%$ of its unstrained value. This would increase the gap from the value found above when including self-energy effects. On the other hand, it may be possible to eliminate the

gap by straining in the opposite sense. This could result in a material with a pressure-induced metal-semiconductor transition, which may have applications in fabricating modulators, polarizers, and microwave switches [95].

In order to get an estimate of what the relative mobility would be in the R8 phase compared to that in the cubic phase we examine the effective masses. Silicon in the cubic phase is characterized by 6 equivalent conduction band minima, which at 300K give a density-of-states mass of $1.09m_0$ [96]. The valence band maximum at Γ is double degenerate between a “heavy-hole” band and a “light-hole” band in addition to a “split-off” band slightly lower in energy as a result of the spin-orbit interaction. The density-of-states effective mass for this valence band is $1.15m_0$ [96].

In R8 the density-of-states within LDA does not exhibit a gap due to the indirect overlap. We separate out the states associated with the conduction and valence bands before integrating and find a thermally averaged density-of-states effective mass for the conduction band of $0.54m_0$. This is lower than unstrained cubic silicon, and is comparable to biaxially strained silicon, which lowers two of the 6 equivalent conduction band minima and reduces the DOS effective mass by approximately a factor of $3^{2/3}$. We calculate a value of $0.61m_0$ for the valence band at a temperature of 300K. This value is obtained neglecting spin-orbit, which may decrease the effective mass by splitting the degeneracy at T.

In R8 the L point in the Brillouin zone is threefold degenerate. This degeneracy can be lifted by applying a strain of 1% in the $[0\bar{1}1]$ direction, taken with respect to the equivalent hexagonal lattice. This results in one of the previously equivalent minima being lowered by approximately 90 meV relative to the other two and a decrease in the conduction band DOS effective mass to the value of $0.27m_0$. For the holes, the valence band degeneracy at T is lifted by the strain. The consequence of this is that the valence band DOS effective mass is reduced to the value of $0.36m_0$.

3.4.2 BC8

Upon further decompression the R8 phase reversibly transforms to the BC8 phase at a pressure of 2 GPa, and remains metastable down to ambient pressure [84]. Previous theoretical work done on BC8 has found it to be a semimetal within LDA with a direct overlap of 0.8 eV at the H point in the bcc BZ [76]. Experimental measurements on BC8 agree with this classification, although the nature of the overlap was suggested to be indirect with a value of 0.3 eV [97]. Impurity effects can sometimes make it difficult to determine transport properties in experimental measurements accurately and previous theoretical work done with empirical pseudopotentials show BC8 to have a direct gap of 0.43 eV at the H point [98]. For these reasons the possibility that BC8 might still be a small band gap semiconductor has not been precluded. In addition to quasiparticle energy corrections that can be included, it may also be possible to eliminate the overlap by straining the structure.

As noted previously, BC8 is specified by a single internal parameter x . Calcula-

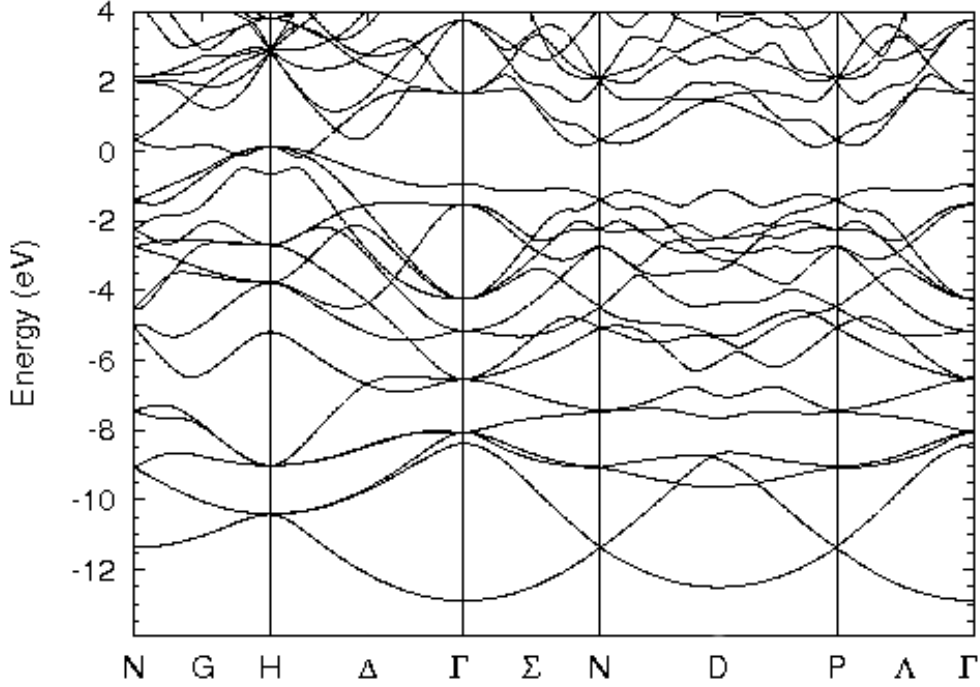


Figure 3.3: Band structure of silicon in the BC8 phase.

tions of the dependence of the LDA band overlap on the parameter x show the band overlap being reduced upon reduction of the parameter from its experimental value of 0.1025. At $x=0.093$, for example, we find a direct gap of 0.23 eV at the bcc H point, resulting in an electron density-of-states effective mass of $0.07m_0$. The small value of effective mass can be related to the existence of the small direct gap by standard $\vec{k} \cdot \vec{p}$ arguments. If a strain could be found which effectively changed this internal parameter, it would allow for the tuning of the band gap and effective masses to situations where a small gap, high-mobility semiconductor was desirable to cases where a larger gap might be preferred at the expense of the mobility.

Our electronic structure calculations are done by relaxing the structure from the experimental values. We find an x parameter of 0.1016, in good agreement with the experimental value of 0.1025. The calculated zero-pressure volume is $17.78 \text{ \AA}^3$ and is in good agreement with the experimental value of $18.26 \pm 0.04 \text{ \AA}^3$ per atom [99] and in excellent agreement with the calculated value of $17.724 \text{ \AA}^3$ per atom by Needs and Mujica [78]. As is typical with LDA calculations the lattice constant and volume are underestimated. The calculations of the electronic structure are in good agreement with the results of Pfrommer *et al.* [76] and we find an overlap of 0.77 eV at the bcc H point. This overlap can be seen in the band structure plotted in Fig. 3.3.

Motivated by the existence of a gap in previous empirical pseudopotential calcu-

lations [98], it is of interest to look at the effect of quasiparticle energy corrections to the direct gap within the GWA. However, this computation is complicated by the fact that within LDA BC8 has a direct overlap. This would require a detailed Brillouin zone summation to adequately represent the Fermi surface. To avoid this difficulty, we have used another approach to estimate the size of the quasiparticle corrections. Our approach makes use of the fact that the pressure coefficients for the band gap in cubic silicon have been found to be similar for both calculations done within the LDA and those done within the GWA [94]. Thus it may be reasonable to expect that the self-energy correction to the band gap doesn't depend too strongly with strain. Using this approximation, we perform a GW calculation on the BC8 structure with an x parameter that has been reduced such that the spectrum exhibits a small band gap. Choosing a value of $x=0.0947$ results in a vanishingly small band gap and taking the shift at the H point at this value of x as representative of the shift in the original structure we estimate a reduction of the direct overlap by 0.33 eV due to self-energy corrections. This estimate might also be argued as an upper bound for the quasiparticle energy correction. This follows because even if the GW correction does depend strongly on the value of the x parameter, changing the value to one in which we eliminate the gap should decrease the screening and thus enhance the screened interaction and consequently the electron self-energy correction Σ . Therefore we conclude that BC8 is indeed a semimetal judged by the quasiparticle energy spectrum calculated within the GWA.

In searching for a strain that would produce the effects of changing the internal parameter x we have considered a number of strains. This includes simple volume strain, uniaxial strain up to 4%, and the substitutional doping of germanium and carbon atoms. The band overlap was found to be resilient to these changes. Through uniaxial strain and doping we were able to reduce, although not eliminate, the BC8 overlap. In calculating the volume dependence of the x parameter, we find that for realistic volume strains the x parameter doesn't decrease to a value corresponding to one which would lift the band overlap.

3.4.3 Lonsdaleite

The hexagonal diamond, or lonsdaleite, structure is obtained from the heating of BC8 samples above 470K and is found to be stable up to ~ 800 K [97]. In addition, it has been formed by applying a shear stress to the cubic phase, by nonhydrostatic stress in indentation experiments, and with ultraviolet laser ablation with a cubic silicon target at lower pressures [100]. Experimental evidence suggests lonsdaleite being semiconducting with a band gap similar to that of the cubic silicon phase [97]. Previous theoretical work using empirical pseudopotentials have predicted an indirect band gap of 0.85 eV between the Γ and M points of the Brillouin zone [98]. Our calculations done within the LDA find an indirect gap of 0.33 eV at this point. Again, noting the large effect that self-energy corrections have on the indirect

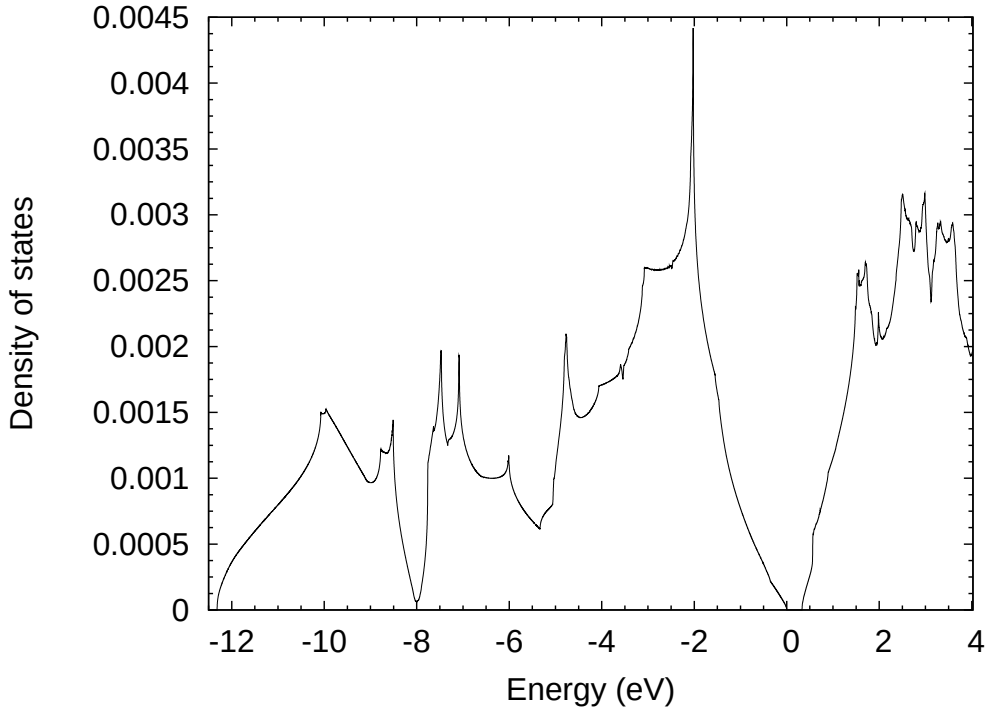


Figure 3.4: Density-of-states per unit volume, per spin, and per eV for silicon in the lonsdaleite phase

gap in the cubic phase, we have also calculated the electronic spectrum within the GWA. This results in an indirect gap of 0.95 eV, similar to both the results derived from empirical methods and experimental observation. Calculations of the electronic density-of-states utilizing the Wannier-interpolated eigenvalues are shown in Fig. 3.4. These results are in excellent agreement with those found earlier by Joannopoulos and Cohen [98], including the sharp structure near the Fermi level. Calculations of the thermally averaged DOS effective mass in the absence of the spin-orbit interaction result in the values of $0.94m_0$ for the electrons and $0.58m_0$ for the holes. Similar to the cubic case, we find that applying a biaxial strain results in a reduction of the conduction band effective mass. A biaxial strain applied in the plane perpendicular to the optical axis causes the 3 equivalent conduction band minima located at M in the hexagonal Brillouin zone to be split in energy, resulting in an electron effective mass of $0.63m_0$. The hole effective mass is not significantly affected by this strain.

3.4.4 ST12

As was noted in the introduction, the ST12 phase has been seen upon decompression from β -Sn in germanium, but has yet to be found in silicon. Despite this, the ST12 structure has been the subject of numerous theoretical studies [85, 101] due to

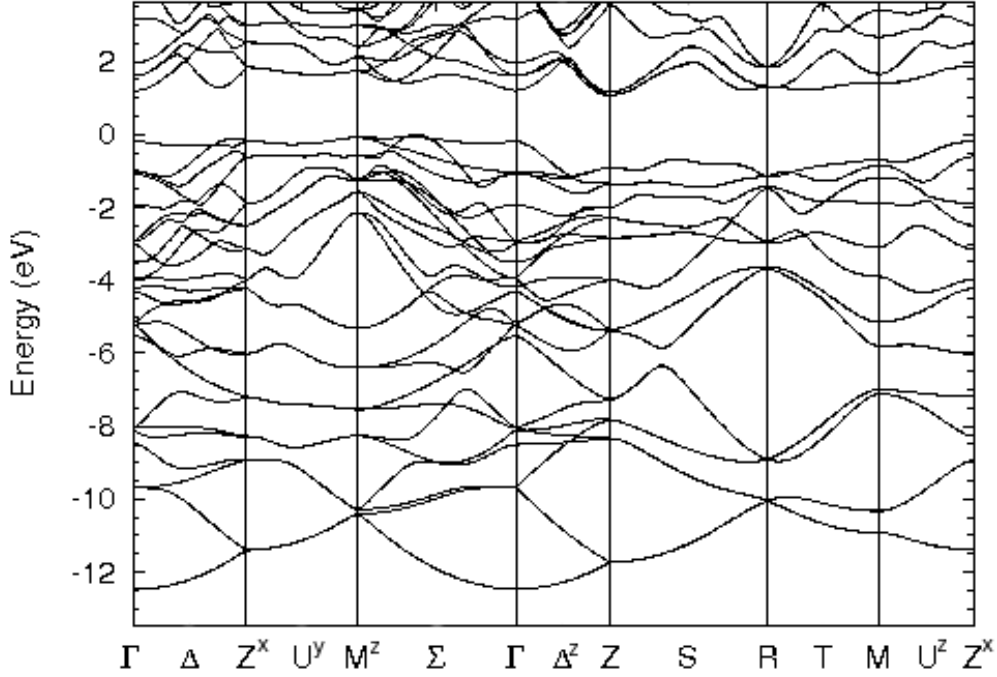


Figure 3.5: Band structure of ST12 silicon.

its usefulness in studying the effects of disorder of amorphous materials.

Our calculations for silicon in the ST12 structure are done with fully relaxing the cell and the internal degrees of freedom. The band structure of ST12 calculated in LDA is shown in Fig. 3.5 and is in qualitative with that found by Joannopoulos [98] using empirical pseudopotentials, although we find the indirect gap to be somewhat smaller at 1.1 eV. In ST12 we find that the bands are characterized by a large number of extrema near the gap edges. This results in the large density-of-states near the band gap as seen in FIG. 3.6. The observed number of states near the gap edge decreases the mobility of the carriers by providing states with which to scatter into. Computing the DOS effective mass for the electrons we find a value of $0.89m_0$, which while smaller than the value of $1.09m_0$ for unstrained cubic silicon, is still larger than that occurring in the case of the biaxially strained cubic phase. The DOS effective mass of the holes in the unstrained structure takes the large value of $2.69m_0$.

The relative location of the extrema in the conduction and valence bands can however be altered by applying a strain to the structure. We have applied strains up to 4% on the structure in the [100] and [010] directions. A clear trend could be seen upon increase of strain in both cases. Strain is found to reduce m_{DOS} for the holes because of other states in the zone near the gap are decreased in energy with respect to those at Γ . But this strain is also seen to result in the lowering of conduction band

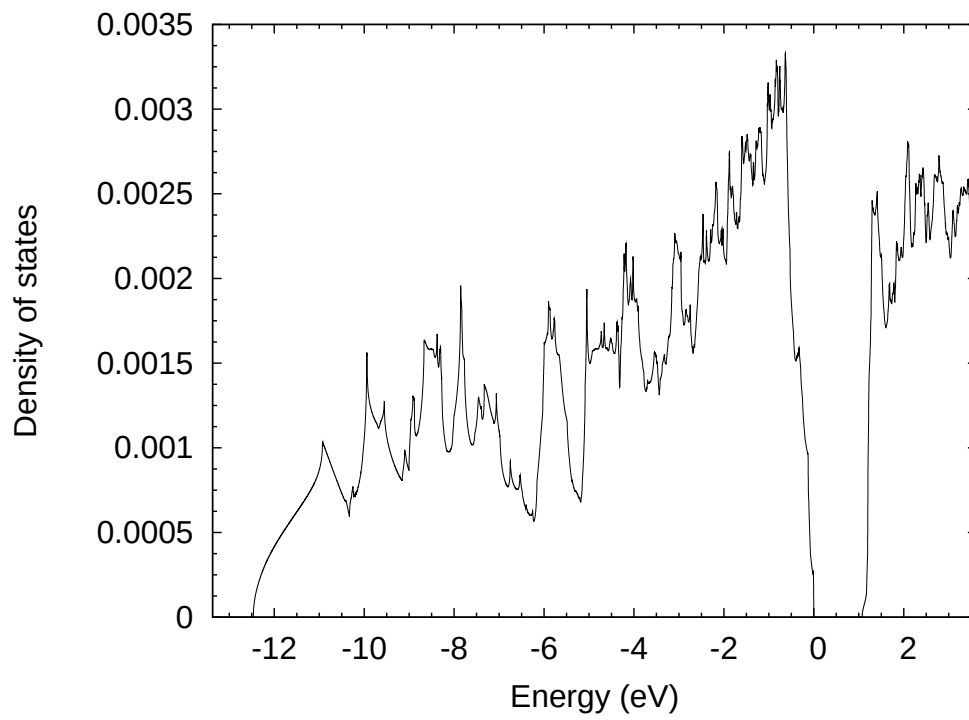


Figure 3.6: Density-of-states per unit volume, per spin, and per eV for silicon in the ST12 phase.

states higher in energy which leads to a large increase in m_{DOS} for the electrons and hence decreases their mobility.

Although the large density-of-states near the gap in ST12 makes it unpromising for high-mobility applications, it could have interesting consequences for superconductivity if it were possible to dope into one of the large peaks in the DOS. If we consider rigid-band doping, it would take a unrealistic amount of doping to dope into the largest peak in the valence band. The conduction band, however, can be reached with approximately 4% electron doping. In order to get an estimate of the effect that the observed density-of-states can have on T_c we use *ab initio* results from a study on superconductivity in doped cubic silicon [102], in which they calculate a λ of 0.3 for phosphorus n-doped cubic silicon, and employ the following relation connecting the electron-phonon coupling constant to the density-of-states at the Fermi level [103]

$$\lambda = \frac{N(0)\langle g^2 \rangle}{M\langle \omega^2 \rangle} \quad (3.4)$$

The electron-phonon matrix elements and phonon frequencies, being associated with the Si-Si sp^3 bond, should be similar for the cubic phase and ST12. This approximation allows us to obtain an estimate for T_c by a scaling of $N(0)$, the density-of-states at the Fermi level. Scaling the 4.2% rigid-band doped ST12 to that of phosphorus doped cubic silicon we obtain a λ of 0.79. Following the work of Bourgeois and Blase [102] we use a modified McMillan expression [104] and values of the Coulomb coupling constant, μ^* , in the range 0.08-0.12. This results in a T_c between 27-39 K. Though this is a rough estimate of the superconducting transition temperature, it does suggest that if silicon in the ST12 phase were experimentally available superconductivity might be an interesting property to study in more detail.

3.5 Conclusion

Ab initio calculations using DFT/LDA were performed on the silicon phases formed under decompression from β -Sn: R8 (Si XII), BC8 (Si III), hexagonal diamond (Si IV), and the unobserved phase ST12. Quasiparticle energy corrections were implemented using the GW approximation [42] for phases in which the electrical nature, metallic or semiconducting, could have been inaccurately determined in previous calculations due to the ‘band-gap’ problem inherent to DFT calculations.

Our calculations done on the electronic structure of R8 within the GW approximation find that LDA incorrectly predicts an indirect overlap in R8 and that the quasiparticle spectrum exhibits a gap. The effective masses found in R8 are smaller than those found in the cubic diamond phase, and can be made even smaller by the application of strain. This creates the possibility that R8 might be found useful in applications where high carrier mobility is needed.

In studying the BC8 phase we find a direct overlap in the quasiparticle spectrum, in qualitative agreement with calculations done within LDA. This overlap is found to be resilient to attempts at its elimination by the use of strain and doping. We find that the electron effective mass in the hexagonal diamond phase is reduced, as in the case of cubic silicon, with the application of a biaxial strain. Lastly, the yet unobserved ST12 phase is found to have a large number of electronic states near the band edges, which upon doping might provide a sufficient number of states at the Fermi level to result in a higher T_c than what has previously been found in other silicon polytypes.

Chapter 4

Ab initio study of the optical properties of Si-XII

4.1 Introduction

The importance of silicon in electronic and photovoltaic devices has stimulated many studies of its electronic properties. While the cubic diamond phase is by far the most commonly studied, silicon has a large number of polytypes. A study of the structural and electronic properties of these phases was seen in Chapter 3 and additional work can be found in the literature [27, 75–80, 105]. One of these phases, Si-XII, also called R8 because of its rhombohedral unit cell containing 8 atoms, is the focus of this chapter.

Silicon in the R8 structure can be formed upon decompression from the high-pressure metallic β -Sn phase at ~ 10 GPa. The R8 structure remains the dominant phase until approximately 2 GPa when the BC8 (Si-III) structure begins to form [86]. First obtained in angle dispersive x-ray powder diffraction experiments [84], Si-XII has more recently been found in nanoindentation experiments performed on silicon wafers [106–108]. Previous calculations [76, 105] found Si-XII to be a semimetal, exhibiting an indirect band overlap within the local-density approximation (LDA) of density-functional theory (DFT). However, in Chapter 3 we have shown that upon inclusion of quasiparticle self-energy corrections within the GW approximation [32, 40, 42] this indirect overlap is lifted and Si-XII exhibits an indirect band gap of approximately 0.24 eV at ambient pressure [105]. With the current extensive research being done on silicon-based photovoltaic (PV) devices, the presence of another semiconducting silicon polytype with small band gap motivates an interest in its optical properties. In this chapter we present calculations of the optical spectrum of Si-XII with the inclusion of electron-hole interactions. We then compare the obtained optical constants with those found for other forms of silicon.

4.2 Method

Accurate calculations of optical spectra from first-principles involves the inclusion of electron-hole interactions which takes the computation beyond the level of a single-particle picture. A calculation of the optical spectrum performed without taking into account the mutual attraction of quasielectron and quasihole excitations can show substantial deviations from experiment. In the calculation of the frequency-dependent dielectric function $\epsilon(\omega)$ this can result in incorrect positions and amplitudes of characteristic peaks [33]. By incorporating the effects of the electron-hole interactions in the same manner as is done in this study, accurate optical spectra have been obtained for a wide variety of materials. Perhaps most relevant to the current study is the fact that for the optical spectra of silicon in the cubic phase, first studied by theorists over 40 years ago [109], this method reproduces the experimental data [33] to a high degree of accuracy.

The electron-hole interaction results in excitations, characterized by the state $|S\rangle$, that are a linear combinations of free quasielectron and quasihole pair configurations [110]:

$$|S\rangle = \sum_{\mathbf{k}} \sum_v^{hole} \sum_c^{elec} A_{v\mathbf{k}}^S \hat{a}_{v\mathbf{k}}^\dagger \hat{b}_{c,\mathbf{k}+\mathbf{Q}}^\dagger |0\rangle \quad (4.1)$$

where $\hat{a}_{v\mathbf{k}}^\dagger$ creates a quasihole in state $\mathbf{k}$ and $\hat{b}_{c,\mathbf{k}+\mathbf{Q}}^\dagger$ creates a quasielectron in state $\mathbf{k} + \mathbf{Q}$ following the absorption of a photon with wavevector $\mathbf{Q}$. The electron-hole excitations represented by $|S\rangle$ and their energies Ω_S can be calculated by solving the equation of motion for the two-body Green's function using the the Bethe-Salpeter equation (BSE) [111]:

$$\begin{aligned} (E_{c,\mathbf{k}+\mathbf{Q}}^{QP} - E_{v\mathbf{k}}^{QP}) A_{v\mathbf{k}}^S + \sum_{\mathbf{k}'} \sum_{v'}^{hole} \sum_{c'}^{elec} \langle v\mathbf{k} | K^{eh} | v'\mathbf{k}' \rangle A_{v'\mathbf{k}'}^S \\ = \Omega_S A_{v\mathbf{k}}^S \end{aligned} \quad (4.2)$$

We apply the BSE within the framework described by Rohlfing and Louie [33], in which a DFT calculation for the electronic ground state is used together with a calculation for the electron self-energy within the GW approximation to obtain the quasiparticle spectrum. The BSE is then solved yielding the correlated electron-hole excitations $|S\rangle$ and their energies Ω_S . Once the states describing the excitations and their associated energies have been solved for, it is then possible to evaluate the optical spectrum.

4.3 Results

4.3.1 Optical Spectrum

In our solution of the BSE in Eq. 4.2 we have included the highest 11 valence bands (v=6-16) and the lowest 12 conduction bands (c=17-28). The electron-hole interaction was calculated for 216 $\mathbf{k}$ points in the Brillouin zone. We then used an interpolation scheme [33] to obtain the matrix elements on a finer mesh of 1000 $\mathbf{k}$ points in the Brillouin zone. This fine mesh is not symmetric with respect to the Brillouin zone, but rather is shifted off the high-symmetry directions to obtain a finer sampling of the spectra. An artificial Lorentzian smearing of 150 meV is included in these calculation to account for the finite sampling of the Brillouin zone.

Once the BSE is solved, the imaginary part of the dielectric function, $\epsilon_2(\omega)$, is calculated from

$$\epsilon_2(\omega) = \frac{16\pi^2 e^2}{\omega^2} \sum_S |\vec{\lambda} \cdot \langle 0 | \vec{v} | S \rangle|^2 \delta(\omega - \Omega_S) \quad (4.3)$$

In the above expression $\vec{\lambda}$ is the polarization of the incident light and $\vec{v}$ is the single-particle velocity operator. In the absence of electron-hole interactions, Eq. 4.3 reduces to the following expression involving photon-induced vertical transitions between the independent electron and hole states

$$\epsilon_2^{(0)}(\omega) = \frac{16\pi^2 e^2}{\omega^2} \sum_{v,c} |\vec{\lambda} \cdot \langle v | \vec{v} | c \rangle|^2 \delta(\omega - (E_c - E_v)) \quad (4.4)$$

where the superscript is used to distinguish the fact that electron-hole interactions are not included.

In Fig. 4.1 we plot $\epsilon_2(\omega)$ both with and without the inclusion of electron-hole interactions. The differences between these two curves might be interpreted to arise from the attractive nature of the electron-hole interaction, resulting in the shift of the spectrum to lower energies. This explanation is seen to be incorrect upon the examination of the joint density of states (dotted line) and the density of excitonic states (solid line) as shown in Fig. 4.2. We have divided the joint density of states by ω^2 to aid comparison with $\epsilon_2(\omega)$. These two density of states are seen to be nearly identical and are not related to each other by a shift in energy. This is not unexpected as the energy scale of the electron-hole interaction is related to the exciton binding energy in the material, which in Si-XII is on the order of meV and is thus small on the scale of our plot. Thus the difference between the interacting and the non-interacting spectrum arises from the constructive superposition of oscillator strengths at low energies and the destructive superposition at higher energy. This is precisely the behavior that has been found in calculations for other semiconductors [33, 112–115].

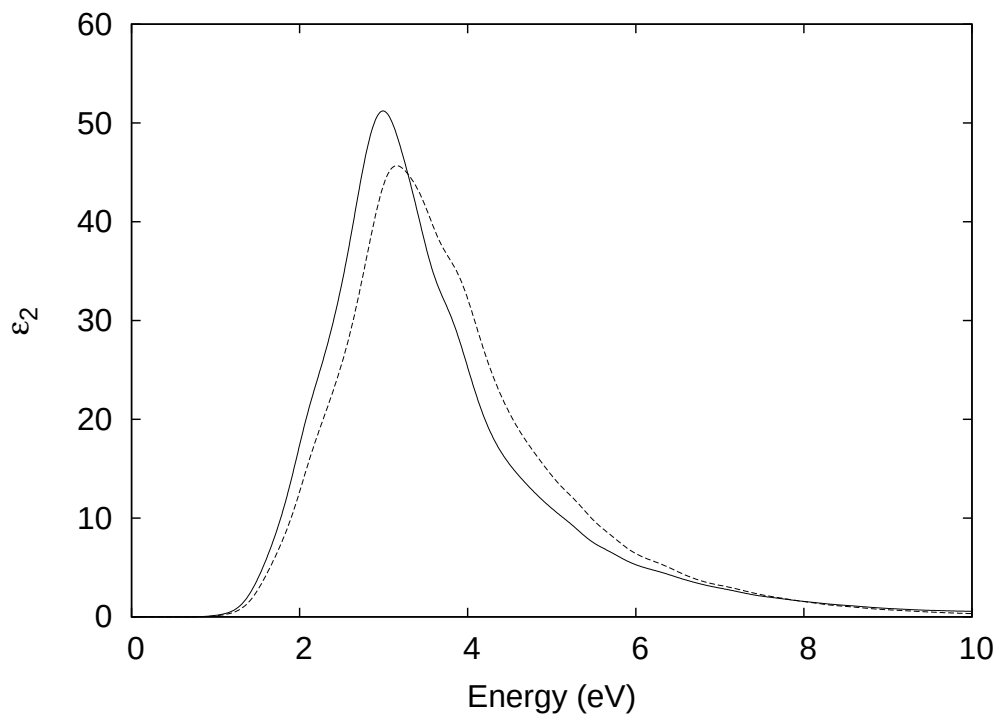


Figure 4.1: Calculated imaginary part of the dielectric function of Si-XII. The solid line (dotted line) is with (without) the inclusion of electron-hole interactions.

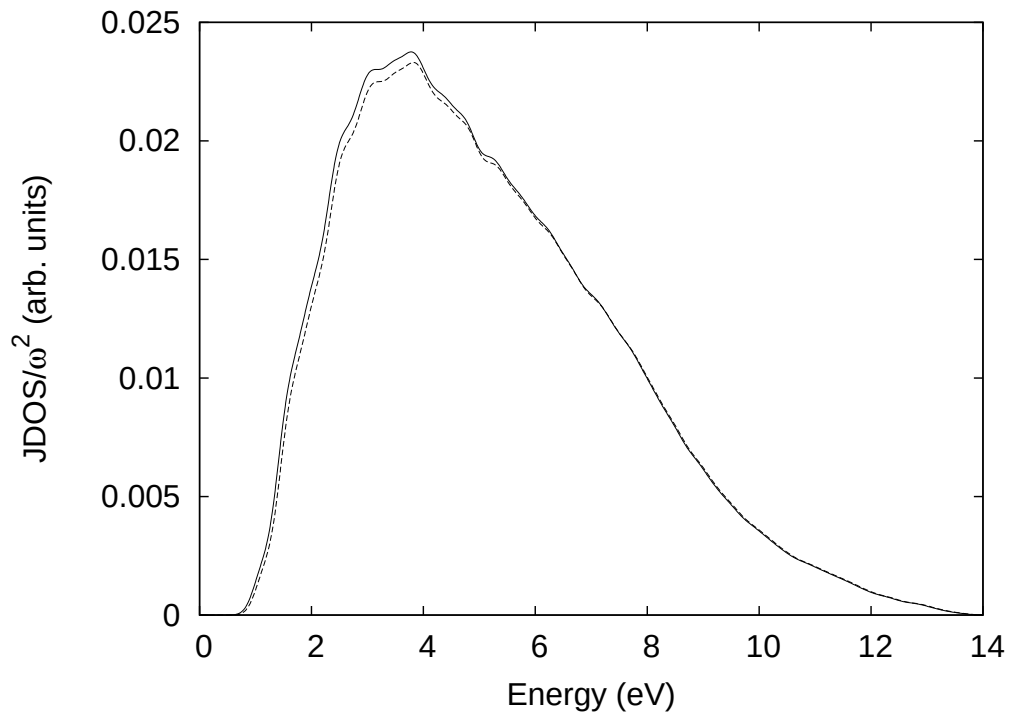


Figure 4.2: Calculated joint density of states (divided by ω^2) for Si-XII. The solid line is the density of excitonic states, while the dashed line is joint density of states in absence of the electron-hole interaction.

4.3.2 Photovoltaic application

To facilitate the evaluation of Si-XII as an absorber material in a photovoltaic device we compare the optical constants relevant to photon absorption with some other forms of silicon: amorphous, polycrystalline, and the cubic crystalline phase. These three forms currently dominate the market for photovoltaic modules in commercial applications [116]. In Fig. 4.3 the imaginary part of the dielectric function of Si-XII has been compared with these other forms of silicon commonly used in solar applications. The measured dielectric functions for these materials is taken from Ref. 117 which are obtained from spectroscopic ellipsometry. The calculated value of $\epsilon_2(\omega)$ for Si-XII clearly has more weight at lower energies. As mentioned previously, this is expected because of the smaller direct quasiparticle gap of Si-XII (~ 1.2 eV) compared to that of cubic silicon (~ 3.4 eV). This is advantageous for solar applications because the incident photon flux is concentrated at energies less than 3 eV. In Fig. 4.4 we show the optical absorption coefficient α for these four materials along with the AM 1.5 solar spectral irradiance [118] to highlight the energies relevant to photovoltaic applications. Silicon in the R8 phase is seen to absorb light more effectively than the other forms of silicon in this energy range. We point out that this result holds in spite of the fact that our calculated values for Si-XII do not include any indirect, i.e. phonon-assisted, transitions, whereas the experimental data that it is being compared to does include these transitions. Since the gap in R8 silicon is indirect, we should expect that if these indirect transitions were included in the calculation of the absorption coefficient that R8 would have even more weight in the photovoltaic relevant region of the energy spectrum.

Since Si-XII has a larger absorption coefficient at lower energies, it could in principle allow the use of photovoltaic absorber layers of smaller thickness to be used. This would result in less material being needed for photovoltaic devices of similar absorptive power. This difference can be seen in a calculation of the absorbed energy flux, W , as a function of the length of the absorber material,

$$W(L) = \int_0^\infty (1 - e^{-\alpha(E)L}) I(E) dE \quad (4.5)$$

where α is the absorption coefficient, I is the solar spectral irradiance, and L the length of the absorber film. In this expression we have neglected the effects of reflection off the front and back surfaces of the film. The reason for this is that reflection often plays a small role in PV applications because of the use of anti-reflection coatings, often made of silicon nitride or titanium oxide. A coating of this type applied to a silicon surface can reduce the percentage of light reflected by the surface from the bare value of $\sim 30\%$ down to only a few percent [119]. In Fig. 4.5 we have compared this quantity among the forms of silicon studied. The data has been normalized to the incident energy flux ($L \rightarrow \infty$ in Eq. 4.5). Si-XII clearly absorbs a greater fraction of the incident flux, with a very marked contrast at small values of

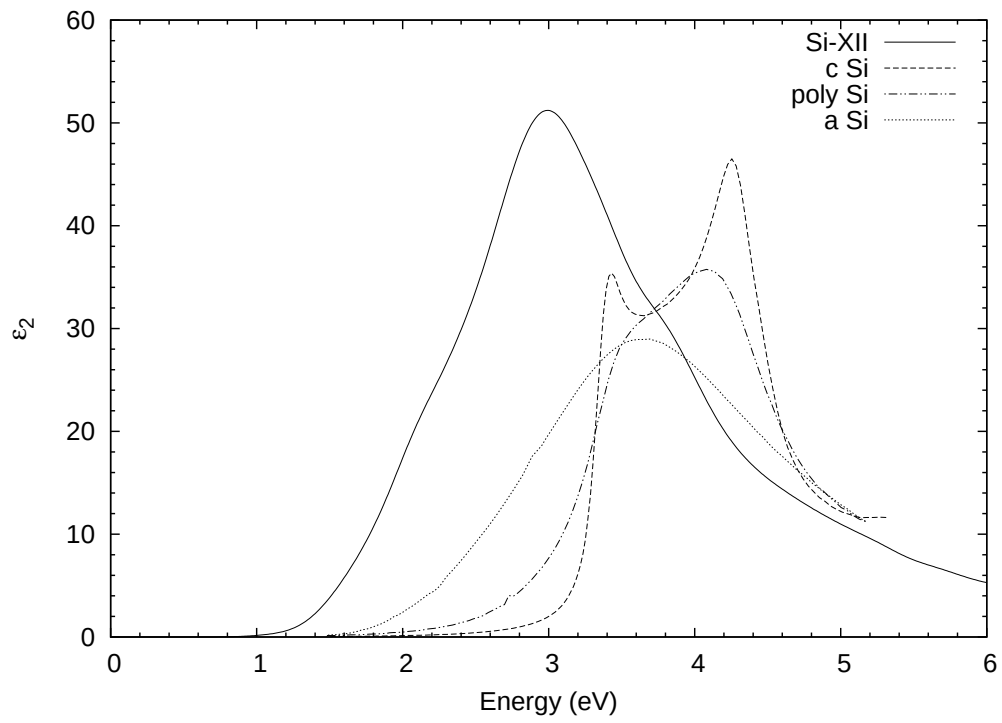


Figure 4.3: Calculated $\epsilon_2(\omega)$ for Si-XII along with the experimental values for silicon in the amorphous (a Si), polycrystalline (poly Si), and crystalline cubic phase (c Si). The experimental data is taken from Ref. 117.

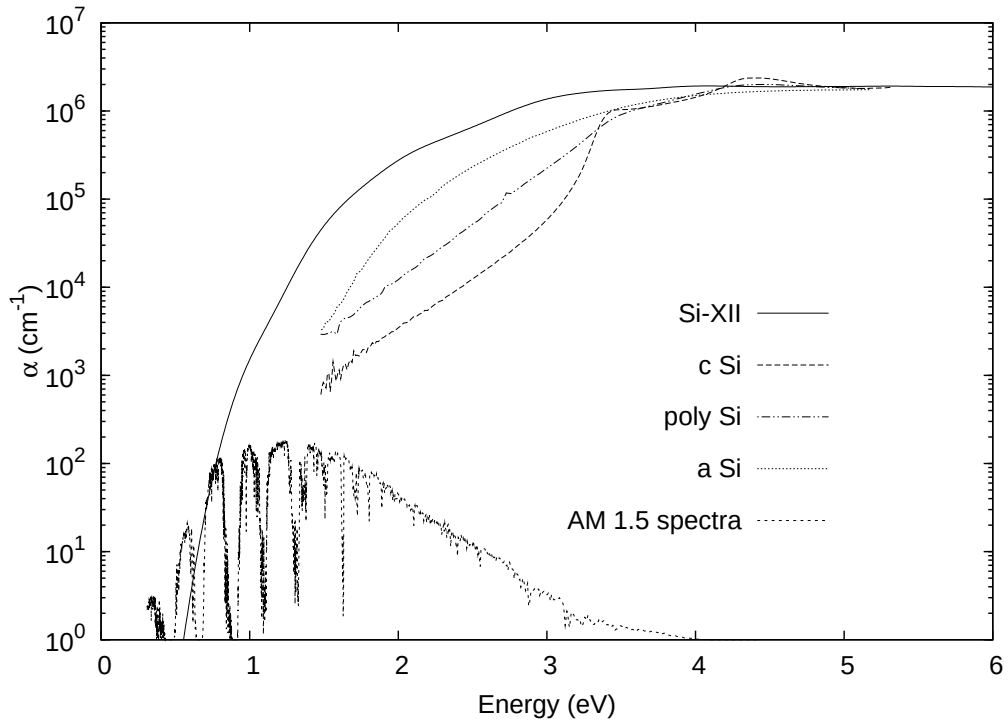


Figure 4.4: Optical absorption coefficient α for for Si-XII and three other silicon forms commonly used in photovoltaic devices. Shown inset (in arbitrary units and linear scale) is the solar spectral irradiance for AM1.5. The absorption coefficients for c-Si, a-Si, and poly-Si are experimental and are taken from Ref. 117.

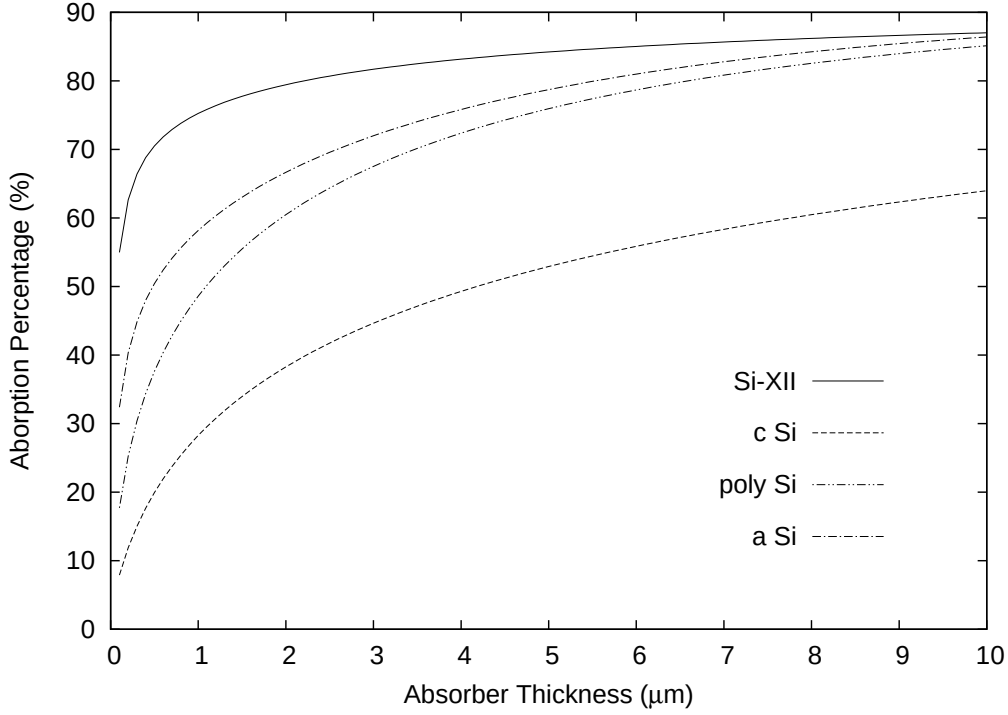


Figure 4.5: Percentage of incident energy flux which is absorbed as a function of film thickness for Si-XII and the other forms of silicon studied. Reflection at the surfaces has not been included.

the absorber thickness.

4.4 Conclusion

In conclusion we have calculated the optical absorption spectrum of the silicon polytype Si-XII (R8) using a first-principles method which is based on the GW approximation to solve for the independent particle quasiparticle spectrum followed by a solution of the Bethe-Salpeter equation to account for electron-hole interactions. We note that to our knowledge this is the first published calculation of the optical spectrum of Si-XII. We find that the imaginary part of the dielectric function, $\epsilon_2(\omega)$, in addition to the absorption coefficient $\alpha(\omega)$ shows greater overlap with the solar spectrum than three other silicon forms used for comparison. These silicon forms are the cubic, amorphous, and polycrystalline forms which commonly find use in photovoltaic application. There are of course other features besides absorption which are relevant to determining the efficiency in a photovoltaic device. One of these is the band gap which places an upper bound on the open-circuit voltage V_{oc} . The small, indirect band gap of Si-XII of ~ 240 meV might then limit its efficiency when used

for photovoltaic applications.

Chapter 5

Electronic and optical properties of body-centered tetragonal Si and Ge

5.1 Introduction

Silicon and germanium play an important role in countless electronic and photovoltaic devices that are at the heart of modern technological development. Their importance necessitates an understanding of their properties, in particular their electronic and optical characteristics. The cubic phases of both of these materials have been extensively studied over the years and have been the topic of a large number of publications. However, while much about the cubic phases is known, much less is known about the various polytypes that these elements form. The interest in obtaining a better understanding of the other polytypes that these materials form is not only interesting from a scientific perspective, but because a number of these phases have common characteristics with their cubic counterparts they may have possible applications in the same domains of use. Additionally, many of these phases are the subject of current experimental study and thus a good theoretical understanding of these materials is desired.

One of the most recent developments in the understanding of the phase diagram of silicon and germanium has been the prediction of a new four-fold coordinated phase in the body-centered tetragonal (bct) structure [120]. This phase had been recently discovered in studies of carbon through molecular-dynamics simulations as a crystalline phase that carbon nanotubes form under a pressure of 20 GPa. [121] Because of the similarities between the group-IVA materials, this phase was expected to occur for silicon and germanium. The *ab initio* calculations performed in the first study of the bct Si and Ge structures suggest that this structure would be metastable with a total energy higher than the cubic structure by only ~ 0.1 eV/atom [120].

Upon the inclusion of quasiparticle corrections using the GW approximation (GWA) [40, 42] silicon in the bct structure was found to be semiconducting with an indirect band gap of 0.47 eV [120]. Germanium, on the other hand, was found to have an indirect overlap in the LDA which persists even after the quasiparticle corrections. While the bct phase in Si and Ge has yet to be observed experimentally, it has been suggested that a possible path for obtaining this structure is hetero-epitaxial growth on the (111) surface of a Ge substrate in the case of Si and the (111) surface of a Ge-Sn buffer layer in the case of Ge [120].

The discovery of a small band gap silicon polytype motivates an interest in its optical properties since silicon-based materials currently dominate the photovoltaics market. In Chapter 4 we investigated the optical properties of another silicon polytype, the high-pressure phase Si-XII [122]. It was found that the optical absorption coefficient showed a greater overlap with the incident solar spectra than that of three other forms of silicon used for comparison, namely the cubic, amorphous, and polycrystalline forms. This characteristic could result in the material finding use in photovoltaic applications. As another small band gap silicon semiconductor, the bct phase is also expected to have an optical spectrum conducive to its use in photovoltaic applications. Similar expectations are reasonable for bct germanium.

In this chapter we first examine the quasiparticle excitation spectra of the bct phases of Si and Ge using the GW approximation (GWA) [40, 42]. We then solve the Bethe-Salpeter equation for the optical spectra which incorporates the mutual attraction of quasielectrons and quasiholes [33]. Taking into account this mutual interaction can be crucial when comparing to experimental measurements [33, 123–126]. We then examine the optical spectra of these materials and compare the bct phases to their cubic counterparts.

5.2 Method

We have performed *ab initio* pseudopotential calculations within the framework of density-functional theory (DFT) in the local density approximation (LDA) to the exchange-correlation functional. The interaction of the valence electrons with the ion cores is represented by a norm-conserving pseudopotential [24]. These calculations are used to evaluate total energies, for structural optimization, and as inputs to the subsequent quasiparticle and many-body calculations. The electronic wave functions are expanded in a plane-wave basis [11, 29] with a kinetic energy cutoff of 30 Ry.

Quasiparticle spectra are calculated in the “one-shot” GW approximation to the electron self-energy following Hybertsen and Louie [42]. In this approach, the self-energy is approximated by the product of the single-particle Green’s function and the screened interaction [40]. The Green’s function is approximated using states and energies resulting from the LDA calculations and the frequency-dependent dielectric

Table 5.1: The minimum indirect gap E_i and direct gap E_d of bct Si and bct Ge, calculated within the local density approximation (LDA) and within the GW approximation (GWA). The results from the present work are compared with those found by Fujimoto *et al.* [120]

	E_i (eV)		E_d (eV)	
	Ref. 120	Present Work	Ref. 120	Present Work
bct Si (LDA)	0.10	0.15	0.25	0.32
bct Si (GWA)	0.47	0.86	0.74	1.07
bct Ge (LDA)	-0.39	-0.23	0.11	0.15
bct Ge (GWA)	-0.12	0.38	0.51	0.79

matrix needed to compute the screened interaction is obtained by extending the static dielectric matrix to finite frequencies using a generalized plasmon-pole model [42]. This approach has been seen to dramatically improve the agreement with experiment with regards to the band gap of materials than results from simply interpreting the DFT eigenvalues as the quasiparticle energies.

In order to calculate the optical spectra of materials accurately from first principles, the effects of the electron-hole interaction must be included. This necessitates going beyond the quasiparticle calculations discussed above to examine excitations that are characterized by linear combinations of free quasielectron and quasihole pair configurations. These excitations, denoted by $|S\rangle$, are given by

$$|S\rangle = \sum_{\mathbf{k}} \sum_v^{hole} \sum_c^{elec} A_{v\mathbf{k}}^S \hat{a}_{v\mathbf{k}}^\dagger \hat{b}_{c,\mathbf{k}+\mathbf{Q}}^\dagger |0\rangle \quad (5.1)$$

where $\hat{a}_{v\mathbf{k}}^\dagger$ creates a quasihole in state $\mathbf{k}$ and $\hat{b}_{c,\mathbf{k}+\mathbf{Q}}^\dagger$ creates a quasielectron in state $\mathbf{k} + \mathbf{Q}$ following the absorption of a photon with wavevector $\mathbf{Q}$. These states and their corresponding energies Ω_S can be calculated by solving the equation of motion for the two-body Green's function using the Bethe-Salpeter equation [111]:

$$\begin{aligned} (E_{c,\mathbf{k}+\mathbf{Q}}^{QP} - E_{v\mathbf{k}}^{QP}) A_{v\mathbf{k}}^S + \sum_{\mathbf{k}'} \sum_{v'}^{hole} \sum_{c'}^{elec} \langle v\mathbf{k} | K^{eh} | v'\mathbf{k}' \rangle A_{v'\mathbf{k}'}^S \\ = \Omega_S A_{v\mathbf{k}}^S \end{aligned} \quad (5.2)$$

The solution of this equation in the limit of $Q \rightarrow 0$ is performed within the framework described by Rohlfing and Louie [33]. Once this equation is solved for the correlated electron-hole excitations $|S\rangle$ and their energies, it is possible to evaluate the optical spectrum.

5.3 Results and discussion

The structures of bct Si and Ge are described in the original work of Fujimoto *et al.* [120] We have relaxed the structures from their original parameters and find that their original structural description is consistent with ours. The calculated gaps in the bct phase are shown in Table (5.1) where they are compared with the results of Fujimoto *et al.* [120] The calculated values of the gaps within the local density approximation are in good agreement. We note, however, that our indirect LDA gap is somewhat larger than that calculated in Ref. 120 for bct Ge. As in Ref. 120, we find the direct gap to be at the Z point for both materials. The indirect gap is between the top of the valence band at Z, for both materials, and the bottom of the conduction band which is along a line connecting Γ to Z in the case of Si and at the P point in the case of Ge [120].

Despite the good agreement between the calculated values of the gaps within the LDA, the values of the gaps calculated within the GW approximation are noticeably different. The values calculated in the present work are larger by an average of 0.38 eV. This difference may be attributed to the different implementations of the GW approximation used in calculating the effect of the electron self-energy. As discussed earlier, in this work we utilize a planewave implementation of the GWA in which the static dielectric matrix is extended to finite frequencies with the use of a generalized plasmon-pole model [42]. We calculate the static dielectric matrix $\epsilon_{G,G'}^{-1}(\mathbf{q}, \omega = 0)$ up to a kinetic energy cutoff of 24 Ry and the self-energy Σ by summing over 59 q points in the irreducible Brillouin zone and 700 bands (8 occupied and 692 unoccupied). In contrast, the calculations done in Ref. 120 are implemented in an all-electron, augmented-wave implementation of GWA using eigenfunctions generated by the full-potential linear muffin-tin orbital method (FP-LMTO) [120, 127]. In addition to this difference, the calculations of Ref. 120 also involve the computation of the full-frequency dependence of the screened Coulomb interaction W and do not make use of the plasmon-pole approximation [127]. It is found that in the method of the present work, the calculated band gap of cubic silicon is 1.27 eV whereas that found in Ref. 120 is 1.03 eV (experimental measurement is 1.17 eV). This difference in the results of the two methods is also seen in other calculations in the literature, giving an indication of the precision of this approach. The original calculations of Hybertsen and Louie [42] find a indirect gap of 1.29 eV in the case of cubic silicon whereas calculations done by Faleev *et al.* using the FP-LMTO method find a one-shot GW energy gap of 0.84 eV [128]. Thus the fact that we find the bct quasiparticle gap to be somewhat larger in our pseudopotential plane-wave approach than that found in Ref. 120 using an all-electron implementation is consistent with previous results.

The difference in the magnitude of the quasiparticle corrections for the bct phases from the current work and that of Ref. 120 gives rise to a qualitative difference in the nature of the gap in the case of germanium. In Ref. 120 it was found that bct Ge is semimetallic with an indirect overlap of 0.12 eV. Our calculations suggest

that the indirect overlap present in the LDA is increased to an indirect gap of 0.38 eV, thereby predicting that bct Ge should be a small band gap semiconductor. It should be noted that this conclusion doesn't rest on the fact that the calculations of Ref. 120 find a somewhat larger overlap in the LDA as our calculated quasiparticle corrections are larger than that of Ref. 120 by more than the difference between our LDA indirect gaps. However, it's important to note that in the quasiparticle calculations for germanium, the occupancies of the states are as they would be in the semiconducting case. This avoids the complexity of representing the intraband transitions in the small region of overlap where the Fermi level crosses the band in the LDA as it would in a true metal.

Once the quasiparticle spectrum has been obtained, we are in a position to solve the Bethe-Salpeter equation (BSE) shown in Eq. 5.2. We have included all 8 valence bands ($v=1-8$) and the lowest 12 conduction bands ($c=9-20$) in our solution. The calculation of the electron-hole interaction matrix elements is a computationally intensive step and so they are calculated on a coarse grid of 512 $\mathbf{k}$ points before an interpolation procedure is used to obtain them on a finer grid of 1000 $\mathbf{k}$ points [33]. We then solve the BSE yielding the coupled excitations $|S\rangle$ and the associated energies Ω_S .

Once these excitations have been solved for using the BSE, the optical spectrum can be evaluated. The imaginary part of the dielectric function, $\epsilon_2(\omega)$, is calculated using

$$\epsilon_2(\omega) = \frac{16\pi^2 e^2}{\omega^2} \sum_S |\vec{\lambda} \cdot \langle 0 | \vec{v} | S \rangle|^2 \delta(\omega - \Omega_S) \quad (5.3)$$

In this expression $\vec{\lambda}$ is the polarization of the incident light and $\vec{v}$ is the single-particle velocity operator. In the absence of electron-hole interactions, Eq. 5.3 reduces to the following expression involving photon-induced direct transitions between the independent electron and hole states

$$\epsilon_2^{(0)}(\omega) = \frac{16\pi^2 e^2}{\omega^2} \sum_{v,c} |\vec{\lambda} \cdot \langle v | \vec{v} | c \rangle|^2 \delta(\omega - (E_c - E_v)) \quad (5.4)$$

where the superscript is used to distinguish the fact that electron-hole interactions are not included.

In Fig. 5.1 we plot $\epsilon_2(\omega)$ for silicon and germanium in the bct structure using the method described earlier. For both materials it is clear that the electron-hole interaction shifts the spectra to lower energies. This behavior is seen in other semiconductors and occurs not because of a shift to lower energies due to the attractive electron-hole interaction, but rather due to a constructive superposition at lower energies of the non-interacting matrix elements that make up the coupled excitations $|S\rangle$ [33].

In comparing the absorption spectra between these two materials it is interesting that while both materials have direct gaps near 1 eV, only the germanium phase has

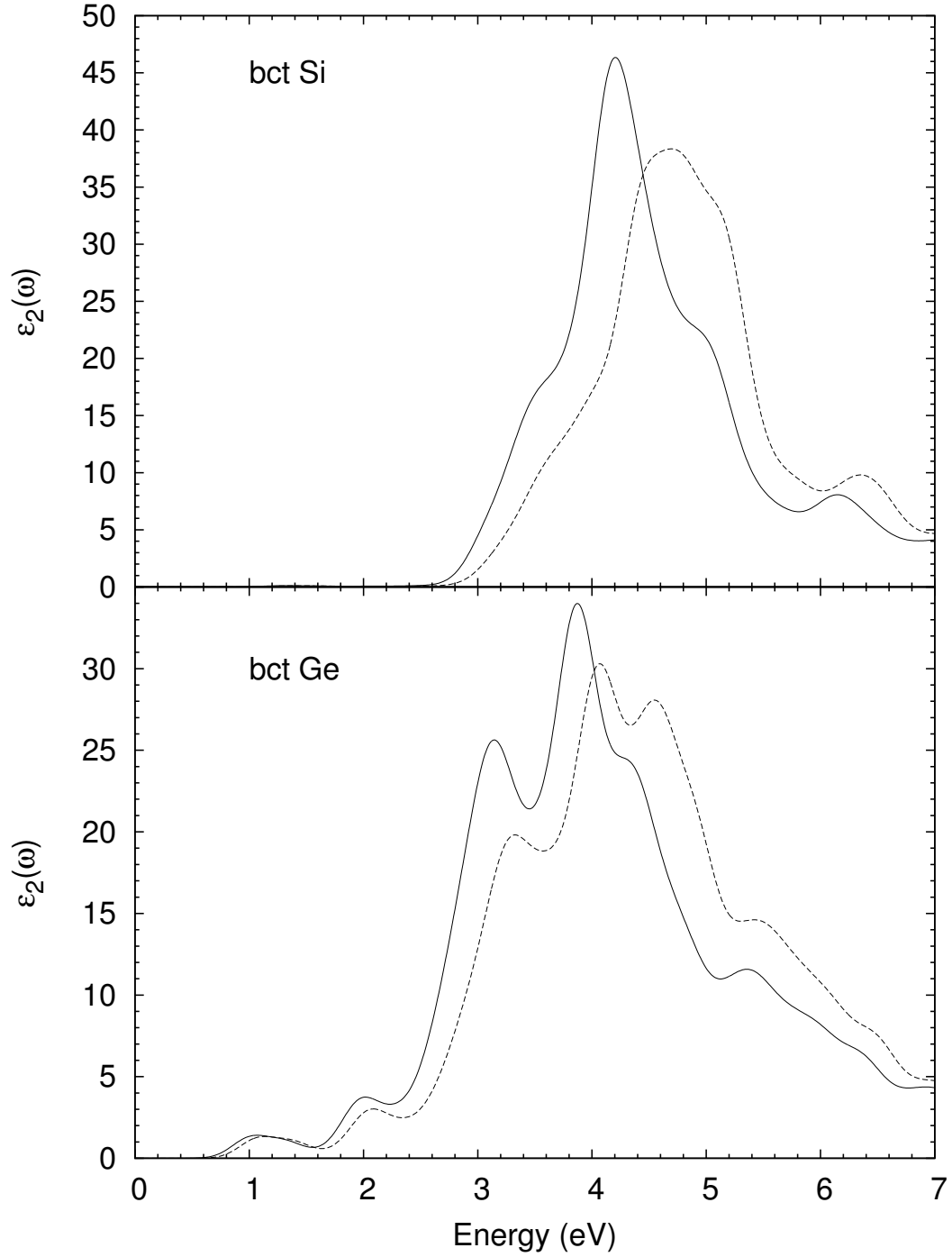


Figure 5.1: Calculated $\epsilon_2(\omega)$ of the bct phase of silicon (upper panel) and germanium (lower panel). The solid line is including electron-hole interactions (Eq. 5.3) while the dotted line is the non-interacting calculation (Eq. 5.4).

any substantial weight in this energy region. The cause of this can be understood by examination of Fig. 5.2. In Fig. 5.2 we have plotted the joint density of states (divided by ω^2 to make it comparable to the absorption spectra in Fig. 5.1) as well as the average squared matrix element, $|\vec{\lambda} \cdot \langle v | \vec{v} | c \rangle|^2$, as a function of energy. While the joint density of states for the two materials are extremely similar, the average squared matrix element for the silicon phase is very small up until after 2.5 eV, which explains why the absorption curve in Fig. 5.1 doesn't become substantial until around this energy.

In order to obtain a measure of the possible utility these phases have for photovoltaic applications we compare the absorption spectra with their cubic counterparts. The cubic phases of both silicon and germanium find widespread use in photovoltaic applications. Silicon-based solar cells dominate the photovoltaics market while germanium is a component in highly efficient multijunction solar cells [130]. While the efficiency of a solar cell certainly depends on other characteristics, our analysis will focus only on the overlap of the absorption coefficient with the incident solar spectrum. In Fig. 5.3 we plot $\epsilon_2(\omega)$ for silicon in the bct and cubic structures calculated using the method described in this work. The experimental dielectric function for cubic silicon is taken from Ref. 129 and shows the reliability of the method in calculating optical spectra from first principles. The calculations for cubic silicon also show good agreement with others in the literature [124, 125]. The calculated $\epsilon_2(\omega)$ for bct silicon shows more weight than cubic silicon at the lowest energies for which the dielectric function is nonzero. However, above ~ 3.3 eV the value for the cubic phase has a greater magnitude. In terms of the relevance of these curves to photovoltaic absorption in materials, the important energy range is that for which the incident solar flux is appreciable.

In Fig. 5.4 we show the calculated absorption coefficients of silicon in the bct and cubic phase. Shown also is the air mass (AM) 1.5 solar spectral irradiance [118]. It can be seen that the absorption coefficient of bct silicon is larger for energies where the incident photon flux is greater. The more complete the overlap between the absorption coefficient and the incident solar spectrum the less material that is needed to capture a given quantity of light. It is also important to note that in the calculations only direct transitions are taken into account. With the inclusion of phonon-assisted, i.e. indirect, transitions we would expect the absorption coefficient of silicon in the bct phase to have a finite values at lower energies than for the case of cubic silicon arising from the smaller indirect band gap found in the bct phase. While there are other considerations that must be taken into account when building a solar device, the strong overlap between the absorption coefficient of bct silicon and the solar spectrum owing to the smaller band gap of this form of silicon shows promise for use in photovoltaic devices.

We now turn to the case of germanium in the bct phase. In Fig. 5.5 we show the calculated $\epsilon_2(\omega)$ for germanium in the bct structure compared with experimental data on its cubic phase from Ref. 131. Compared to its cubic counterpart, bct Ge

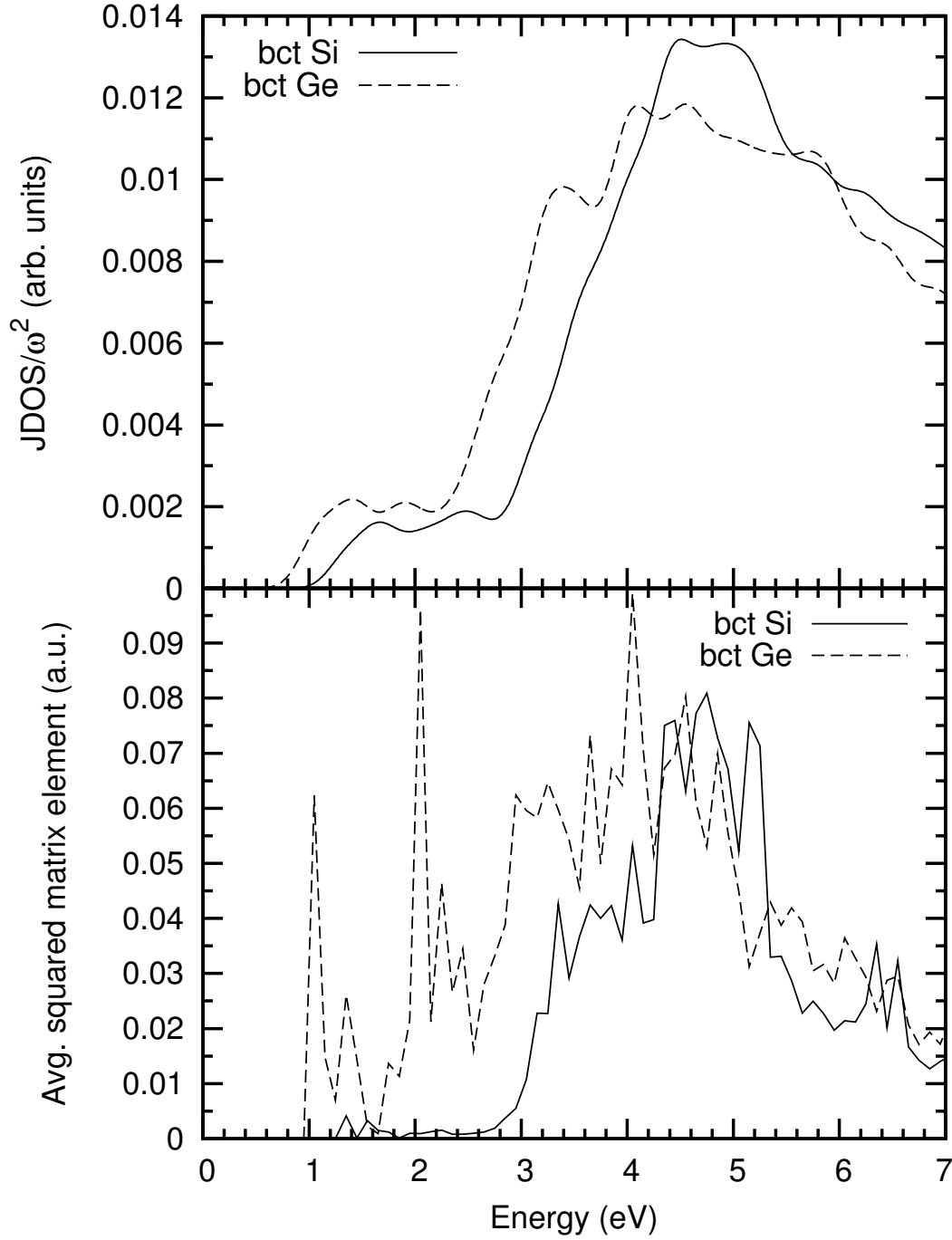


Figure 5.2: Upper panel: calculated joint density of states (JDOS), divided by ω^2 , for silicon and germanium in the bct phase. Lower panel: average squared matrix element $|\vec{\lambda} \cdot \langle v | \vec{v} | c \rangle|^2$ for the two calculations. The size of the energy bins in the lower panel is 0.1 eV.

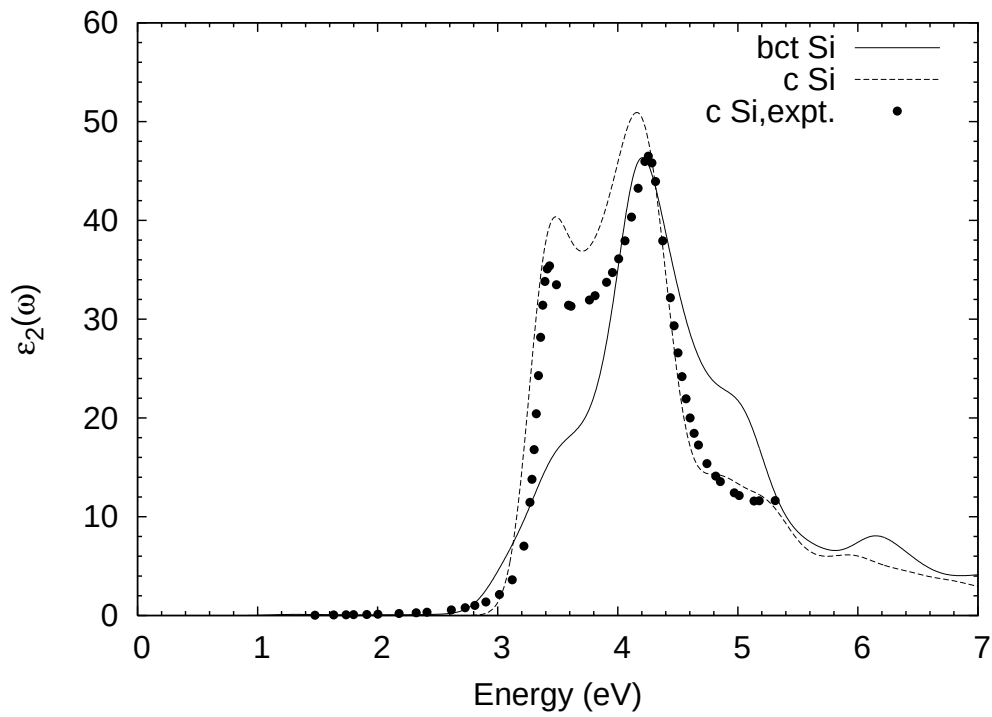


Figure 5.3: Calculated $\epsilon_2(\omega)$ of silicon in the bct and cubic phases. The calculated curve for cubic silicon (c Si) is compared to the experimental data from Ref. 129 and indicates the reliability of the present method for calculating optical spectra. The calculated curves include an artificial broadening of 0.15 eV

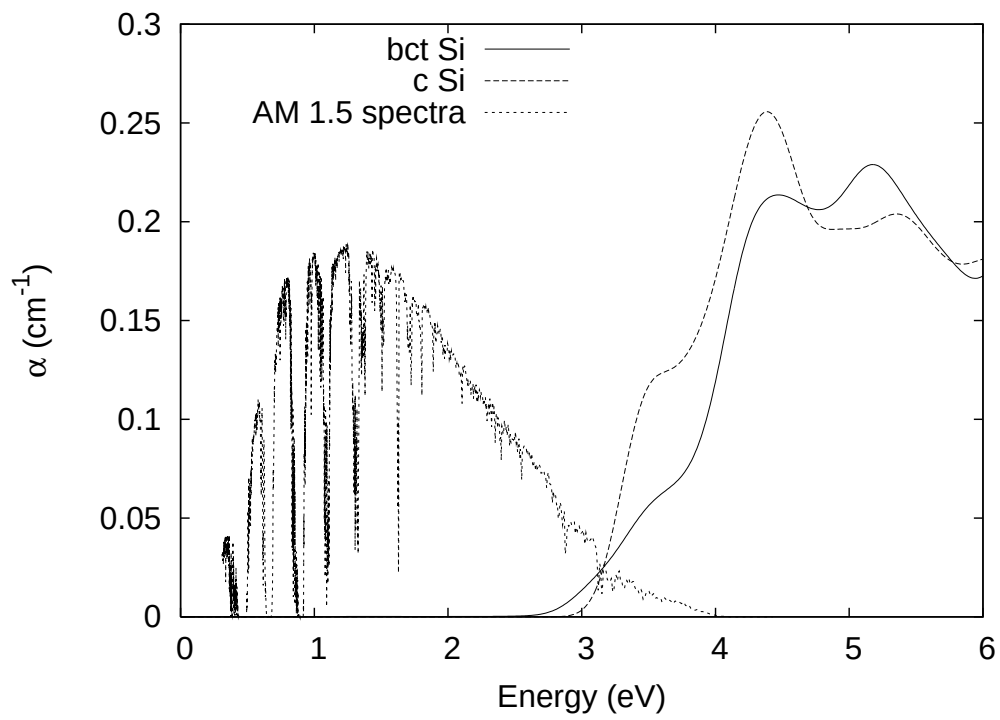


Figure 5.4: Calculated optical absorption coefficient α of silicon in the bct and cubic structures. Shown also in arbitrary units is the solar spectral irradiance for AM 1.5.

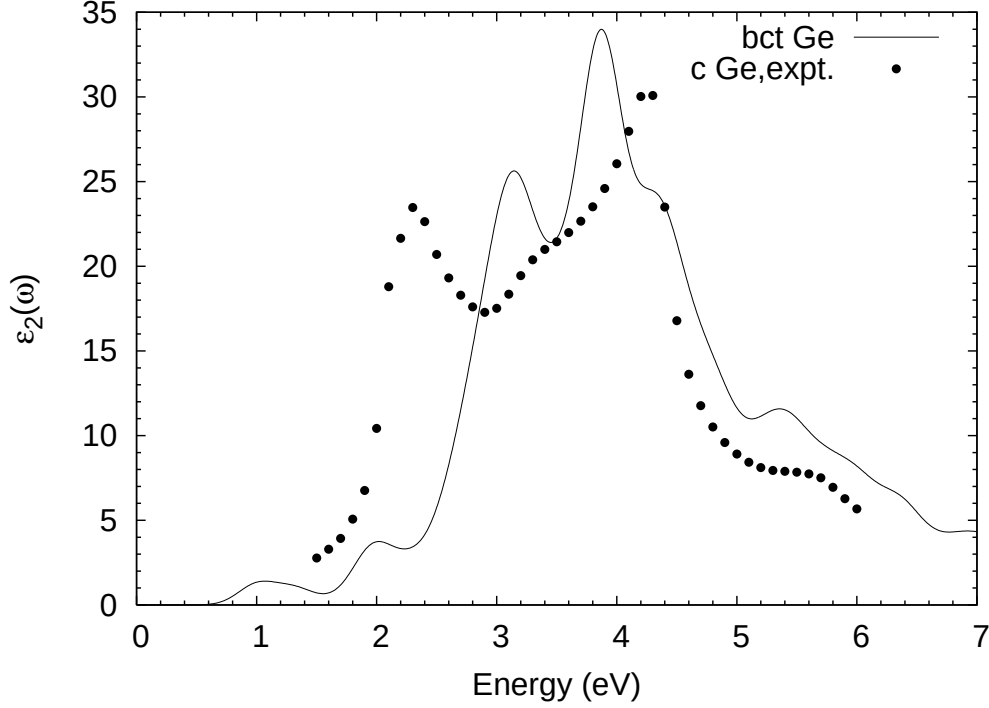


Figure 5.5: Calculated $\epsilon_2(\omega)$ of germanium in the bct and compared to experimental data on cubic germanium taken from Ref. 131. The calculated curves include an artificial broadening of 0.15 eV.

appears to have substantially less weight at lower energies at which the incident solar flux is large. Thus even though germanium's bct phase compares favorably to the bct phase of silicon as seen in Fig. 5.1, the fact that the cubic phase of Ge has a much smaller direct gap of 0.89 eV compared to the direct gap of 3.4 eV in cubic silicon [65, 132] results in the bct Ge phase comparing unfavorably in terms of overlap with the solar spectrum whereas the bct Si phase compares favorably in relation to their respective cubic counterparts.

5.4 Conclusion

In conclusion we have calculated the quasiparticle excitation and optical absorption spectra of the recently predicted bct phase of silicon and germanium. The quasiparticle spectra are calculated using a planewave implementation of the GW approximation which predict that both Si and Ge in the bct structure are semiconductors with indirect gaps of 0.86 eV and 0.38 eV, respectively. The optical absorption spectrum is obtained by solving the Bethe-Salpeter equation to take into account the effect of the electron-hole interactions. The obtained optical coefficients of silicon

in the bct structure show a better overlap with the incident solar spectrum than do those for the cubic structure. In addition to providing optical characterization of this predicted silicon and germanium polytype, the calculations suggest that this phase of silicon may find possible applications within the realm of photovoltaics.

Chapter 6

Electronic structure, equation of state, and lattice dynamics of low-pressure Ge polymorphs

6.1 Introduction

The phase diagrams of group IVa materials such as carbon, silicon, and germanium have generated substantial research interest not simply because of their status as prototypical insulators/semiconductors, but also because of their immense technological importance. The silicon and germanium phase diagrams have strong similarities in the structures reported both upon compression and decompression of experimental samples [79]. Both materials have the cubic diamond (cd) structure under ambient conditions and convert to the metallic β -Sn phase near a pressure of 10-12 GPa. Upon further compression both materials go through a similar sequence of high pressure metallic phases, with the possible exception of the highest pressure transition being hexagonal close-packed (hcp) to face-centered cubic (fcc) in Si while in Ge the highest pressure structure obtained thusfar remains the hcp phase [133].

If samples of Si or Ge are decompressed from the metallic β -Sn phase, they do not return to the cubic form at low temperatures, but instead go to a series of metastable phases. In Si, the slow decompression from the β -Sn phase results in the rhombohedral R8 phase at 9.4 GPa [134]. Further decompression results in the BC8 structure at 2 GPa [86]. In Ge, rapid release from the β -Sn structure to ambient pressure results in the BC8 structure whereas slower release of pressure has resulted in the ST12 structure being recovered [89]. The ST12 phase has not yet been observed in Si and the R8 phase has not yet been definitively observed in Ge. The R8 phase of Ge has been reported in a single experiment but the result was not reproducible [135]. The phase transition pressures in Si are well determined experimentally and have been corroborated with theoretical calculations [76, 136]. In Ge, on the other

hand, fewer experimental and theoretical studies have been carried out. In particular, experimental studies of Ge often quench the pressure to ambient and then study the resulting phases with x-ray diffraction [89, 137]. This is in contrast with experiments performed on Si in which the pressure has been reduced in small steps with the diffraction studies being performed at intermediate pressures in order to detect any resulting change in structure [84].

From a theoretical standpoint, early calculations on the phase stability of Ge polymorphs report that the ST12 phase has a lower enthalpy than the BC8 phase over the pressure range 0-20 GPa [138]. At the time of these calculations the R8 phase of Si had yet to be discovered and was not considered in the study on Ge. Later, after the discovery of R8 Si [84], DFT calculations predicted that the R8 phase of Ge was lower in enthalpy than that of the BC8 phase at moderate compression, and suggested that R8 might be obtained experimentally from compressing the BC8 phase [139]. The small differences in energy shown in the energy-volume relations for the three metastable phases made it difficult to estimate the transition pressure and thus no estimate was given.

The possibility of obtaining Ge in the R8 phase is an interesting one because R8 Si is believed to be a semiconductor [105, 140] with interesting optical properties [122, 141]. Additionally, the theoretical prediction of transition pressures between the metastable germanium phases would be of interest in helping to guide future experimental work in searching for metastable Ge structures, especially those which more carefully explored the pressure regime intermediate between ambient pressure and that where the β -Sn phase is stable. As noted earlier, the analogous theoretical transition pressures in Si are in good agreement with experimental observation [76, 136], and thus it might be expected for the same to be true in Ge.

In this chapter we focus on obtaining more information on the properties of Ge in the BC8, ST12, and R8 structures. In addition, we have performed calculations for the cubic, the hexagonal diamond, and β -Sn phases of Ge to have a complete description of the phases expected to be relevant below pressures of ~ 10 GPa. We carry out detailed calculations on the structural properties, the electronic structure, the equations of state, and give estimates for the transition pressures between phases. Additionally, since except for the frequencies of the zone-center phonons for the BC8 and ST12 phases, little is known about the lattice dynamics of these phases [142]. Hence, we compute the full phonon dispersions for the phases ST12, BC8, R8, and hexagonal diamond.

6.2 Methods

Our calculations of the structural and electronic properties are done within the local density approximation (LDA) to density functional theory (DFT). The plane-wave pseudopotential method [11, 29] is used with a norm-conserving pseudopotential

utilizing a nonlinear-core correction (NLCC) [143] to take into account the spatial overlap between the $3d$ electrons treated as core states and the $4s^2 4p^2$ electrons that were treated as valence. A kinetic energy cutoff of 40 Ry was found to result in highly accurate total energies and phonon frequencies. When structural relaxations were being performed, a higher cutoff of 70 Ry was used in order to have a highly converged stress tensor. This methodology is implemented in the **QUANTUM ESPRESSO** package [46]. Calculations of the phonon frequencies were also done within **QUANTUM ESPRESSO** using the density functional perturbation theory (DFPT) approach [22]. We are very careful to ensure adequate sampling of the electronic and phonon Brillouin zones. Details of these samplings will be given later with the discussion of the corresponding results.

6.3 Results

6.3.1 Structural descriptions

We have relaxed all structures with respect to all degrees of freedom to obtain the relaxed structures at zero pressure. Tests have been performed on the cubic phase of germanium to evaluate the quality of the pseudopotential in providing a good structural description. For cubic Ge a lattice constant of 5.58 Å was found, which underestimates the experimental lattice parameter of 5.66 Å by 1.4% [144]. The calculated bulk modulus is 77.2 GPa, which is in excellent agreement to the value of 75.8 GPa found experimentally [145]. Below we will provide the structural descriptions obtained and used in this work. More extensive descriptions of the space groups, bond length and angle distributions, etc. can be found in the references.

The ST12 structure of Ge can be described as a tetragonal unit cell with a 12 atom basis. The basis is fully specified with four internal parameters x , y , z , and α . We take as our starting point the structural parameters of Mujica and Needs which result from their theoretical relaxation as given in Ref. 138. Structural relaxation from these parameters changes little. We have used a $8 \times 8 \times 8$ grid to sample the Brillouin zone which ensures highly accurate forces and components of the stress tensor. Our final relaxed parameters for the tetragonal lattice are $a=5.823$ Å with $c/a=1.1812$ Å. These compare well with the reported experimental values of 5.93 Å and 1.17707 [99]. Our internal parameters of $x=0.1705$, $y=0.3704$, $z=0.2520$, and $\alpha=0.08679$ compare well with the experimental parameters of $x=0.1730$, $y=0.3784$, $z=0.2486$, and $\alpha=0.0912$ [99]. The bulk modulus of ST12 Ge is calculated to be 65.3 GPa, which is in excellent agreement with the value of 66 GPa calculated by Mujica and Needs [138].

The BC8 structure of Ge is a body-centered cubic lattice with 8 atoms in the unit cell. It is completely specified by its lattice constant a and a single internal parameter x . We begin with the experimental lattice parameter as determined in Ref. 137, in

which $a=6.92$ Å, and take the internal parameter x to be that of the BC8 structure in Si, which experimentally is found to be 0.1033 [99]. We then relax this starting structure with a mesh of size $12 \times 12 \times 12$ in the Brillouin zone. Our relaxed parameters are $a=6.82$ Å and $x=0.1019$, in good agreement with both experiment and previous theory [138]. The calculated bulk modulus is determined to be 73.9 GPa, slightly smaller than the value for the cubic diamond phase slightly higher than the value of 68 GPa calculated in Ref. 138.

As mentioned in the introduction, no definitive experimental report has been made on the existence of the R8 phase in Ge and thus no experimental starting description is available. In this case we use the lattice constants known for Si in the R8 structure and scale them by the ratio of the Ge to the Si lattice constants parameters in the cubic phase. There are 4 internal parameters which specify the atomic positions in the R8 structure: u , x , y , and z . These values were given in Ref. 139 as a result of DFT calculations on the R8 structure in Ge and we take these values as our starting point (the lattice constants were not given in this work). As in the calculation of the BC8 structure, a mesh of $12 \times 12 \times 12$ was used for the integration over the Brillouin zone. Our relaxed lattice parameters for R8 Ge are $a_r=5.909$ Å for the rhombohedral lattice vector with an angle of 109.95° . The obtained internal parameters for the R8 structure are $u=0.28521$, $x=0.47042$, $y=-0.02660$, and $z=0.2771$. The values are similar, although slightly different, from those obtained in Ref. 139. Finally, the calculated bulk modulus of Ge in the R8 structure is 70.5 GPa.

Finally, in this work we also examine the hexagonal diamond structure of Ge. The initial parameters of the hexagonal lattice were taken to be $a=3.96$ Å and $c=6.57$ Å [146] with an internal parameter u taking the “ideal” value of 0.0625 (i.e., $1/16$). Our relaxed parameters of $a=3.93$ Å, $c=6.49$ Å, and $u=0.0631$ are in excellent agreement with these values. The bulk modulus is obtained to be 78.2 GPa, in very close agreement to other LDA results in the literature [88]. The relaxation of this structure was performed with a mesh of $12 \times 12 \times 12$ in the Brillouin zone.

6.3.2 Electronic structure

The band structure of the ST12 Ge is shown in Fig. 6.1. Previous density functional calculations obtain a direct gap of size 0.7 eV about 70% of the way along the Γ -M line [138], whereas those obtained via the empirical pseudopotential method (EPM) obtain a value of 1.47 eV [147]. Due to the well-known band gap problem inherent in DFT, the estimate provided by the EPM calculation is likely closer to the experimental value [148]. Our results are very similar to prior DFT results, although we find a direct gap of 0.56 eV just off the Γ -M line at (0.3333,0.345,0). However, we also find that the fundamental gap is indirect. It is of magnitude 0.54 eV between the valence band maximum at (0.31666,0.3333,0.08333) and the conduction band minimum at the location of the non-fundamental direct gap given above. This difference is simply attributed to the fact that we carried out a very fine sampling

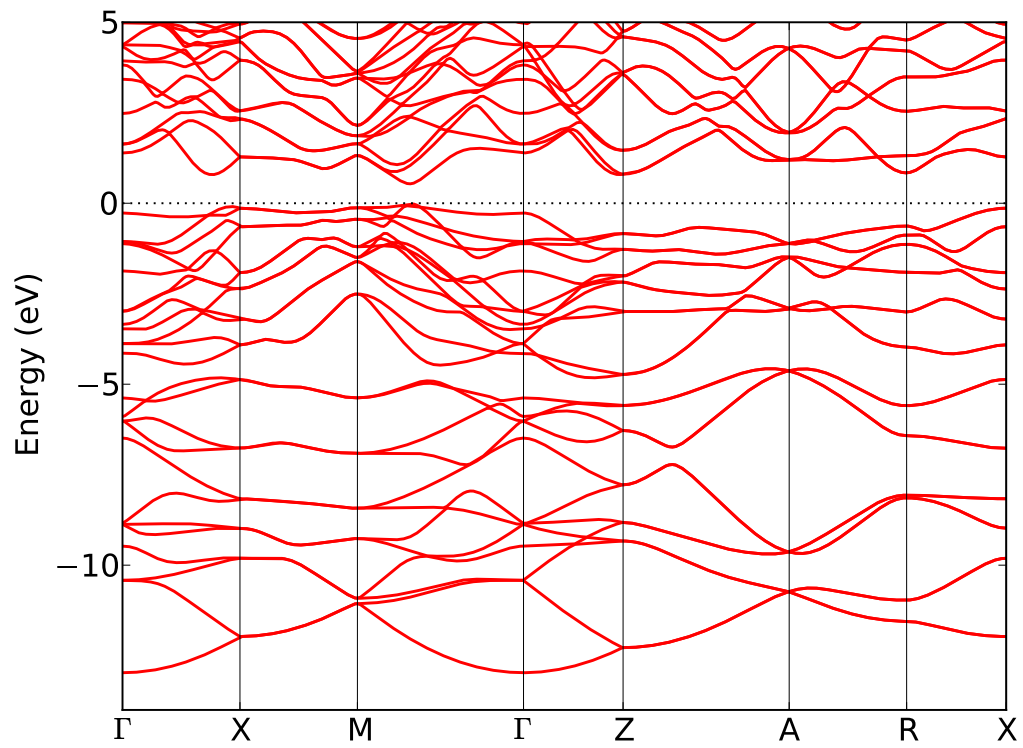


Figure 6.1: Band structure of the ST12 phase of Ge plotted throughout the Brillouin zone. The zero is located at the valence band maximum (VBM) which is just off the Γ -M line (see text).

of the Brillouin zone ($60 \times 60 \times 60$) and therefore could resolve small deviations of the extremum positions off of the high-symmetry lines. Additionally, it should be noted that the valence band manifold in the vicinity of the valence band maximum has little dispersion and so the precise position of this gap could vary slightly depending on the details of the calculation.

The BC8 phase of Ge, like that of Si [76, 105], is found to be semimetallic within DFT. The band structure of this phase is plotted in Fig. 6.2 throughout the high-symmetry points in the equivalent rhombohedral Brillouin zone. There is a direct overlap at the Z point (the H point in the body-centered cubic Brillouin zone). There is also a conduction-band like state which drops below the Fermi level and has a minimum around 50% of the way along the Γ - Z line. This result is in good agreement with the results of Mujica and Needs [138]. The early results of Joannopoulos and Cohen with the EPM method find this conduction band like state to be higher in energy than that of the other conduction-like states at the Z point, which just touch the valence band maximum there and result in a zero-gap material [147].

To our knowledge, the electronic structure of germanium in the R8 structure has not been explored previously. Using the close correspondence to silicon [76], one would guess that R8 Ge would be semimetallic within DFT with a indirect overlap. Our calculations, shown in Fig. 6.3, confirm this expectation. The lowest conduction band occurs at the point L, and overlaps with the highest valence band at the point Z, by 0.57 eV. These are the same points where the band overlap occurs in the R8 phase of silicon. In the case of Si, both GW [105] and hybrid HSE [149] calculations predict that the band overlap present within DFT is lifted to obtain a small band gap semiconductor, and this prediction is consistent with recent experiments [140]. The examination of the electronic structure of R8 Ge at a level of theory beyond DFT is needed to evaluate this possibility, but lies outside the scope of the present work.

The hexagonal diamond phase is a phase which is believed to be typically recovered upon annealing depressurized samples of Si and Ge [89]. The band structure of this material is shown in Fig. 6.4 and we find a small overlap of 0.19 eV at Γ . This is similar to previous DFT theory suggesting a zero band gap at Γ [88] as well as EPM results which predict a direct gap of 0.55 eV at the zone center [147]. Again, because of the band gap underestimation within DFT, the EPM estimate is likely closer to what would be obtained experimentally.

6.3.3 Lattice dynamics

We have calculated the phonon dispersions for the ST12, BC8, R8, and hexagonal diamond phases of Ge throughout the entire Brillouin zone as shown in Figs. 6.5-6.8. The lattice parameters used in the calculations correspond to the zero-pressure theoretical relaxed structures. As a test, we have computed the zone-center optical phonon mode in the cubic phase of Ge and obtained a value of 298.2 cm^{-1} , which is in excellent agreement with the experimental result of 300.6 cm^{-1} [150].

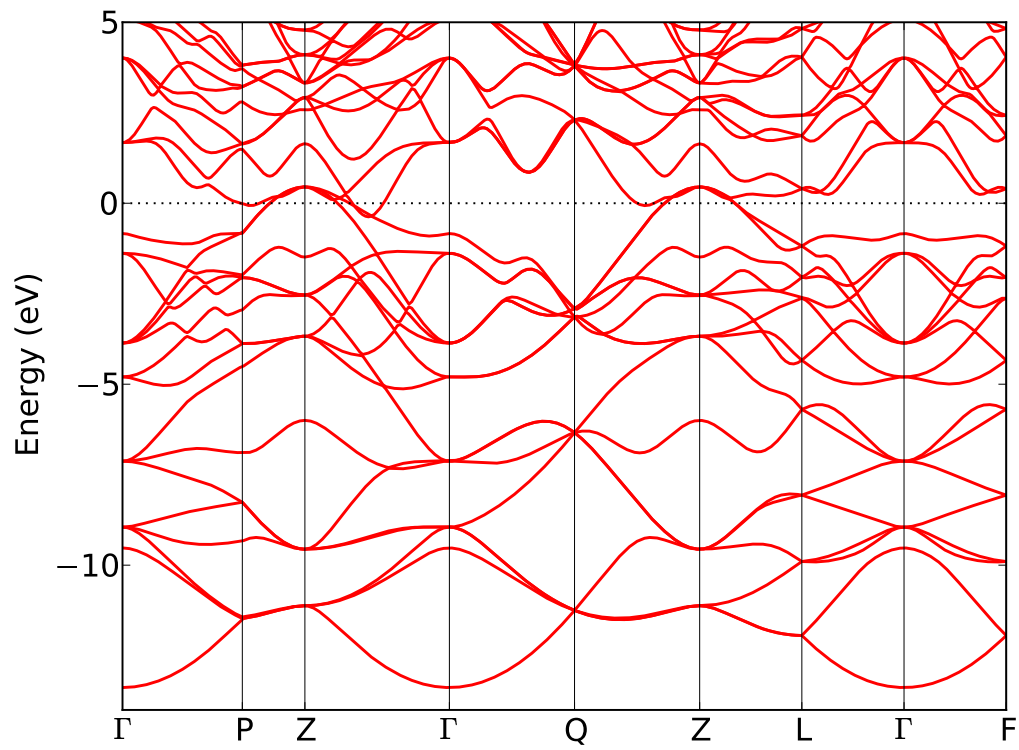


Figure 6.2: Band structure of germanium in the BC8 phase. The path is labeled according to the equivalent rhombohedral description. The Fermi level is located at zero energy.

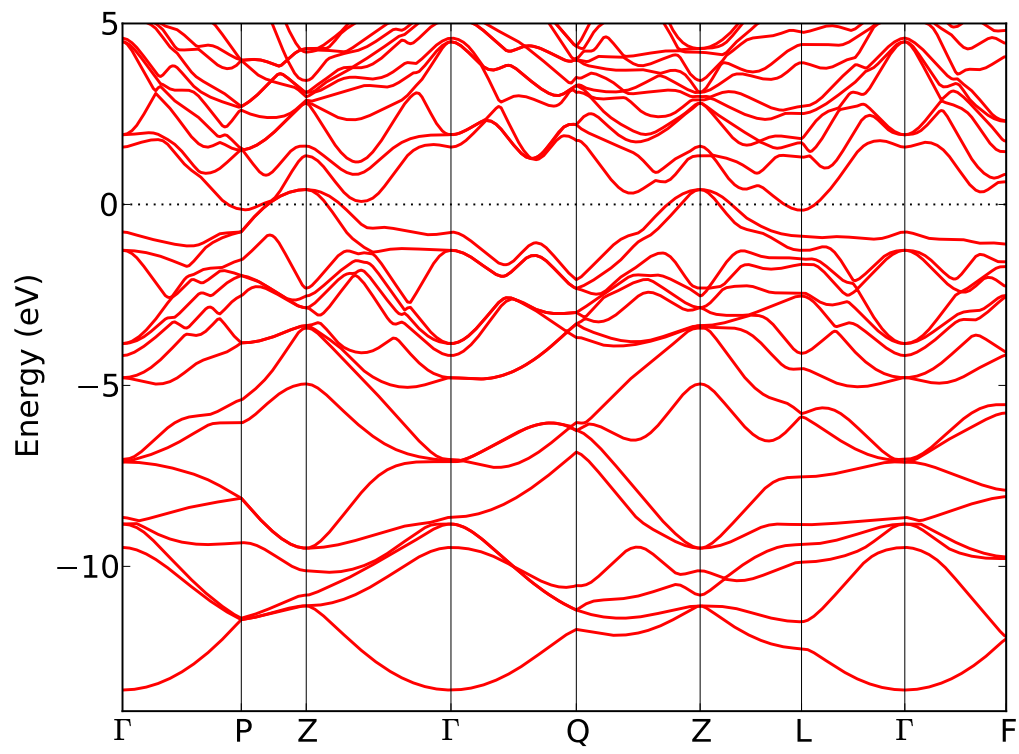


Figure 6.3: Band structure of germanium in the R8 phase. The Fermi level is located at zero energy.

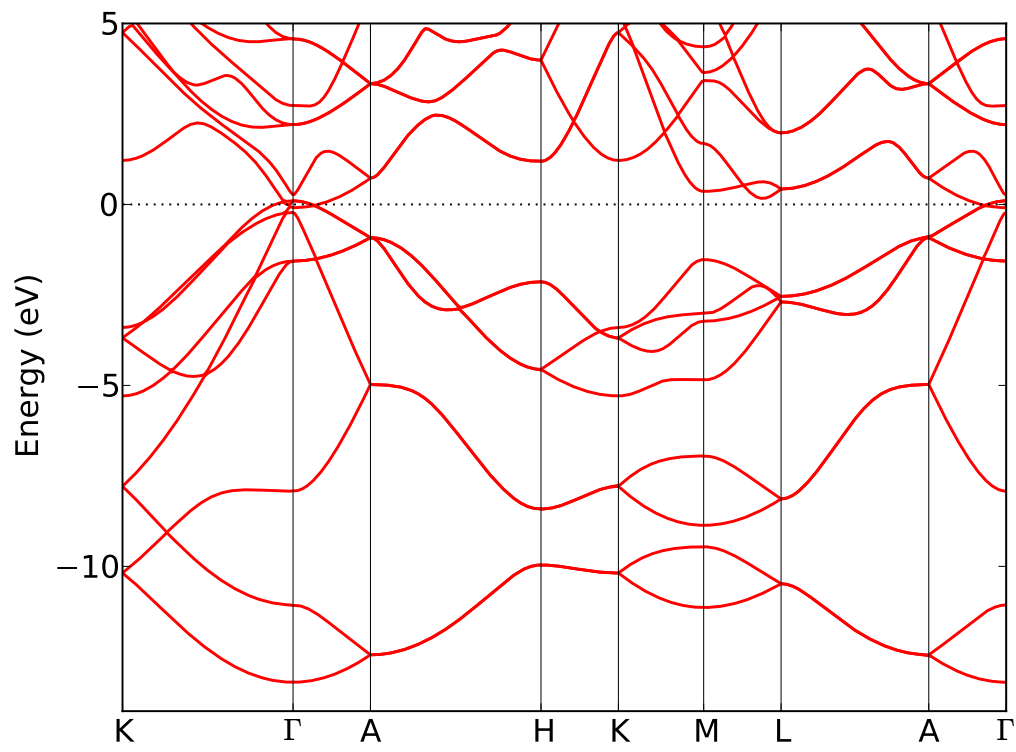


Figure 6.4: Band structure of germanium in the hexagonal diamond (lonsdaleite) phase. There is a small direct overlap of 0.19 eV at Γ .

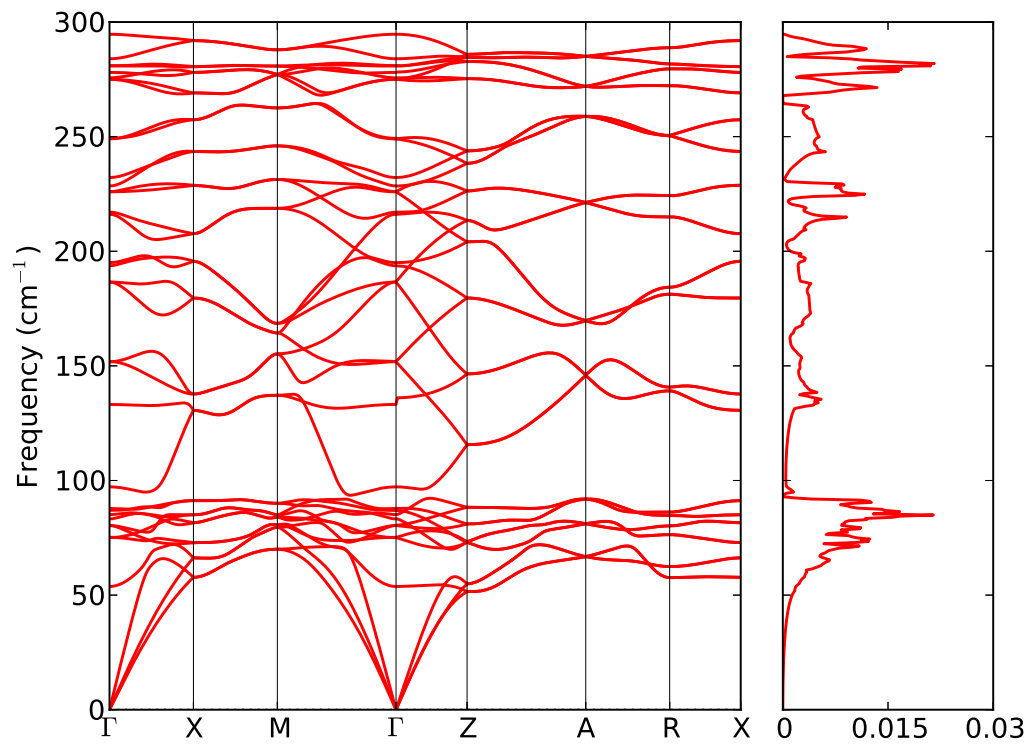


Figure 6.5: Phonon dispersion of the ST12 phase of germanium. The phonon DOS is normalized to 1.

The phonon dispersion of ST12 Ge is shown in Fig. 6.5. The full phonon dispersion was obtained by Fourier interpolation of a $4 \times 4 \times 4$ uniform grid of q-points in the Brillouin zone. The electronic sampling in these calculations was on a shifted $5 \times 5 \times 5$ grid. The phonon dispersion of ST12 exhibits relatively flat bands in the region around 85 cm^{-1} and 280 cm^{-1} leading to peaks in the phonon DOS in those regions. An interesting feature is the presence of a slight discontinuity of approximately 3 cm^{-1} in the phonon frequency near Γ at $\sim 135 \text{ cm}^{-1}$ as one goes from the $M \rightarrow \Gamma \rightarrow Z$ direction. The source of this discontinuity is the nonanalytic part of the force constant matrix which is present due to macroscopic polarization associated with this phonon mode. Typically one does not expect such contributions related to polar materials in a phase composed of only one element, but here it occurs due to the specific structure present in ST12. Phonon calculations we have performed on silicon in the ST12 structure show a similar discontinuity at this point. The calculated Raman active modes at Γ are (all frequencies in cm^{-1}): $54(B_1)$, $75(E)$, $80(E)$, $85(A_1)$, $87(B_1)$, $88(B_2)$, $97(A_1)$, $152(E)$, $187(E)$, $194(B_2)$, $195(A_1)$, $216(B_1)$, $217(B_2)$, $226(E)$, $232(B_1)$, $249(E)$, $275(E)$, $276(A_1)$, $281(E)$, $284(B_1)$, and $295(B_2)$.

The phonon dispersion for BC8 was obtained throughout the zone via Fourier interpolation from a uniform grid of $6 \times 6 \times 6$ phonon wavevectors in the Brillouin zone and a $10 \times 10 \times 10$ shifted grid for the sampling of the electronic states and is displayed in Fig. 6.6. The Raman-active modes in the BC8 structure are obtained as (all frequencies in cm^{-1}): $87(T_g)$, $214(T_g)$, $227(A_g)$, $246(T_g)$, and $259(E_g)$.

The phonon dispersion for the R8 phase of Ge is shown in Fig. 6.7. The electronic integration mesh, like in BC8, was performed on a $10 \times 10 \times 10$ shifted grid and a uniform grid of $6 \times 6 \times 6$ phonons were calculated to obtain the interpolated dispersion. Qualitatively, the phonon spectra for the R8 phase is similar to that found in the BC8 phase, although the differences present manifest themselves clearly in the calculated phonon DOS where the difference in the peak structure can be seen. The zone-center Raman active modes are found to be (all frequencies in cm^{-1}): $83(A_g)$, $90(E_g)$, $203(A_g)$, $212(E_g)$, $223(A_g)$, $244(A_g)$, $247(E_g)$, and $278(E_g)$.

Finally, the phonon dispersion of the hexagonal diamond (lonsdaleite) phase of Ge is shown in Fig. 6.8. These calculations were performed for a uniform grid of $6 \times 6 \times 6$ phonons with an electronic sampling on an $10 \times 10 \times 10$ shifted grid. The phonon DOS exhibits a prominent peak near a frequency of 275 cm^{-1} . The Raman-active modes for this phase are determined to be (all frequencies in cm^{-1}): $287(E_{2g})$, $301(E_{1g})$, and $302(A_{1g})$.

6.3.4 Equation of state

In order to calculate coexistence pressures that may aid future experimental investigations, we have calculated the energy-volume relations for the cubic, BC8, R8, ST12, hexagonal diamond, and β -Sn structures of Ge. These calculations are performed by calculating the total energies of each structure while fully relaxing the

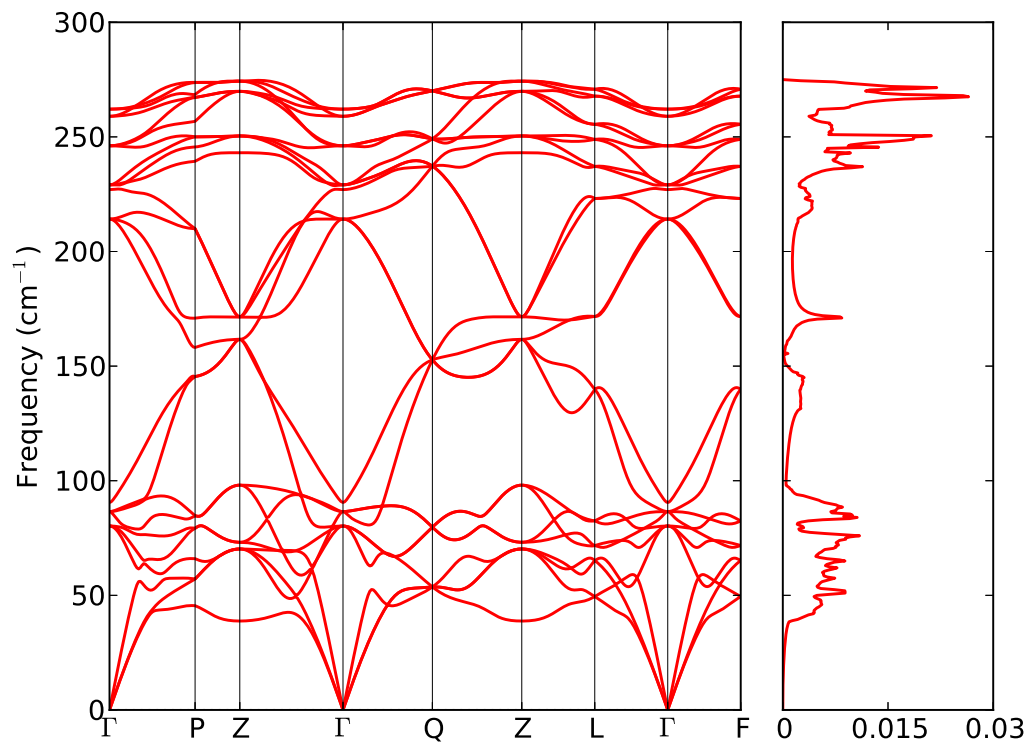


Figure 6.6: Phonon dispersion of the BC8 phase of Ge, plotted along the directions in the equivalent rhombohedral Brillouin zone. The phonon DOS is normalized as in Fig. 6.5.

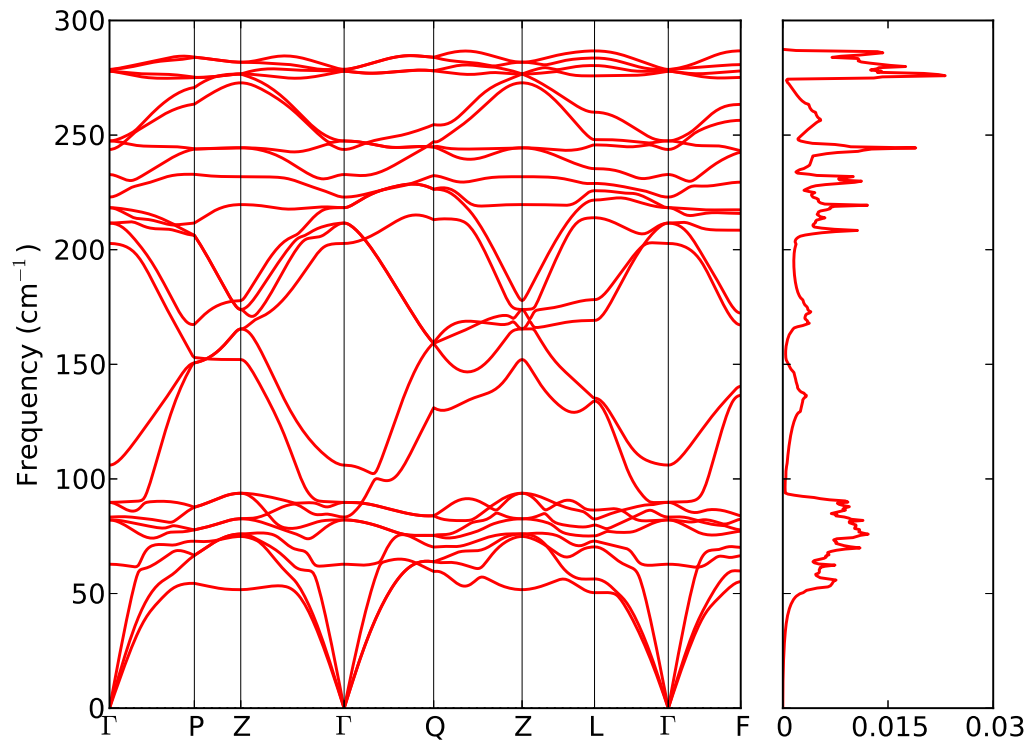


Figure 6.7: Phonon dispersion of the R8 phase of germanium. The phonon DOS is normalized as in Fig. 6.5.

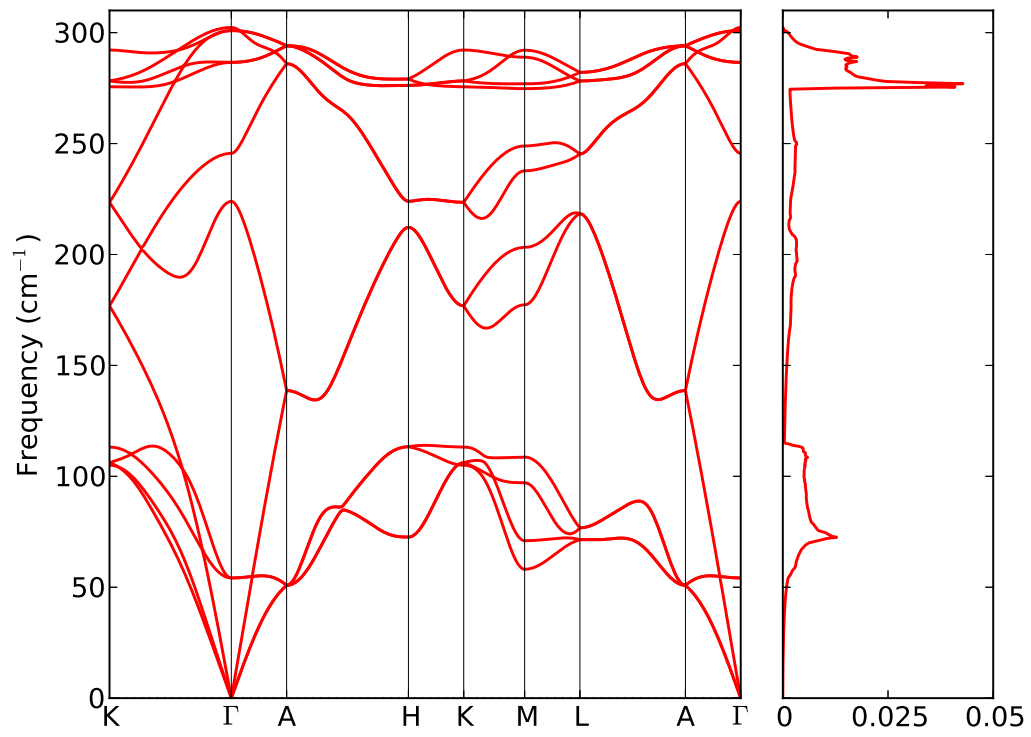


Figure 6.8: Phonon dispersion of the hexagonal diamond phase of germanium. The phonon DOS is normalized as in Fig. 6.5.

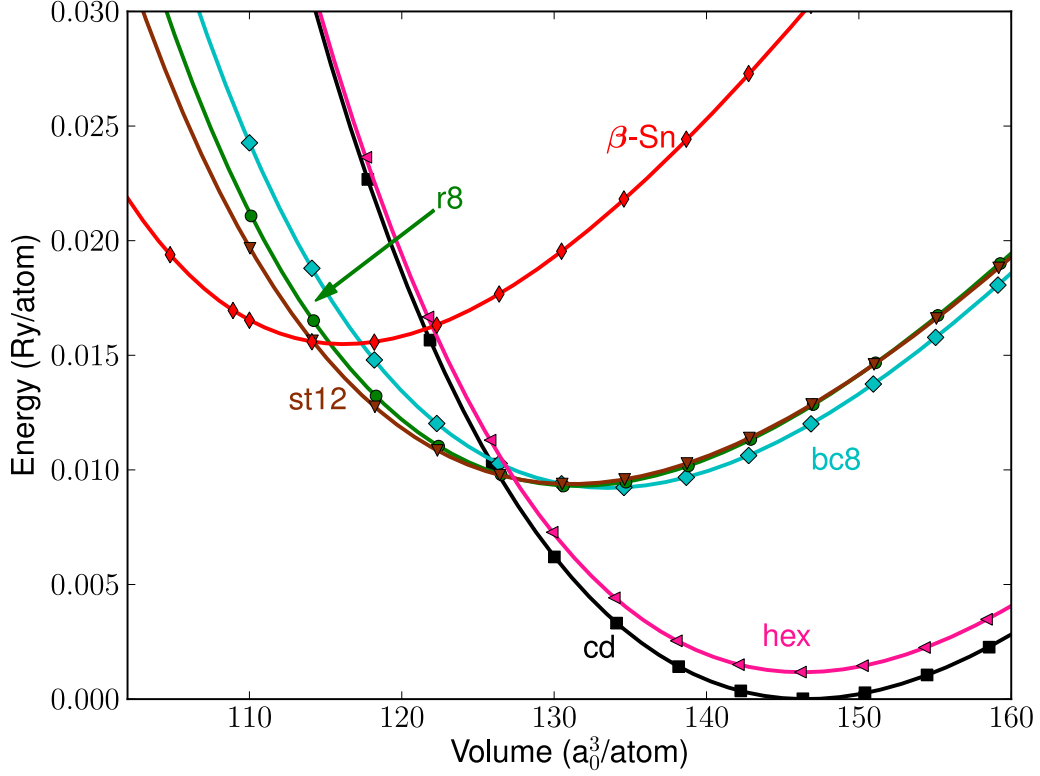


Figure 6.9: Energy-volume relations for low-pressure polymorphs of germanium. The markers correspond to the calculated data and the lines are fits to the Birch-Murnaghan equation of state. The equilibrium energy of the cubic phase is located at zero energy.

cell and ion positions at fixed volume. The resulting $E(V)$ curve is fit with a Birch-Murnaghan equation of state [151]. The results are shown in Fig. 6.9. Pressure-volume curves are shown in Fig. 6.10 and are obtained via numerical differentiation of the $E(V)$ relations. Knowledge of $E(V)$ and $V(P)$ allows for the determination of the enthalpy as a function of pressure, $H(P) = E(V(P)) + PV(P)$, which is shown in Fig. 6.11. It is simpler to obtain coexistence pressures from the enthalpy-pressure curves than from the equivalent “common tangent” approach to the energy-volume relations.

From the results of Fig. 6.11, we can calculate transition pressures between the various phases. The first of these we will examine is that of the cubic $\rightarrow \beta$ -Sn transition that occurs experimentally at 10.5 GPa [152]. We find this transition at a pressure of 7.9 GPa, which underestimates the transition in comparison to the experimental value. This result is in reasonable agreement with that obtained in previous

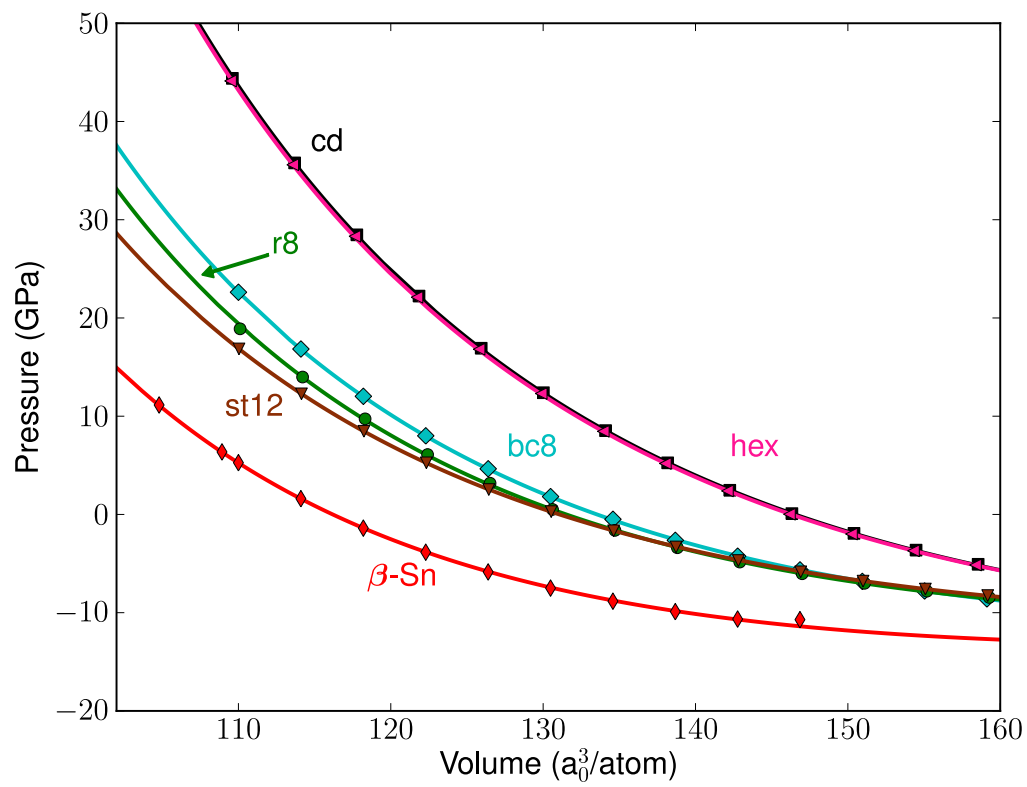


Figure 6.10: Pressure-volume relations for low-pressure polymorphs of germanium. The markers correspond to the calculated data and the lines are taken from numerical differentiation from the $E(V)$ curves shown in Fig. 6.9.

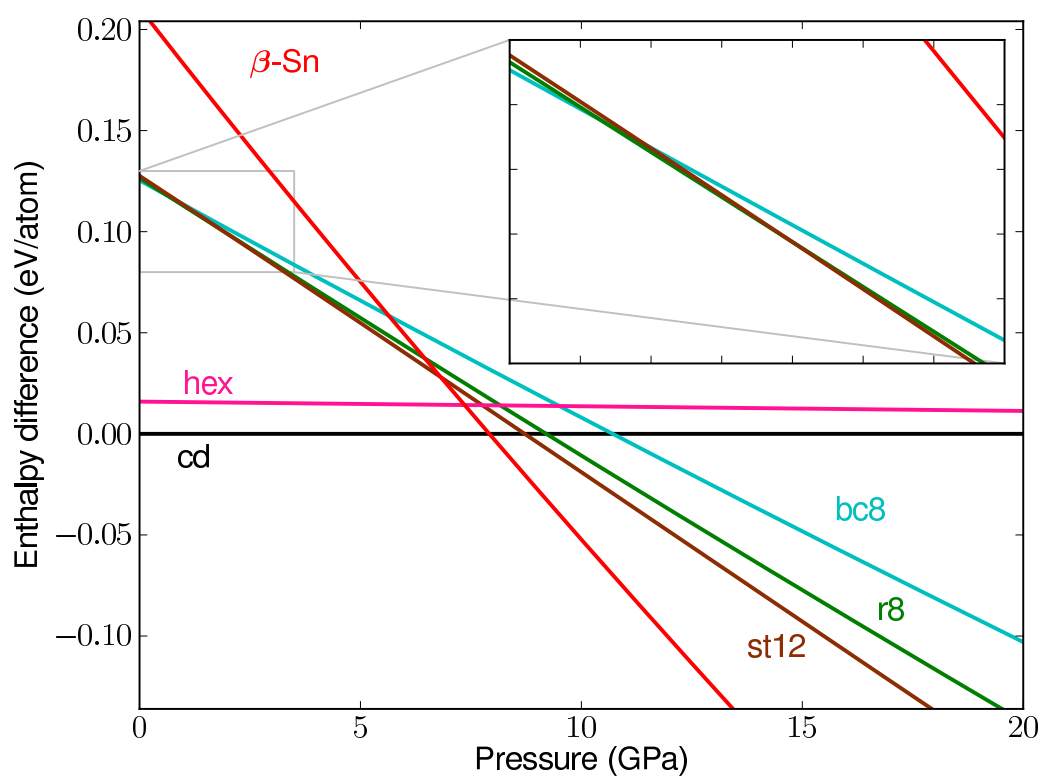


Figure 6.11: Enthalpy as a function of pressure for low-pressure polymorphs of germanium. All enthalpies are taken relative to that of the cubic phase, which is shown at zero. The inset is a zoomed-in view showing the BC8-R8-ST12 transition region.

theoretical calculations by Mujica *et al.* within the LDA, who also underestimate the experimental transition pressure, predicting the transition at 6 GPa [138]. The early LDA results of Yin and Cohen [153] also underestimate the transition pressure at a value of 9.6 GPa, although in this work the coexistence pressure is higher and closer to the experimental value partly because of the use of the Wigner formula [154] for the exchange-correlation energy [155]. The underestimation of the cubic $\rightarrow \beta$ -Sn transition pressure within LDA functionals occurs in both Si and Ge and is improved upon the inclusion of gradient corrections to the exchange correlation functional which increase the transition pressure and bring it closer to the experimental value [155, 156]. This behavior was explained as being the result of a greater lowering of the total energy for systems with more inhomogeneous charge densities upon going from the LDA to the GGA [155]. Since the cubic diamond phase has a more inhomogeneous charge density than that of the metallic β -Sn phase, its total energy is lowered by a greater amount upon going to the GGA. This results in a larger difference in energies at equilibrium between the two phases and thus a larger transition pressure [155].

An interesting observation, originally seen in the results of Mujica *et al* [138] is that if the cubic $\rightarrow \beta$ -Sn transition pressure were corrected to the experimental value of 10.5 GPa, that this would result in a region of stability for the ST12 phase because the enthalpy for the cubic diamond phase becomes equal to that of ST12 at 8.7 GPa. This is just under 2 GPa from the experimental cubic- β -Sn coexistence pressure. Although there have been reports recently in nanocrystalline samples of Ge that the ST12 phase appears to be more stable than previous experiments have suggested [157], we are not aware of any experimental results that have reported this possible region of stability at higher pressures.

With regards to the phases obtained upon decompression from β -Sn, namely the R8, BC8, and ST12 phases of Ge, we obtain regions of metastability (compared to the cubic phase) for all three phases. This can be seen from the inset of Fig. 6.11. The ST12 phase of Ge is found to transition to the R8 structure at 2.0 GPa, which remains the lowest enthalpy of these three phases until 0.65 GPa when it transitions to the BC8 phase.

Predictions of phase stability are known to be sensitive, in some cases strongly so, to the underlying treatment of exchange and correlation effects. This was discussed above in relation to the cubic $\rightarrow \beta$ -Sn transition. The small enthalpy differences between the BC8, R8, and ST12 phases near the predicted transitions as seen in Fig. 6.11 suggest that these transitions might also be sensitive to the functional used. In order to illustrate this, we have repeated the equation of state calculations with another pseudopotential which includes both the semicore $3d$ states explicitly and in addition treats the correlation-exchange energy in the PW91 GGA functional of Perdew and Wang [25, 158]. The purpose of this is not to systematically evaluate the effects of gradient corrections, the inclusion of semicore states, etc., but to simply obtain the results that would be expected if different choices were made in the pseudopotential construction. As expected with the GGA, the lattice constant

of cubic germanium is overestimated, in this case by 1.7%, in comparison to experiment. The bulk modulus obtained is 60.2 GPa, which compares less favorably to the experimental value of 75 GPa that we obtained with our calculations within the LDA. This underestimation of the bulk modulus in Ge within GGA is consistent with results elsewhere in the literature [88, 155]. With the GGA functional, the transition pressure of the cubic $\rightarrow \beta$ -Sn transition is pushed up to 10 GPa, in close agreement with the experimental value of 10.5 GPa, as expected due to the effect of the GGA on systems with large differences in their degree of charge homogeneity. In the GGA calculations the pressure region where R8 is favored over BC8 or ST12 is no longer present, and BC8 is found to have a coexistence pressure with ST12 at 3.6 GPa. The enthalpies of the ST12 and β -Sn phases become equal at a pressure of 8.4 GPa, which differs from the 6.8 GPa we find in the calculation within the LDA.

The fact that changing the exchange-correlation functional and altering the split between core and valence electrons changes the prediction of a metastable R8 phase in Ge means that we cannot draw firm conclusions that this is the case. This is not surprising because of the closeness in enthalpy of the three enthalpy curves in both of our calculations. However, it must be pointed out that GGA calculations do not uniformly improve the results in relation to the LDA [159]. The strong effect of simply including gradient corrections on the cubic- β -Sn transition due to the differences in charge homogeneity is not expected to exist in relation to the transitions between BC8, R8, and ST12, which all have similar bonding characteristics. Furthermore, the calculations performed with the LDA result in a much more accurate bulk modulus than that in the GGA, which has a direct and obvious relationship to the $E(V)$ relations and thus the calculated transition pressures. Finally, it is well-known that DFT has shortcomings in its treatment of localized states. In germanium this results in binding energies which are too shallow in relation to experiment by about 5 eV for the semicore $3d$ states [160], and thus their inclusion into the valence might also induce some errors in the computed quantities.

6.4 Comparison to experiment

In section 6.3.3 we reported the Raman-active modes for the low pressure phases of Ge studied in this chapter. These theoretical predictions have recently been used to provide support for the existence of the R8 phase of Ge in indented samples of amorphous Ge [161]. In Raman spectra taken on an indented sample of amorphous Ge, a prominent line is found to occur at 202.7 cm^{-1} . This mode is in very good agreement with the Raman active mode A_g mode predicted above at 203 cm^{-1} . Other modes characteristic of R8 are also present in the sample. For example, the theoretically predicted E_g mode at 212 cm^{-1} is well produced by a peak in the experimental Raman spectrum at 213.2 cm^{-1} .

However, there has been previous experimental work done on amorphous Ge in

diamond anvil cells (DAC) by Coppari *et al.*[142]. This work obtains a very similar Raman spectra as that of Ref. 161. In the work of Ref. 142, however, they attribute their results to that of the ST12 phase, which is known from previous studies on crystalline germanium to result in slow compression from the β -Sn phase [89]. This determination is made in Ref. 142 by comparing to first principles calculated phonon frequencies and their dependence on pressure. Yet, an examination of the comparison between the theoretical Raman frequencies and those determined experimentally in the work of Coppari *et al.* shows that the agreement between the two is actually quite poor. In the study of Ref. 142 only the ST12 and BC8 phases of Ge were considered. However, we have found that the phonon frequencies corresponding to the R8 phase of Ge, along with their pressure evolution, corresponds exceptionally well to the results in Ref. 142. This data is shown in Fig. 6.12, in which the experimental pressure evolution of 5 Raman-active modes is shown. The bottom two modes match exceptionally well with the theoretically predicted modes of R8, both in absolute value and in their trend of decreasing energy separation as pressure is increased. The ST12 modes in the vicinity, in addition to being either higher or lower in energy, also display the incorrect behavior as pressure is increased: the energy difference between the modes increases. The third and fourth highest energy modes seen experimentally could in principle be attributed to either the R8 or the ST12 phases, as the theoretical curves are very similar in those regions and thus a clear determination is not possible. However, when one looks at the pressure dependence of the highest energy experimental mode, it is apparent that its rapid rise in frequency as pressure is increased can only be explained as being either a mode corresponding to BC8 or one corresponding to R8. All theoretically calculated ST12 modes in that frequency range have much smaller pressure coefficients than the experimental mode exhibits, and thus it is highly unlikely that this mode corresponds to the ST12 phase.

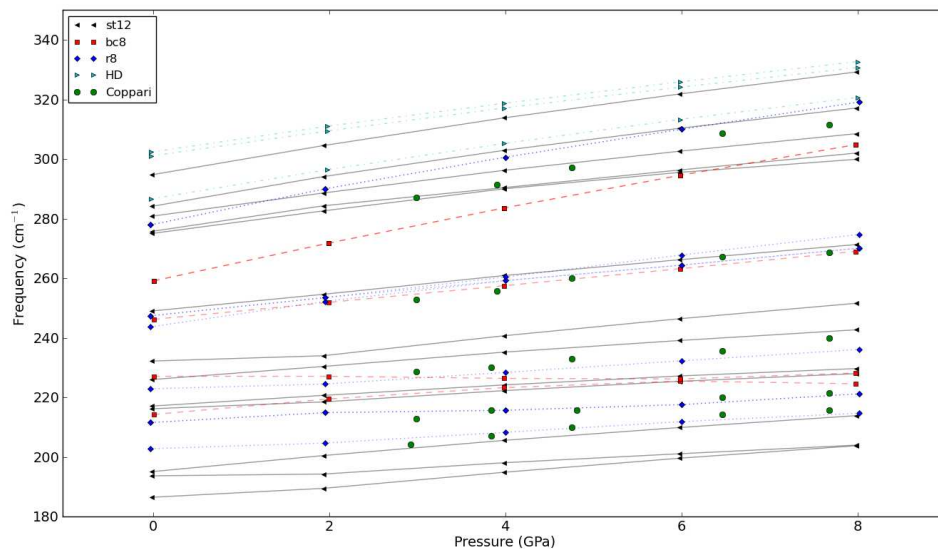


Figure 6.12: Calculated pressure dependence Raman-active modes pertaining to the Ge phases studied in this chapter along with the experimental data of Coppari *et al.* (Ref. 142).

6.5 Conclusion

In this chapter we have calculated the electronic structure, phonon dispersions, and the phase diagram of the low pressure phases of germanium, namely those of hexagonal diamond, ST12, R8, and BC8. Our electronic structures are in good agreement with previous calculations and suggest BC8 and R8 to be semimetallic, at least within DFT, while ST12 is semiconducting. Calculations of the lattice dynamics of these phases have been performed and Raman-active phonon frequencies have been identified. These frequencies were then used to support experimental evidence of the R8 phase of Ge occurring in indented samples of amorphous Ge [161]. Furthermore, prior experimental work which identified the obtained phase in DAC experiments as ST12 was likely mistaken, and it is more likely that they were actually seeing signatures of the R8 phase.

Calculations of the equation of state of these phases find regions of metastability of the BC8, R8, and ST12 phases of germanium which suggests that the R8 phase of Ge might be obtained from compression of the BC8 phase or a partial reduction of pressure upon obtaining ST12. However, because of the small energy differences involved, this result is sensitive to choices made in constructing the pseudopotential, and thus firm predictions on the metastability of the R8 structure in Ge are difficult to make at this time. Further experimental studies could help shed light on this

matter, and the Raman-active phonon modes calculated here will hopefully be of use in helping to identify these phases in those experiments.

Chapter 7

Prediction of a new metastable phase of silicon in the *Ibam* structure

7.1 Introduction

The high-pressure phase diagram of silicon has been the focus of numerous experimental and theoretical studies because of its extreme technological importance. A major goal is to find new structural phases with desirable electronic properties such as direct band gaps and useful optical spectra for applications. High-pressure phases of Si have been studied experimentally up to at least 248 GPa [162] and the observed phase transitions have in general been in good agreement with theoretical predictions [28, 78, 79, 82, 86, 134, 163–166]. It is also known that upon slow pressure release from the metallic β -Sn phase that Si does not return to the cubic diamond phase but instead transforms into a number of metastable phases with distorted tetrahedral bonding, namely the r8 and bc8 phases of Si [76, 79]. However, other phases have been reportedly obtained with different conditions of pressure release. Upon rapid release of pressure (<100 ms) two other phases of silicon were obtained, referred to as Si-VIII and Si-IX, however only incomplete structural information could be obtained from the x-ray diffraction pattern [167] and hence electronic structure calculations are not possible. Furthermore, nanoindentation experiments in silicon have resulted in the observation of another phase, Si-XIII, although little is known regarding its structure [168, 169].

The lack of information about the low-pressure phase diagram of Si is unfortunate because while cubic silicon has many characteristics which make it favorable for modern technological applications, it should be possible to improve on some of these by using a closely related silicon polytype. Because it is possible that one of the phases previously obtained but unidentified may have desirable electronic, optical, or other

properties that would be an improvement on existing silicon technology based on cubic silicon, several unidentified phases are being studied. This idea is behind recent theoretical studies of the optical spectra in both known metastable phases of Si as well as doped cubic silicon in the hope of obtaining better absorption characteristics relevant to photovoltaic devices [122, 170, 171].

7.2 AIRSS method details

In this chapter we use *ab initio* random structure searching (AIRSS) [172] to further explore the phase diagram of Si. In particular we target the phase Si-IX, which is one of the two phases obtained from rapid depressurization from 12 GPa in diamond anvil cell experiments [167]. The experimental x-ray diffraction lines could be indexed on the basis of a tetragonal unit cell with cell parameters $a = 7.482 \text{ \AA}$ and $c = 3.856 \text{ \AA}$. By assuming the same density as nearby metastable phases it was estimated that the unit cell contains 12 atoms, although the atomic positions were undetermined. It was noted that this phase is not equivalent to the known, metastable silicon phase st12.

We conduct our structure search by randomly generating 5000 structures with 12 atoms in the unit cell with the experimentally determined cell parameters. The atomic positions are then optimized to their lowest energy configuration using the local density approximation (LDA) to density functional theory (DFT) as implemented in VASP using the projector augmented wave method [173–175]. We employ a planewave cutoff energy of 350 eV and a Monkhorst-Pack Brillouin zone sampling grid with a resolution of $2\pi \times 0.04 \text{ \AA}^{-1}$. In order to evaluate the dynamical stability of the resulting structures we calculate phonon frequencies throughout the Brillouin zone using density-functional perturbation theory (DFPT) as implemented in the Quantum-ESPRESSO package [46]. Thermodynamic stability and the equation of states for the various polytypes of Si are also computed within Quantum-ESPRESSO using the planewave pseudopotential method [11, 29] with norm-conserving pseudopotentials and a kinetic energy cutoff of 40 Ry.

7.3 Results

The random structure search results in a single lowest energy structure [176]. Symmetry analysis reveals that the structure corresponds to the orthorhombic space group *Ibam* with atoms in the 8(j) and 4(a) Wyckoff positions. However, calculation of the stress tensor for this structure reveals that there is a slight nonuniform pressure which distorts the cell parameters from a tetragonal to an orthorhombic cell. The fully relaxed structure at 0 GPa retains the *Ibam* symmetry with $a = 3.872 \text{ \AA}$, $b = 7.887 \text{ \AA}$, and $c = 6.877 \text{ \AA}$. The atomic positions are located at the 8(j) sites with $x = 0.25201$

and $y=0.16322$ and the 4(a) sites with $z=0.25$. This structure, shown in Fig. 7.1, has distorted tetrahedral bonding with the interesting feature of having four-fold rings. These four-fold rings are connected in a chain-like manner, similar to that found in SiS_2 and SiSe_2 which also take the *Ibam* structure [177]. The presence of only even-membered rings opens up the possibility of this structure forming in III-V semiconductors because of the absence of energetically unfavorable same-species bonds.

The bond lengths in the *Ibam* structure are similar to those found in the cubic phase of silicon. Around the 4(a) sites there is a single bond length of 2.359 Å, with three distinct angles of 86.4°, 113.9°, and 131.1°. The 8(j) sites have two distinct bond lengths of 2.359 Å bridging to the 4(a) sites and 2.371 Å between neighboring 8(j) sites. Around the 8(j) sites there are four distinct bond angles of 93.6°, 130.7°, 109.5°, and 88.7°. As already mentioned, the fully relaxed *Ibam* structure is similar although not identical to the structural parameters experimentally determined by Zhao *et al.* [167] Relaxation of the cell stress alters the lattice constants by 0.4%, 5.4%, and -8.0%, although there is no change in the symmetry type or nature of the bonding. Simulations of the x-ray diffraction spectra for both the structure with a fixed cell and a relaxed cell show some agreement with the diffraction peaks found in experiment, however the agreement is insufficient to unambiguously identify Si-IX as being in the *Ibam* structure. Phonon frequencies are also calculated for the *Ibam* structure on a $6 \times 3 \times 3$ grid of q -vectors in the Brillouin zone and all frequencies were found to be real, indicating that the structure is locally stable. The resulting phonon dispersion plotted throughout the Brillouin zone is shown in Fig. 7.2.

The bandstructure of *Ibam* Si is shown in Fig. 7.3 plotted throughout the orthorhombic Brillouin zone. *Ibam* Si is predicted within the LDA to be semimetallic as can be seen in Fig. 7.3 and from the small but finite density of states present at the Fermi level in Fig. 7.4. There is a slight direct overlap at Γ of approximately 0.275 eV, although the (occupied) band minimum of the would-be conduction manifold is located off the high-symmetry lines at the point $P=(0.4,0,0.35)$ in reciprocal coordinates. This band minimum is located 0.35 eV below what would be the valence band maximum at Γ in the case of a band gap opening. Calculations based on LDA-DFT approaches are known to inadequately predict quasiparticle energies. However, generally agreement with experiment can be obtained by including quasiparticle self-energy corrections as calculated within the GW approximation [42]. Although such calculations are beyond the scope of the present study, these corrections are of sufficient magnitude in other phases of Si [105] that it is expected that *Ibam* Si would become semiconducting. It is also worth noting that the Si-IX structure obtained by Zhao *et al.* was believed to be either semimetallic or a narrow-band semiconductor [167]. Finally, the sudden large increase in the density of states near the Fermi level as shown in Fig. 7.4 and the many-valley nature of the bandstructure as seen in Fig. 7.3 suggests that *Ibam* Si could be of interest as a superconductor if it were doped. Many-valley phonon scattering processes can give rise to superconductivity

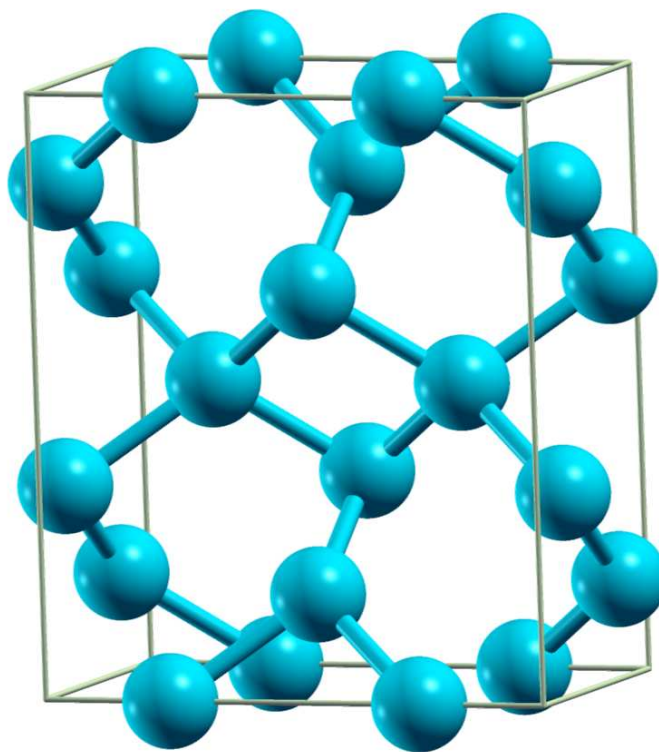


Figure 7.1: Structure of *Ibam* Si.

as a result of the large number of states available for scattering in addition to the fact that intervalley processes involve large momentum transfers and thus tend to be screened less than intravalley processes [178].

In order to evaluate the thermodynamic stability of the *Ibam* phase, we have performed total energy calculations on all experimentally observed silicon phases and have also included the predicted metastable phases st12 and the recently predicted bct phase [120]. For each phase, total energies are calculated while fully relaxing the cell under a fixed volume, and the resulting $E(V)$ data is fit with a Birch-Murnaghan equation of state [151]. The resulting fits are shown in Fig 7.5. Pressure-volume curves were obtained by numerical differentiation of the total-energy curves (as shown in Fig. 7.6) and were used to produce the enthalpy versus pressure curves shown in Fig. 7.7.

For the range of pressures studied, the *Ibam* phase of silicon is never the lowest enthalpy structure and so is always a metastable phase. This does not preclude the possible observation of the *Ibam* structure however, as the conditions present in experiment and the activation barriers between phases play an important role. Although the r8 and bc8 phases also have higher enthalpy than cubic silicon for all pressures, they are present in experiments resulting from decompression from the β -Sn phase. There is also the possibility that germanium could form in the *Ibam*

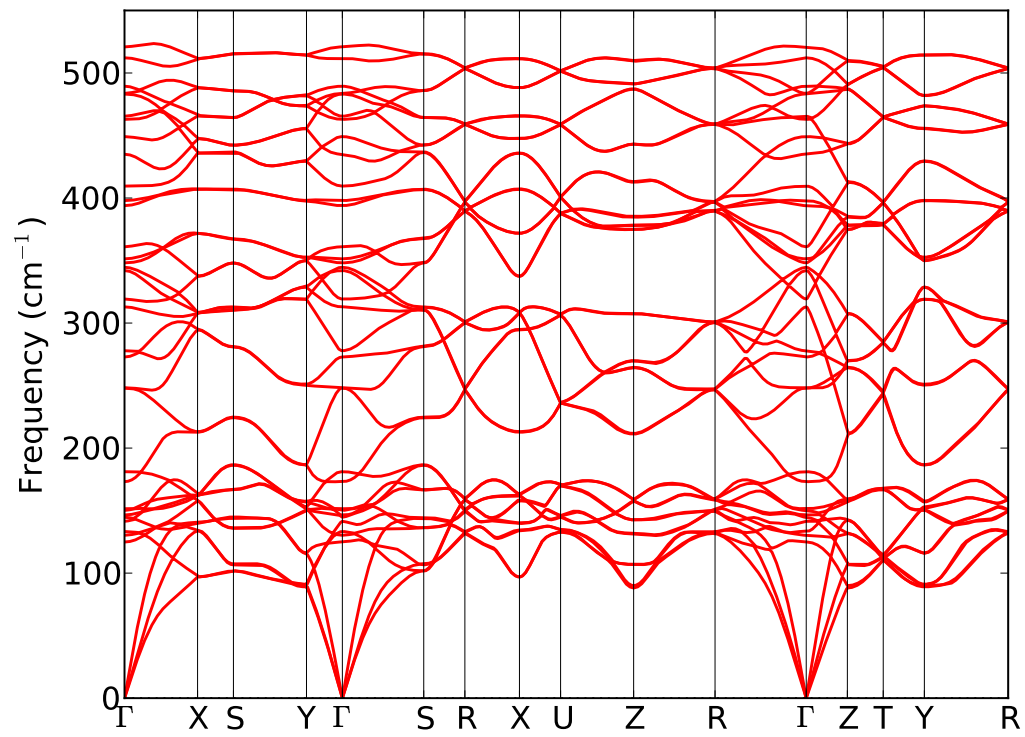


Figure 7.2: Phonon dispersion of *Ibam* Si plotted along the high-symmetry lines of the Brillouin zone

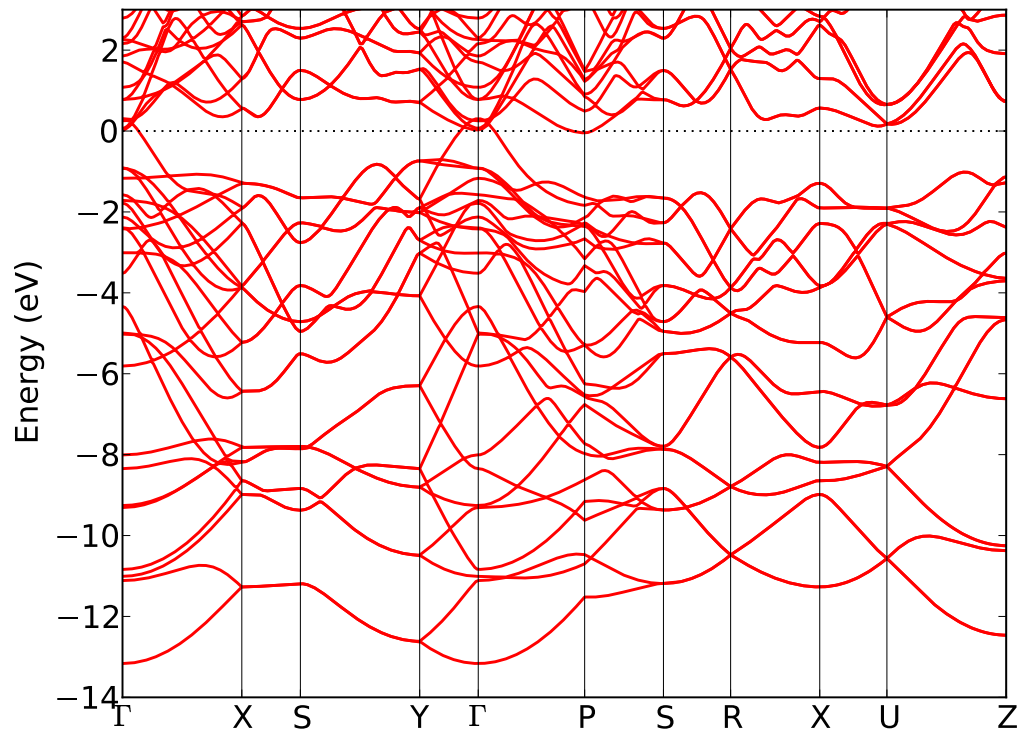


Figure 7.3: Bandstructure of *Ibam* Si. The Fermi level is at zero energy.

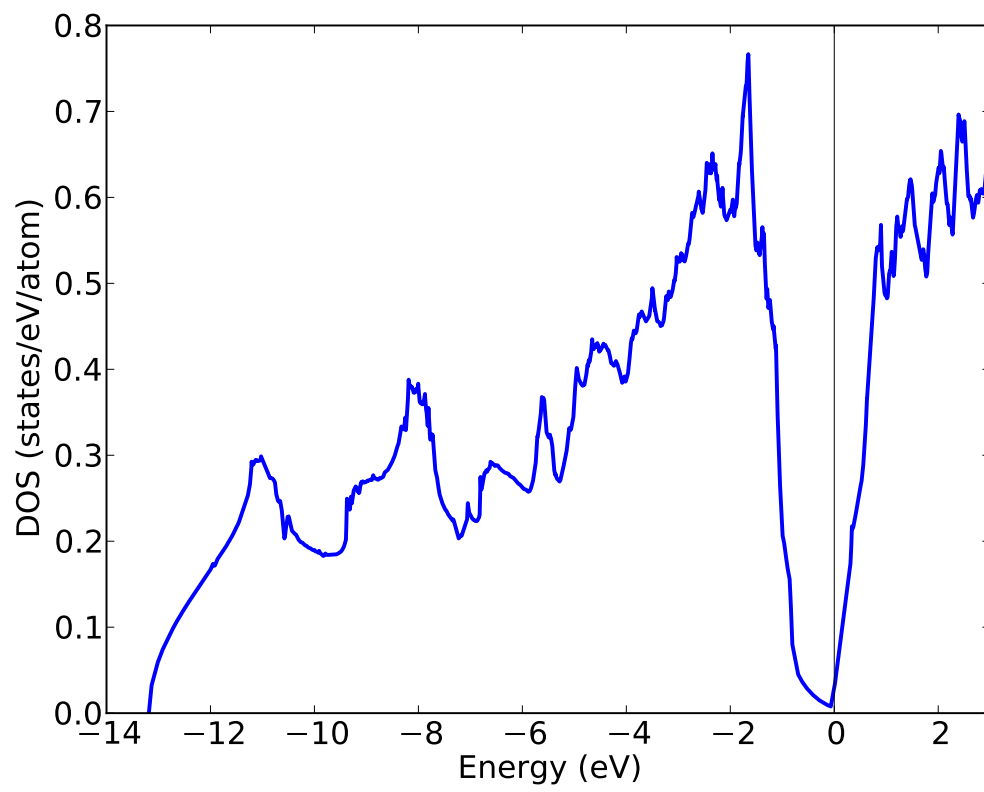


Figure 7.4: Electronic density of states (DOS) of Si in the *Ibam* structure. The Fermi level is at zero energy.

structure because of the similar, but not identical, correspondence between the phase diagram of Si and Ge. For example, the bc8 and r8 phases are favored over st12 in Si, but in Ge the reverse is true and st12 is obtained at zero pressure.

From our calculations we can also compare the calculated coexistence pressures for the other known phase transitions in Si. The cubic silicon to β -Sn transition occurs within a pressure range of 8.8-12.5 GPa [78]. Our calculations predict this transition to occur at a pressure of 9.2 GPa in good agreement with the experimental data and in better agreement than previous theoretical results which lie between 7.8-8.8 GPa [78]. Experimentally, upon increasing the pressure from β -Sn, Si takes on the *Imma* structure at 13.2 GPa [82]. Our calculations predict that the cubic phase transforms to *Imma* at a pressure of 9.1 GPa, lower than that of β -Sn transition. This result is in agreement with the results of Lewis and Cohen [164] who find the total energy of the *Imma* structure to be lower than that of β -Sn for all volumes considered in their calculation. The *Imma* and β -Sn structures are very similar in this pressure region and have nearly degenerate enthalpy curves. It has been argued that this similarity might make them experimentally indistinguishable at finite temperatures which leads to the discrepancy between theory and experiment [164].

The *Imma* structure is found to transform to the simple-hexagonal (sh) structure at 15.4 GPa by McMahon *et al.* [82], although other theoretical calculations have shown that the *Imma* phase is lower in energy than sh for all but the smallest volumes [164]. Our calculations predict a transition from *Imma* $\rightarrow$ sh at a pressure of 30.5 GPa, although this value is extremely sensitive because of the merging of the *Imma* and sh curves at low volumes and high pressures. The sh phase is observed to transform to the *Cmca* phase at a pressure of 38 GPa [165]. We find a transition pressure of 41.4 GPa, which is in somewhat better agreement than that of 33 GPa found in previous calculations [166]. From the *Cmca* phase silicon is found experimentally to transform to the hexagonal close-packed (hcp) phase at 42 GPa [165]. Our calculations predict a transition at 50.6 GPa, which is in slightly larger disagreement with the experimental measurement than a previous calculation by Christensen *et al.* of 41 GPa [166]. Finally, the hcp phase is found to transform at ~ 80 GPa to fcc and remains in this structure up to the highest pressure explored of 248 GPa [162]. Our calculations predict a transition pressure of 87.5 GPa, which is both in good agreement with experiment and with the previous value of 84.3 GPa calculated by Needs and Mujica [78].

Our calculations also allow us to determine the transition pressures obtained upon decompression from the β -Sn phase. We find a transition from β -Sn to the r8 phase at 8.4 GPa, which compares well with the experimental transition at 9.4 GPa [134] and previous calculations from Frommer *et al.* of 7.4 GPa [76]. We also calculate the transition from r8 to bc8 to occur at 1.8 GPa, in excellent agreement with experiment (2 GPa) [86] and prior theoretical calculations (1.9 GPa) [76].

An interesting new prediction from our enthalpy calculations is the fact that the previously predicted bct phase of Si [120] is found to have a lower enthalpy than the

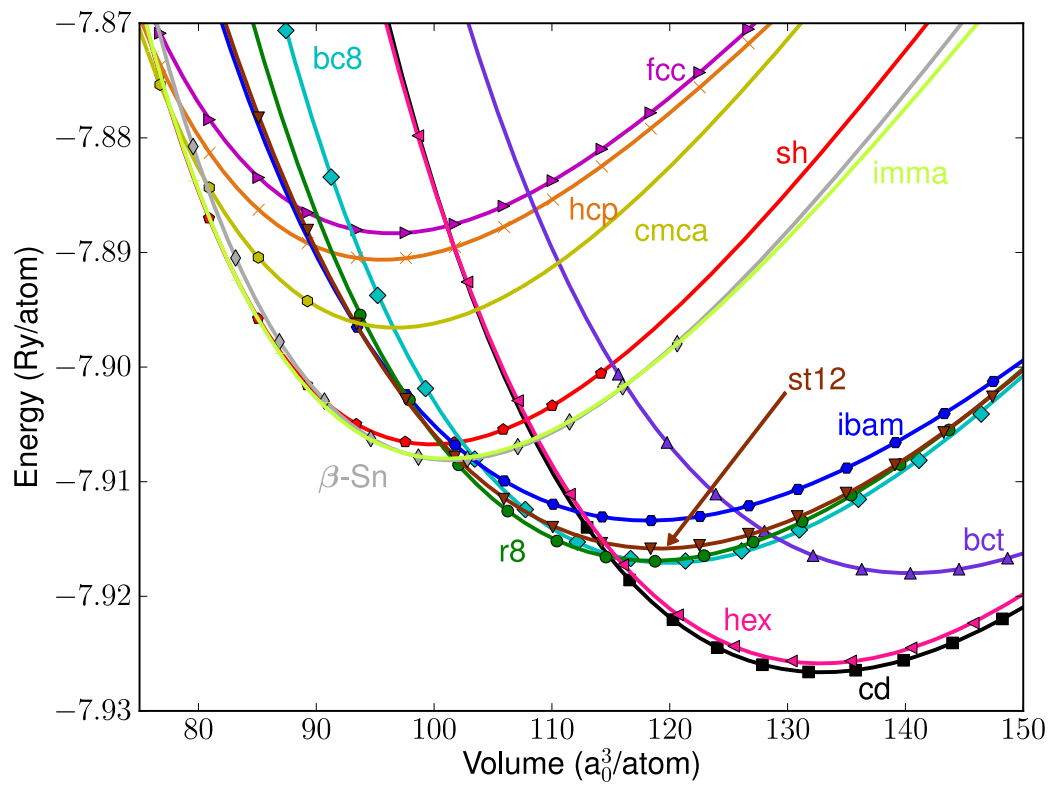


Figure 7.5: Comparison of the energy-volume curves for observed and predicted phases of silicon. The markers correspond to calculated data points and the lines are fits to the equation of state.

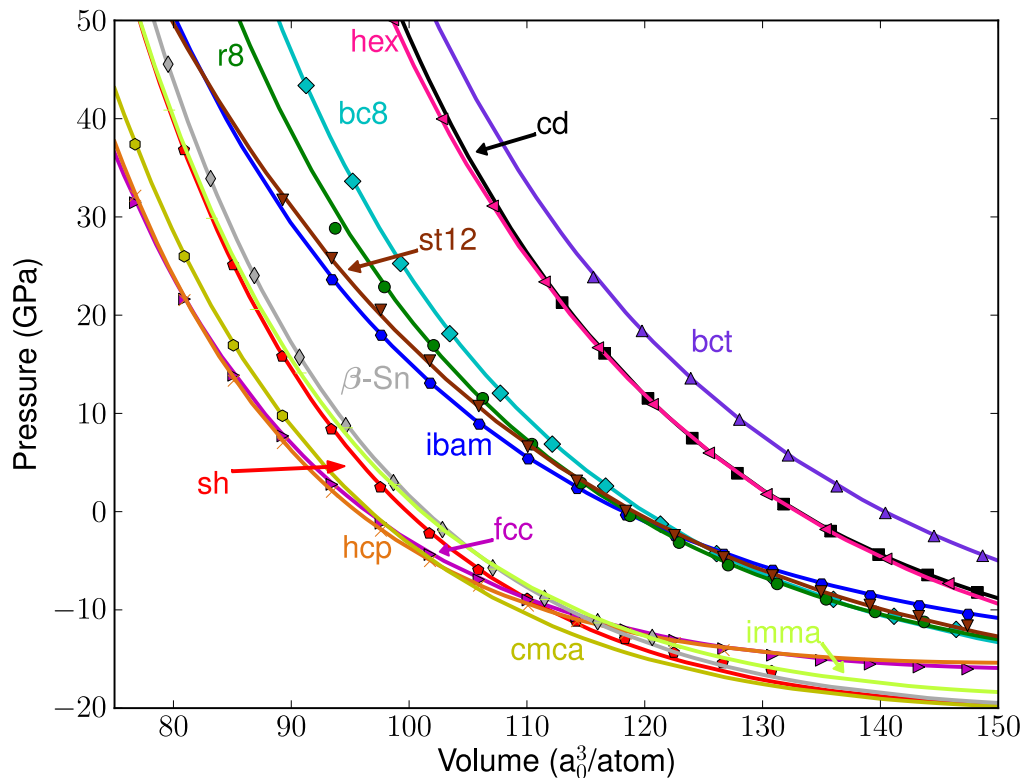


Figure 7.6: Pressure-volume curves for observed and predicted phases of silicon. The markers correspond to data points and the curves are obtained from numerical differentiation from the $E(V)$ equation of state.

bc8 phase at 0.65 GPa. This phase was recently discovered in molecular dynamics simulations on carbon [121] and subsequent calculations predicted its existence in Si [120]. Calculations by Fujimoto *et al.* [120] had compared the bct phase with cubic Si and predicted that cubic Si would transform to the bct structure at -11.2 GPa. This is in good agreement with our calculations suggesting the value of -13 GPa. While, to our knowledge, experiments have not obtained bct Si at ambient conditions upon depressurization, the positive coexistence pressure of 0.65 GPa indicates that bct Si may be easier to obtain than previous calculations have suggested.

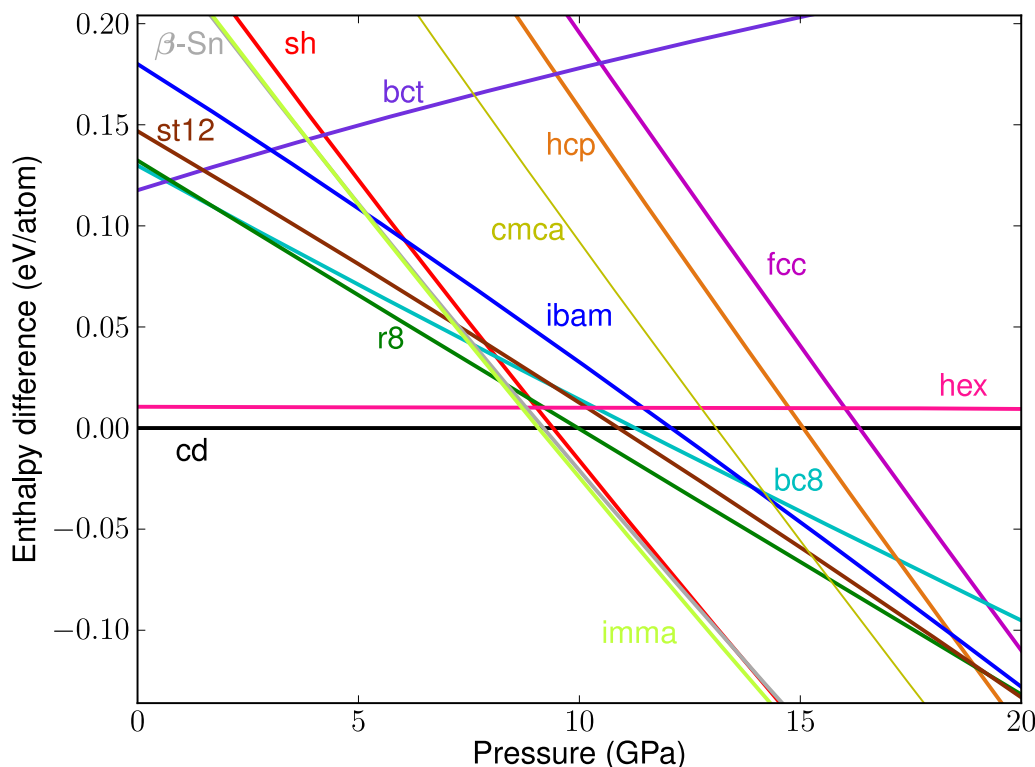


Figure 7.7: Enthalpy-pressure curves for observed and predicted phases of silicon. The enthalpy differences are taken relative to silicon in the cubic diamond phase.

7.4 Evolutionary methods: The search for Si-VIII

As we noted in the introduction to this chapter, another unknown phase was obtained upon rapid pressure release in addition to Si-IX - the phase known as Si-VIII [167]. Experimentally it was believed to consist of a tetragonal cell with $a=8.627$ Å and $c=7.500$ Å. Again, the internal parameters were not able to be determined, but by assuming that the phase had a similar density to that of Si-III (BC8), it was estimated to contain 30 atoms.

Before we begin to approach this problem, it is worth reconsidering our method with which we explore the energy landscape. In the search for Si-IX we used the *ab initio* random structure searching method. This approach was well suited because the unit cell was relatively small and we were able to brute-force explore the many structural arrangements that the atoms could take within the unit cell. However, as the number of atoms increases, the complexity of the problem increases exponentially [179, 180]. This implies that brute-force methods such as a random search may

take an unreasonably large number of attempts in order to find the minimum energy structure. In situations such as this, it is preferable to alter our approach and consider more sophisticated methods.

One approach which is currently widely used in these cases is that of an evolutionary algorithm. An evolutionary algorithm works by beginning with an initial set of structures (which can be taken as random). These structures then “evolve” over successive generations. Offspring are produced in a variety of ways, including heredity (in which new structures are obtained by joining together segments of structures from the previous generation), lattice mutation (where the lattice vectors are strained in some way in relation to those from the previous generation), atomic permutation (where atomic positions are swapped in a structure from the previous generation to form a new structure), and well as other mutations of the coordinate positions (e.g., using the concept of “local order” we could introduce random movements of the atomic positions, with the atoms in positions of lower symmetry – and thus generally higher energy – displaced by a larger amount) [179, 180]. By letting the population of structures evolve through the generations, giving a preferential bias towards structures of higher “fitness” (e.g., this is typically taken to be those with the lowest enthalpy), the population evolves towards the structures which have (typically) the lowest enthalpy in the complex space spanned by all possible structures.

In to examine structures that may correspond to that of experimentally observed Si-VIII we use an evolutionary structure predictor as described above. Specifically, we use the **USPEX** code [181], which interfaces with **VASP** [173–175] and other DFT codes. We perform an evolutionary structure search at the experimental lattice constants for cells with 30 or 32 atoms.¹ We begin the search by generating 100 random structures at the experimental lattice constants. The structures are not generated completely randomly, which is known to lead to a population with very low diversity [180], but are generated randomly within some subcell of the (30 or 32) atom supercell and this subcell is then repeated to fill the cell size of interest. The results of our evolutionary algorithm run for the 30 atom cell is shown in Fig. 7.8. The enthalpy of the population moves throughout the generations to smaller values as we approach the minimum of the energy landscape. By the 10th generation, we have already obtained a lowest energy candidate which does not go to lower enthalpy as the simulation continues, suggesting that this structure is may be the lowest 30-atom cell consistent with the experimental lattice parameters. This lowest energy structure is then relaxed with no constraints on the cell shape within **Quantum-ESPRESSO**. The cell has a slight nonuniform stress which distorts the perfect tetragonal lattice slightly and the resulting structure has space group P1. The cell parameters are given, along with a **Quantum ESPRESSO** input file, in Appendix B. Calculations of the phonon frequencies for this Si-VIII candidate suggest that the material is locally stable. Fur-

¹While the experiment estimated 30 atoms in the unit cell by assuming the density of the BC8 phase, we found that an estimate of 32 atoms could be obtained if one assumed the density to be equal to that of the R8 phase, which was unknown at the time.

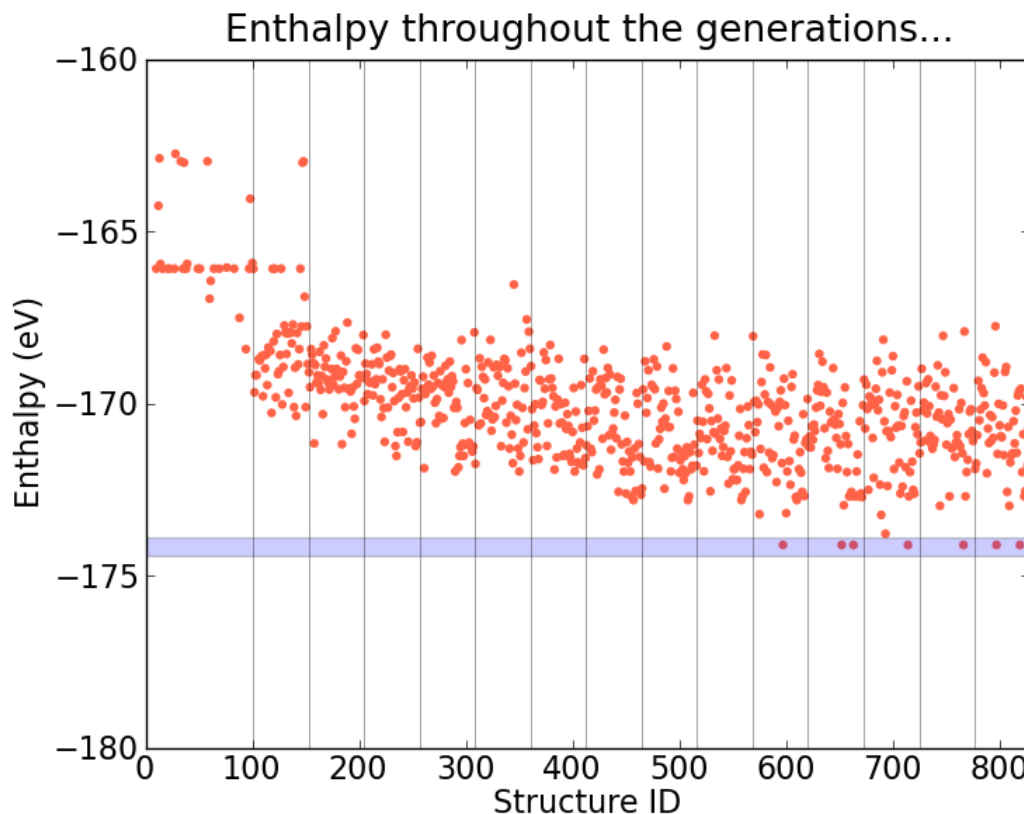


Figure 7.8: Relaxed structures resulting from a 30 atom run at the Si-VIII experimental lattice constants. Vertical lines separate structures belonging to different generations. The lowest enthalpy structure obtained at the end of the run is highlighted.

thermore, calculations of the band structure reveal it to be semiconducting with a gap of 0.66 eV. As we did in the case of the *Ibam* structure, we can look at the equation of state of this structure and compare it with other forms of Si. The energy as a function of volume is shown in Figure 7.9 and the associated enthalpy as a function of pressure is in Figure 7.10. As can be seen, the energy of this structure is actually quite low and very similar to metastable phases of Si that are known to exist. In fact, in Figure 7.10 it is interesting to that this Si-VIII candidate actually becomes very similar in enthalpy to BC8 Si, which is known to be the dominant phase upon decompression from β -Sn at ambient pressure. A simulation of the x-ray diffraction data for this structure shows some agreement with the x-ray diffraction profile obtained experimentally, although discrepancies remain which make a clear identification of this candidate as the experimentally observed phase difficult to make.

We now discuss the results of the evolutionary algorithm results on the 32 atom structural candidates. Our evolutionary run (figure not shown) results in two lowest

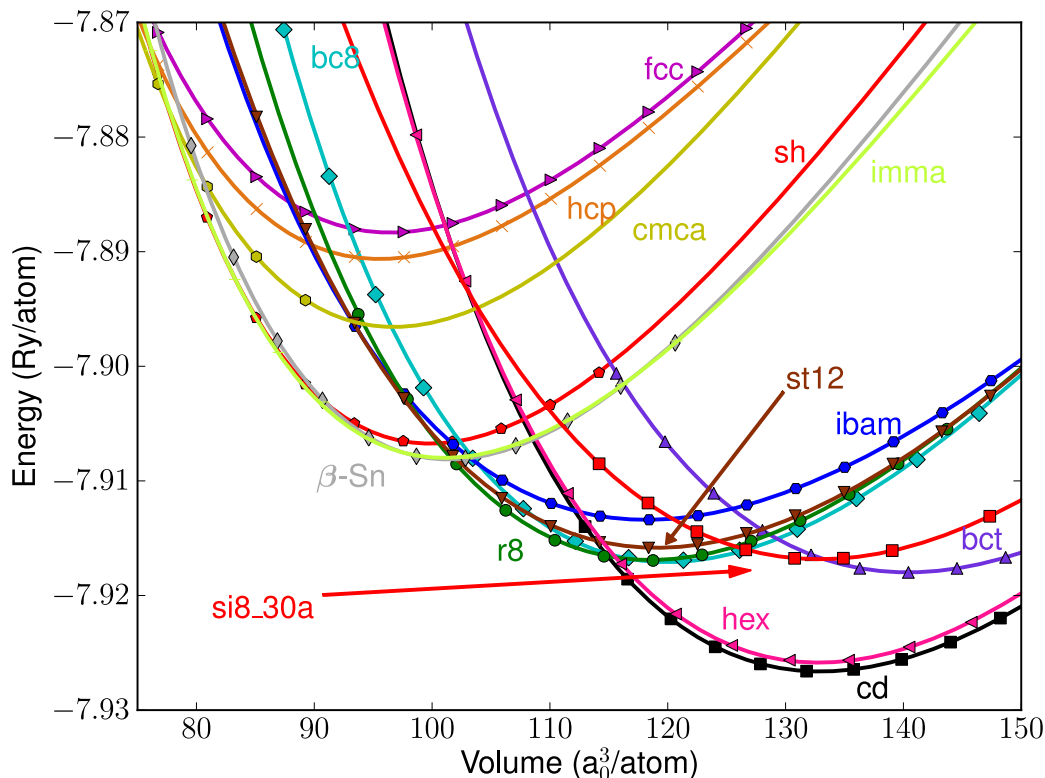


Figure 7.9: Shown is the energy vs volume for the 30 atom candidate for the Si-VIII phase as determined from the evolutionary algorithm run.

energy structures which are stable between generations ². The lowest energy of these structures turns out to simply be a supercell of the hexagonal diamond structure (lonsdaleite) that we discussed in Chapter 3 and thus we do not consider it as a structural candidate for Si-VIII. However, the idea that the Si-VIII phase was actually simply an observation of the hexagonal diamond phase is an interesting possibility. Instead, we turn our attention to the higher energy metastable phase with 32 atoms in the unit cell.

The obtained Si-VIII candidate with 32 atoms in the unit cell is actually a supercell of a 8 atom unit cell with space group 17 ($P222_1$), with four atoms at the 4(e)

²One might wonder how this is possible, since it seems that all structures should eventually go to the lowest energy structure. The USPEX code can be forced to take a number of structures from the previous generation and keep them for the later generation without modification. This by itself would lead to all the structures going to the lowest energy structure. However, by the use of “fingerprinting” analyses, which are used to determine whether structures are essentially identical, identical copies of the lowest energy structure can be removed from the population. This allows room for metastable structures to survive throughout the generations.

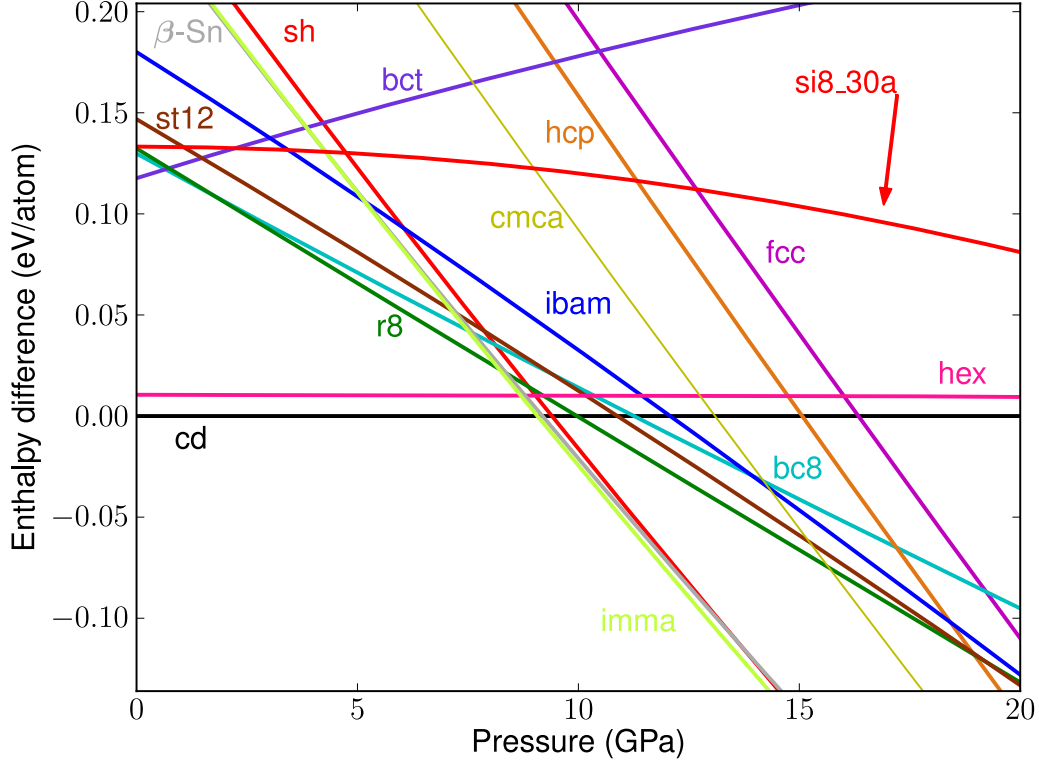


Figure 7.10: Enthalpy as a function of pressure for the 30 atom Si-VIII candidate. The enthalpy differences are taken relative to silicon in the cubic diamond phase.

sites with $x=0.25155$, $y=0.37306$, and $z=0.59475$, two atoms at the 2(c) sites with $y=0.19007$, and two atoms in the 2(a) sites with $x=0.33913$. The lattice parameters are orthorhombic with $a=3.9472$ Å, $b=8.678$ Å, and $c=4.0855$ Å. We contain a full Quantum ESPRESSO input file in Appendix B. A $2 \times 1 \times 2$ expansion of the unit cell of this structure is shown in Figure 7.11.

This structure turns out to also be semiconducting with a gap of 0.37 eV within LDA. Phonon calculations also suggest that this structure is locally stable. The equation of state of this structure is shown in Figure 7.12 and its enthalpy as a function of pressure in Figure 7.13. This structure is very similar from an energetic standpoint to the structure discussed more extensively above. As was the case with the 30 atom Si-VIII candidate, there is some agreement with the experimental x-ray diffraction data but not enough to make the identification conclusive.

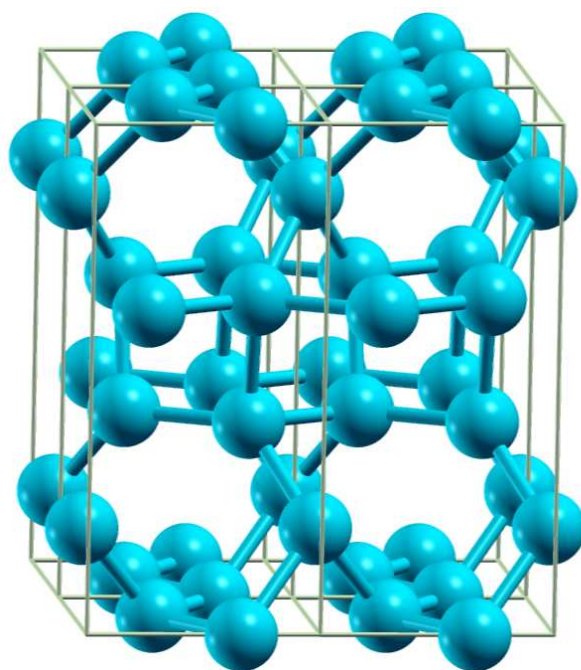


Figure 7.11: Structure of the Si-VIII candidate obtained from the evolutionary algorithm run starting from a 32 atom unit cell. The shown structure is a $2 \times 1 \times 2$ expansion of the 8-atom unit cell to better illustrate the bonding arrangement.

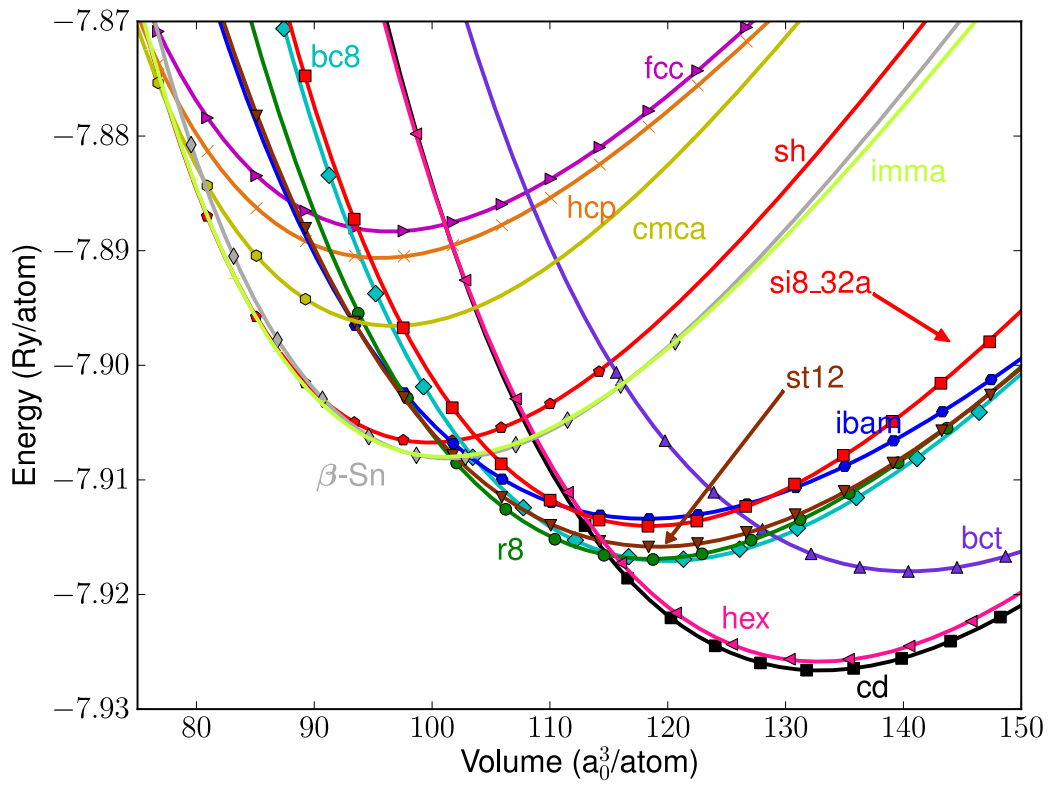


Figure 7.12: Shown is the energy vs volume for the 32 atom candidate for the Si-VIII phase as determined from the evolutionary algorithm run.

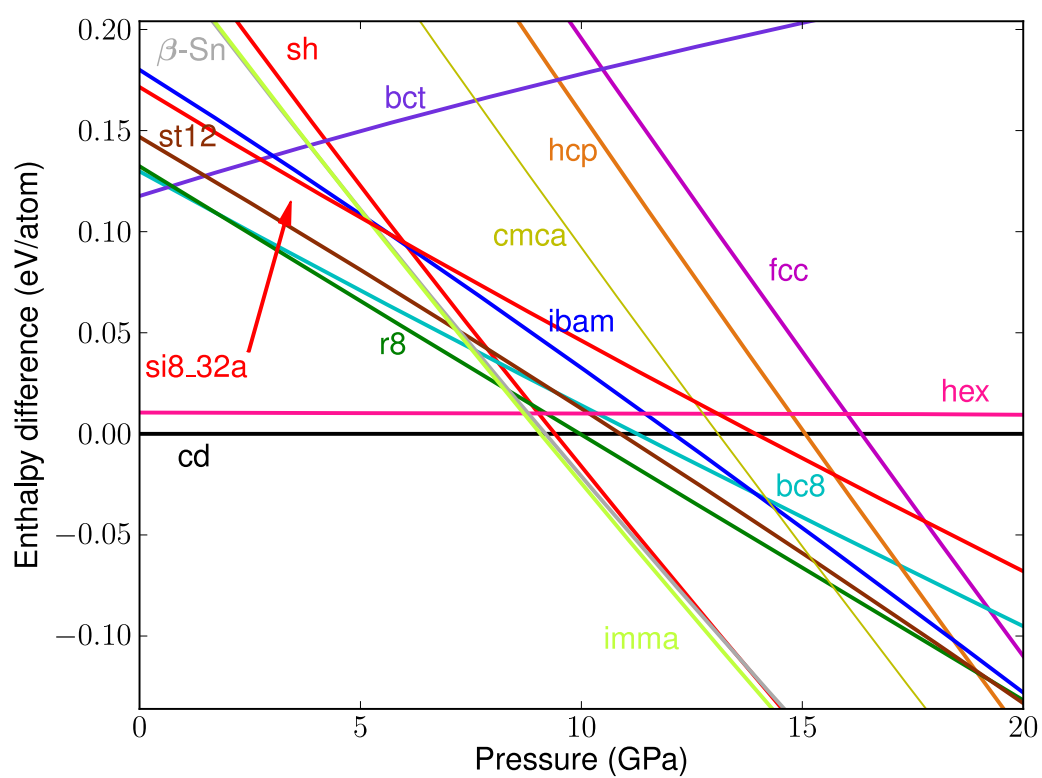


Figure 7.13: Enthalpy as a function of pressure curves for the 32 atom Si-VIII candidate. The enthalpy differences are taken relative to silicon in the cubic diamond phase.

7.5 Conclusion

In summary, we predict a new phase of silicon in the *Ibam* structure using *ab initio* random structure searching. This phase may be accessible experimentally in samples of Si decompressed from the β -Sn phase. Calculations of the electronic structure reveal that it is semimetallic within LDA with a small band overlap of 0.35 eV. It is expected that quasiparticle corrections using the GW approximation would open up the gap and result in a semiconductor. Calculations regarding energetic stability suggest that this phase is metastable over the pressure range studied. In addition, our calculated transition pressures for other known phases of silicon are in good agreement with experimental observation, often improving on previous theoretical results. We also find that the bct phase of Si recently predicted by Fujimoto *et al.* [120] has a calculated coexistence pressure of 0.65 GPa with the bc8 phase of Si, which is known to exist at ambient conditions upon depressurization from the β -Sn phase. Finally, we also discuss the results of an evolutionary algorithm structure search for the undetermined phase Si-VIII.

Chapter 8

First principles study of phosphorus and boron substitutional defects in Si-XII

8.1 Introduction

Silicon has a very rich phase diagram, having been clearly observed to exist in at least ten different phases depending upon the experimental conditions [79]. Most of these phases are formed upon the application of pressure, some of which are metastable at ambient conditions. One such phase which has attracted recent attention is that of Si-XII, a rhombohedral phase of silicon formed upon decompression from silicon's metallic β -Sn phase.

Si-XII has only recently been characterized, having been identified as a distinct phase a little over 15 years ago [84]. Since then there has been both theoretical and experimental interest in its properties. Early theoretical studies of its electronic structure [76] found that within the local density approximation (LDA) to density functional theory (DFT) Si-XII was semimetallic, exhibiting an indirect band gap between the T and the L points in the Brillouin zone. Calculations performed in Chapter 3 of the quasiparticle spectra using the GW approximation [32, 42] to the electron self-energy result in the prediction of Si-XII being a small band gap semiconductor, with an indirect band gap of approximately 0.24 eV[105]. In Chapter 4 it was shown that Si-XII, because of its large optical matrix elements at the low energies relevant to photovoltaics, might have promise for solar energy applications [122]. Early experimental work concentrated on the structural phase transformations induced by nanoindentation into c-Si or a-Si substrates without specifically resolving the metallic or semiconducting nature of Si-XII [107, 182–184]. However, recent experimental evidence has been presented supporting the prediction that Si-XII is semiconducting [140]. Another result from this experimental work is that Si-XII, like

conventional cubic Si (Si-I), can be doped both n- and p-type with phosphorus and boron dopants, respectively. The interesting difference, however, is that the dopants can be electrically activated at room temperature during the indentation process and do not need a subsequent annealing step at 600 °C as is the case with conventional polycrystalline Si-I [185]. Additionally, in hole-doped Si-XII, the carrier mobility is found to be higher than that of polycrystalline Si and more comparable to that of single-crystal cubic Si.

Because of the potential technological utility of a doping process that is possible at lower temperatures, there is interest to obtain more information on the properties of doping Si-XII with defects, in particular with phosphorus and boron. In this chapter, we address some of the basic properties of these dopants in Si-XII. Using the hybrid functional of Heyd, Scuseria, and Ernzerhof (HSE) [186, 187], we calculate the formation energies of substitutional P and B defects in Si-XII. From these formation energies we also estimate thermodynamic charge transition levels and classify the nature of the defects in Si-XII, as well as address any site preference of the substitutional defects to the two, distinct sites in Si-XII. Si-XII, which has the space group $R\bar{3}$ with 8 atoms in the unit cell, has two atoms located on the 2(c) sites and six on the 6(f) sites (see Fig. 8.1). While all the atomic sites have a distorted tetrahedral bonding configuration, the 2(c) sites are characterized by two distinct bond lengths and angles whereas the 6(f) sites are more complex, having three distinct bond lengths and six different bond angles [84]. The degree to which this different bonding environment will affect the formation energies of substitutional defects is unknown and is investigated in this chapter.

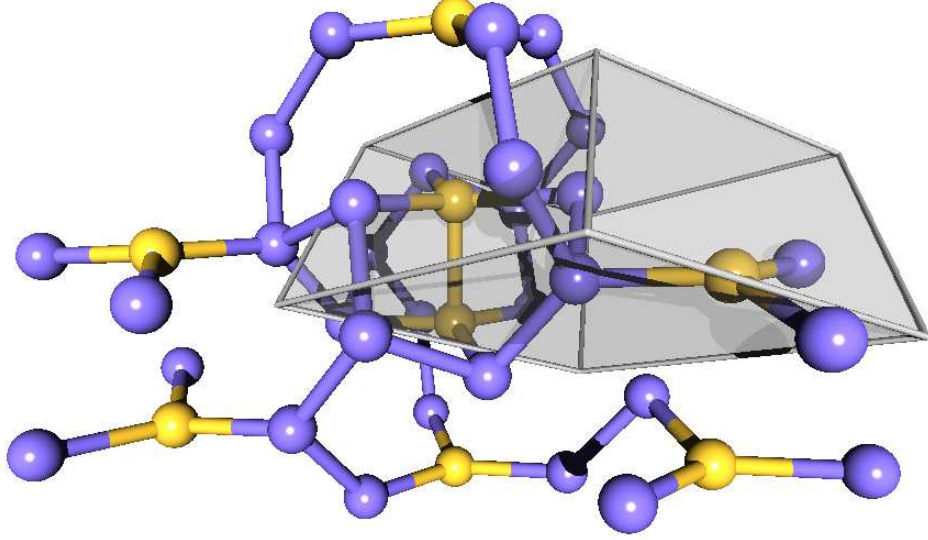


Figure 8.1: Structure of Si-XII. Gold (light) atoms are those at the 2(c) positions and violet (dark) are those at the 6(f) positions. Additional atoms above the 8 in the primitive cell are shown to illustrate the bonding.

8.2 Calculation of defect properties

The formation energy of a defect D in charge state q is calculated as

$$\begin{aligned}
 E^f(D, q) = & E_T(D, q) - E_T(\text{pure}) - \sum_i n_i \mu_i \\
 & + q(E_{VBM} + E_F + \Delta V)
 \end{aligned} \tag{8.1}$$

where $E_T(D, q)$ and $E_T(\text{pure})$ are the total energies of the supercell with and without the defect in charge state q . The defect is created by adding or removing n_i atoms with chemical potentials μ_i from the defect-free “pure” supercell [188]. E_F is the Fermi level referenced to the valence band maximum (VBM) in the pure cell. Because the VBM is referenced to the pure cell, an alignment term ΔV is added to align the electrostatic potentials between the defect and pure cells. In our calculations this term is taken as the difference in electrostatic potentials between the atomic site furthest from the defect in the defect cell with its equivalent site in the pure cell [38].

Despite the apparent simplicity of Eq. 8.1, the calculation of defect formation energies is made more difficult by two principle problems [37, 38, 189]: (i) band gap corrections to the DFT calculation and (ii) finite size effects caused by the use of the supercell approximation [190] and periodic boundary conditions (PBCs). The band gap corrections are caused by the well-known inability of DFT calculations

to reproduce the correct quasiparticle gap [30, 31]. This underestimation leads to problems in placing the defect transition levels relative to the band extrema and can also affect the occupations of the levels which will affect the relaxation of the defect in the host crystal.

In the case of Si-XII, the band gap problem manifests itself in an extreme way - the complete absence of a band gap within DFT. To overcome this problem, we have utilized the screened hybrid functional HSE, in which a fraction of Hartree-Fock (HF) exchange is mixed with the exchange potential of Perdew, Burke, and Ernzerhof (PBE) [26]. In contrast to PBE0, the HSE functional employs a screened Coulomb potential for the exchange interaction [191]. Our calculations were performed using the plane-wave projector augmented-wave (PAW) method with the HSE hybrid functional as implemented in the VASP code [173–175]. The screening parameter used in the splitting of the Coulomb parameter into short- and long-range parts is $\omega = 0.2 \text{ \AA}^{-1}$ while the HF mixing constant is set to 25%, as determined by perturbation theory [192, 193]. In a study by Heyd *et al.*, it was found that HSE improves the value of the band gap substantially over that of the pure DFT functionals [194]. Additionally, Batista *et al.*, in their study of the silicon self-interstitial, found that the HSE functional produces self-interstitial formation energies as well as energy differences between silicon in the diamond and metallic β -Sn phases in very close agreement to diffusion Monte Carlo (DMC) results [195]. The experimental data surrounding the silicon self-interstitial is a matter of controversy, but it has been argued that the DMC results better explain the experimental data than that of the pure DFT functionals [195, 196]. As the bonding environment in Si-XII is similar to that of Si-I, it can be hoped that such an improvement in the calculation of energy differences between similar structures will also be present in the use of the HSE functional for Si-XII defects and may improve the accuracy of the calculated formation energies beyond what would be obtained from pure DFT functionals.

The other difficulty in using Eq. 8.1 to calculate the formation energy is that even though we are often interested in the properties of an isolated defect, computational necessity requires us to consider a single defect in a supercell of finite extent with the application of PBCs. The use of a finite cell can express itself in unphysical constraints on the lattice relaxation around a defect site, although in many cases these elastic errors to the total energy can be made sufficiently small by employing moderately-sized cells. A larger problem, in the case of charged defects, is the spurious electrostatic interactions between defects in one cell and its periodic images implied by the use of PBCs. These errors typically converge much less rapidly with cell size and are difficult to converge “brute-force” by going to extremely large cell sizes.

In order to account for these spurious electrostatic interactions, a number of correction schemes have been used in the literature [197, 198]. The most common is arguably that of Makov and Payne [199], in which the energy of a defect in a cubic supercell of length L is used to extrapolate to the isolated defect limit ($L \rightarrow \infty$) using the expression

$$E_T(L) = E_T(L \rightarrow \infty) - \frac{q^2\alpha}{2\epsilon L} - \frac{2\pi qQ}{3\epsilon L^3} + \mathcal{O}(L^{-5}). \quad (8.2)$$

In the above equation α is the supercell dependent Madelung constant, ϵ is the static dielectric constant of the embedding host medium, and Q is the second radial moment of the electron density difference between the defect and pure supercells. In previous works, Eq. 8.2 has been applied directly and has also been used as the justification for a fitting procedure in powers of $1/L$ [189, 200, 201]. In this work we use a variant of the fitting procedure introduced by Hine *et al.* [202] This method makes use of the fact that the first-order term of the expansion in Eq. 8.2 is the Madelung energy $-q^2\alpha/2L$ of a point charge q in a neutralizing background, divided by the dielectric constant ϵ . Thus if we neglect the higher order terms of the expansion in Eq. 8.2, which are expected to be small because of our large cell sizes, the electrostatic correction can be written as

$$E_T(L) = E_T(L \rightarrow \infty) - bE_M \quad (8.3)$$

$$b = \frac{1}{\epsilon}.$$

Eq. 8.3 can then be fit using the total energies of the defect supercells along with the Madelung energy determined by that supercell's particular geometry. This generalization of the scaling method for the electrostatic corrections has the advantage that arbitrary cell shapes can be treated, unlike the conventional application which assumes the same supercell geometry and hence Madelung constant α . Not only does using the Madelung energies directly allow for the electrostatic correction in non-cubic supercells, but allowing for different supercell geometries in the fitting procedure allows for a much larger number of data points to be considered than in the case where all dimensions of the primitive cell have to be scaled equally in the construction of the supercell. The fitting scheme results in the extrapolated total energy of the defect supercell in the isolated limit as well as the parameter b , which is inversely related to the effective dielectric constant in the host material. Finally, it should be noted that phosphorus and boron substitutional defects in cubic silicon are found to be very shallow. Shallow defects are associated with very extended wavefunctions and result in a spurious interaction because of the overlap of these wavefunctions with their periodic images. We have calculated the properties of these defects in Si-XII with as large of supercells as is currently possible, while at the same time utilizing the HSE functional to correct for the bandgap underestimation, in order to minimize the spurious overlap of defect wavefunctions and provide the best estimate possible for the thermodynamic transition levels.

8.3 Results

Within the LDA, Si-XII has previously been calculated to have an indirect band overlap of ~ 0.3 eV [76, 105]. The calculation of the quasiparticle spectrum using the GW approximation for the electron self-energy results in a band opening to 0.24 eV [105]. Using the hybrid functional HSE we have calculated the Kohn-Sham single particle eigenvalues on a $6 \times 6 \times 6$ Monkhorst-Pack grid centered on Γ with a plane-wave energy cutoff of 320 eV. We find that the band gap of Si-XII within HSE to be 0.234 eV, only 6 meV different from our prior GW result. This result is in agreement with the behavior found by Heyd *et al.* [194] in their study of 40 bulk semiconductors in which they found that HSE works remarkably well in the prediction of semiconductor band gaps, especially that of small band gap materials. It should be noted, however, that this HSE result is obtained on the Si-XII structure relaxed with the HSE functional whereas our previous GW results were done at the lattice constant determined by LDA. Using the hybrid HSE functional on the LDA-relaxed structure results in a lower band gap of 0.13 eV, as expected due to the smaller lattice constant in materials calculated with LDA (5.70 Å) as opposed to HSE (5.73 Å) [194, 203]. For the purposes of the defect calculations the fact that the relaxed HSE structure reproduces the prior GW result simplifies the matter as no ad-hoc band gap correction schemes are needed in the interpretation of the results [204]. This result is also important in that it represents an independent prediction that Si-XII is semiconducting.

For the calculation of the formation energy of the P and B substitutional defects we use a supercell of 216 atoms, which corresponds to a $3 \times 3 \times 3$ expansion of the primitive unit cell of 8 atoms. We use a $2 \times 2 \times 2$ Γ -centered Monkhorst-Pack grid of $\mathbf{k}$ points to represent the Brillouin zone integration, which is the equivalent sampling density used in the calculations of the primitive cell previously discussed. The P or B defect is placed at either the 2(c) or the 6(f) atomic position. The atomic positions are then allowed to fully relax with the hybrid HSE functional until the forces on the atoms are all smaller than 0.01 eV/Å. This supercell of 216 atoms is large enough to contain the relaxation of the defect in the host Si-XII crystal such that any additional elastic energy contributions from using larger cells are found to be insignificant.

As described in Sec. II, we use a variant of the finite-scaling scheme to account for the electrostatic error implied by the use of PBCs with charged defects. This approach requires the calculation of total energies for the defect in a number of host cells of varying size. These cells are constructed by placing the defect (either P or B) in the center of a 216-atom supercell (at either the 2(c) or 6(f) position) and relaxing the atomic positions until the forces are smaller than 0.01 eV/Å. Because of the relaxation, these supercells then have approximately the correct lattice relaxation around the defect. These “seed” supercells are then placed in larger supercells, as indicated in Table (8.1). Importantly, in these larger supercells the atomic positions are not relaxed. Thus, in the fitting procedure of Eq. 8.3 only electrostatic screening is

Table 8.1: The supercells used in the electrostatic scheme of Eq. 8.3. The supercell size is given as an expansion of the primitive cell of 8 atoms. The Brillouin zone (BZ) sampling is given in the final column, and is in all cases greater than or equal to density used in the converged calculations on the Si-XII unit cell.

Supercell size	Number of atoms	BZ sampling
3x3x3	216	2x2x2
3x3x4	288	2x2x2
3x3x5	360	2x2x2
3x3x6	432	2x2x1

present. The resulting electrostatic correction obtained by this means is then applied to the original 216-atom fully-relaxed supercell, which has already been determined to have only small elastic errors. This can be seen in Fig. 8.2 which shows the displacement of the atoms around each defect as a function of their unrelaxed distance from the substitutional defect site. The defect displacements decay to a very small value at the edges of the 216-atom “seed” cell. As is expected with boron substitution, the nearby displacements of the silicon atoms are large, although they too decay to small values far from the substitutional site. The results of the electrostatic fitting scheme are shown in Fig. 8.2 for the case of a P substitutional defect at the 2(c) atomic site of Si-XII. The data points do exhibit a reasonable linear trend, as expected from the dependence on the electrostatic error on the Madelung energy seen in Eq. 8.3. The curves for the P defect at the 6(f) site and the B defect at the 2(c)/6(f) sites are similar.

The formation energies resulting from the 216-atom relaxed supercells along with the electrostatic corrections from the fit are shown in Fig. 8.4, for defects placed at the 2(c) and 6(f) sites, respectively. The band gap of Si-XII is shown shaded, with the valence band maximum (VBM) being located at $E_F = 0$ and the conduction band minimum (CBM) located at $E_F = 0.234$ eV. The thermodynamic charge transition level is defined as the Fermi level position at which the formation energies between two different charge states are equal to each other [37, 38]. It can also be computed from the equation

$$\epsilon(q, q') = \frac{[E_{VBM}^f(D, q) - E_{VBM}^f(D, q')]}{q - q'}, \quad (8.4)$$

where q and q' correspond to two different charge states and the formation energies are evaluated at $E_F = E_{VBM}$. Without the electrostatic corrections of Eq. 8.3 the thermodynamic transitions levels, between the neutral and +1 state for P and between the neutral and -1 state for B, are resonant with the conduction and valence bands, respectively. After correcting for the electrostatic error, however, these levels become are placed inside the band gap of Si-XII. The thermodynamic transition levels and

the formation energies evaluated in the middle of the band gap are given in Table (8.2). Also shown is the effective dielectric constant, $1/b$, from each fit. For P, a shallow donor level between 67 and 69 meV is found, which is very similar to the 45 meV that is found experimentally for cubic Si [3]. For B, we find acceptor levels which differ more strongly depending on the atomic site, located at 94 and 165 meV above the valence band. This greater dependence on the substitutional site may be due to the much larger displacement of silicon neighbors with boron substitution which amplifies differences in the local environment. The effective dielectric constants obtained from the fitting procedure are similar when compared between sites of the same substitutional defect. For P defects, a value is obtained which is somewhat larger than what is known for silicon in the cubic phase. This would be expected because of the smaller gap of Si-XII. For boron defects, however, the effective dielectric constant from the fit is lower for both sites. It is not clear why this should be the case. It could be that for the boron substitutional defect electrostatic corrections beyond the monopole term become important or perhaps that the much larger lattice distortion in the case of boron affects the estimated dielectric constant. Finally, we note that using a Makov-Payne correction for shallow defects will likely overestimate the degree of electrostatic correction. For a perfectly delocalized defect state (i.e, a constant background charge) there is an electrostatic energy correction that exactly cancels that of the leading term in the Makov-Payne expansion [205], although this effect may be minor in our case because of the use of the inverse of ϵ as a fitting parameter instead of using the presently unknown, experimental dielectric constant.

As seen in the table, there is a slight preference, both for the B and the P substitutional defect, to go into the 6(f) atomic site of the lattice, with preferences for the 6(f) site of 0.13 eV and 0.07 eV for B and P, respectively. Also seen in the table is that, for the choice of reference chemical potentials taken here, that the formation energy for B substitutional defects is higher than that of P. This difference is of similar magnitude as what has been found in B- and P-doped cubic silicon nanocluster

Table 8.2: Summary of the formation energy data in the neutral charge state and thermodynamic charge transition levels shown in Figs. 8.3 and 8.4. For each defect the formation energy is calculated at the center of the gap. $\epsilon(q, q')$ is the charge transition level between the relevant charge states, including the electrostatic corrections. $1/b$ is the effective dielectric constant determined from the fitting of Eq. 8.3. All energies listed are in eV.

Defect	$E^f(E_F=0.5E_g)$	$\epsilon(q, q')$	$1/b$
P @ 2(c)	0.18	0.067 below CBM	14.0
P @ 6(f)	0.11	0.069 below CBM	15.4
B @ 2(c)	0.82	0.094 above VBM	6.7
B @ 6(f)	0.69	0.165 above VBM	4.2

calculations which used the same reference phases of B_{50} and black phosphorus [206].

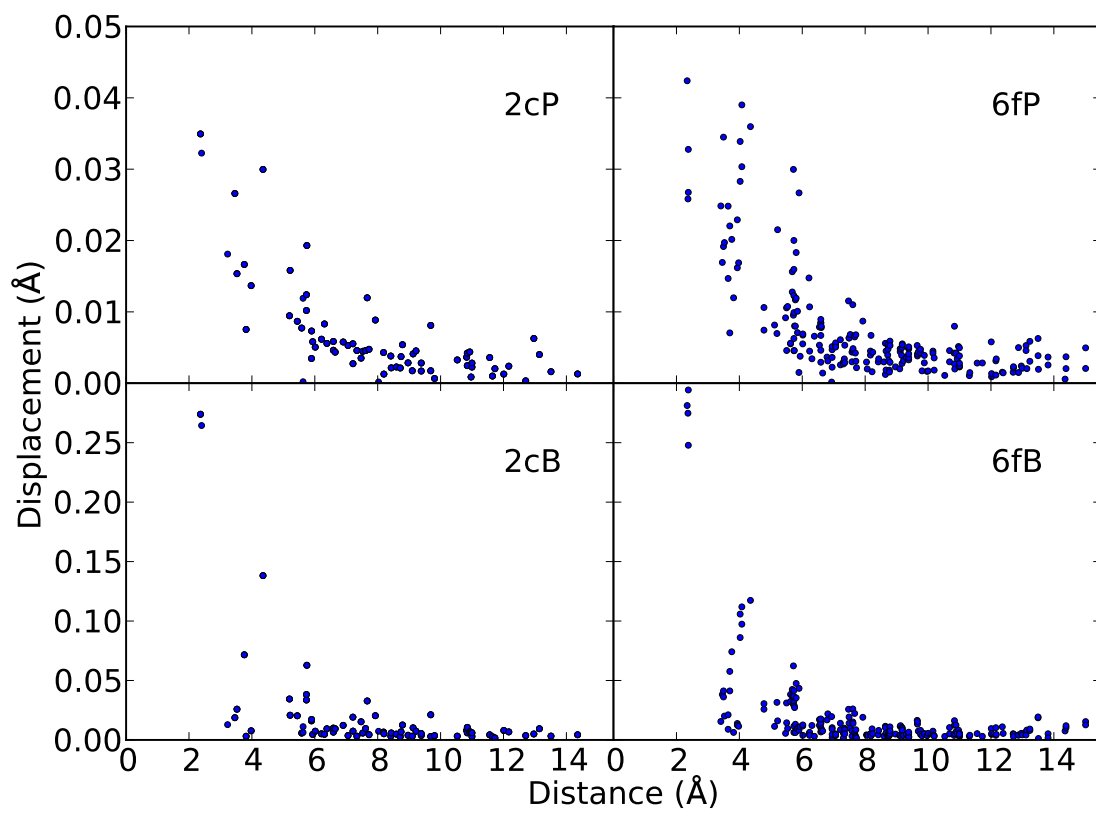


Figure 8.2: Shown are the displacements of the atoms around the defect in relation to their unrelaxed distance from the defect site in the 216-atom relaxed “seed” cells for the phosphorus and boron substitutional defects at the 2(c) and 6(f) atomic sites.

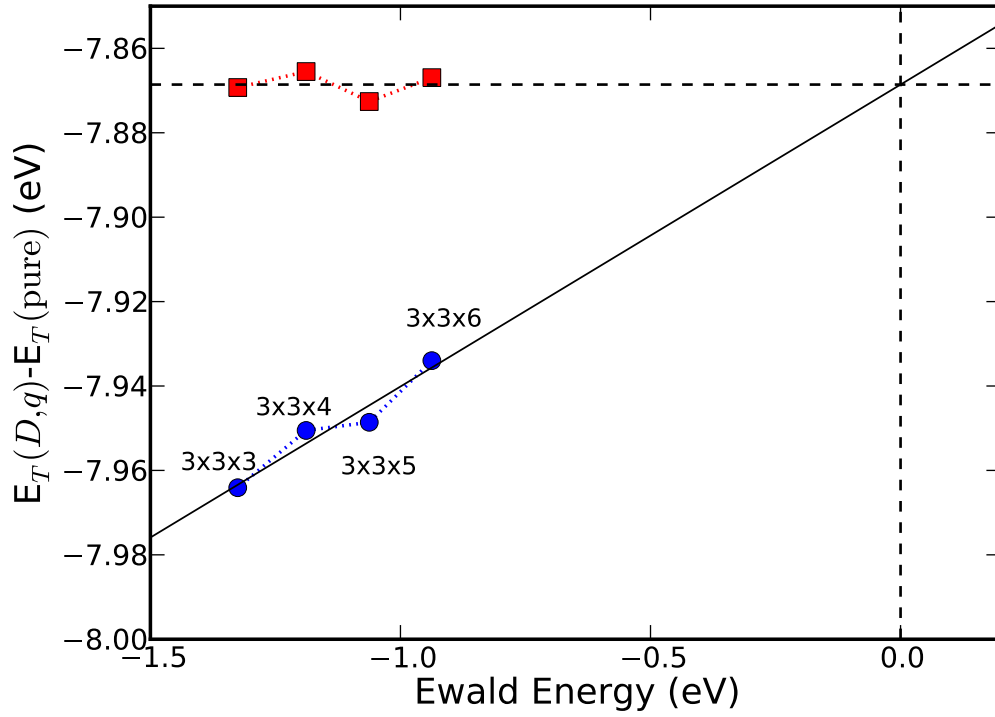


Figure 8.3: The data for the P defect at the 2(c) site in Si-XII used in the electrostatic correction scheme. Blue circles correspond to the total energy differences between the defect and pure supercells shown in Table (8.1) as a function of the Madelung energy. The solid line is a linear fit to the data as given by Eq. 8.3. Red squares denote the same supercells but shifted by the linear corrections. The infinite cell limit is obtained at zero Madelung energy.

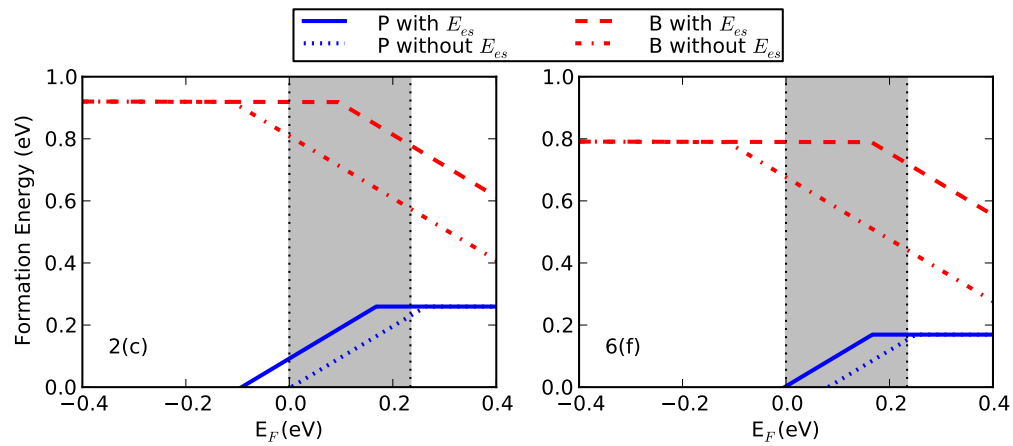


Figure 8.4: Formation energies of the P and B substitutional defects at the 2(c) (left) and 6(f) (right) atomic sites of Si-XII. Shown are both the data with and without electrostatic corrections E_{es} . The energy range of the band gap is shown shaded.

8.4 Conclusion

In conclusion, we have performed a first-principles study of substitutional P and B defects in Si-XII, a high-pressure polytype of silicon metastable at ambient temperature and pressure. We have used the hybrid functional HSE to evaluate formation energies and thermodynamic transition levels. We find P to have shallow donor levels, similar to the case of P in cubic silicon. Substitutional B in Si-XII is estimated to have somewhat deeper levels in relation to what is experimentally found in cubic silicon and depend more strongly on the atomic site. We find a slight preference for the defects to go into the 6(f) atomic site of Si-XII. Additionally, the HSE functional predicts Si-XII to be semiconducting, at odds with DFT-LDA calculations but in agreement with recent experimental evidence [140] and with previously calculated GW results [105].

Bibliography

- [1] J. Kohanoff. *Electronic Structure Calculations for Solids and Molecules*. Cambridge University Press, 2006.
- [2] R. M. Martin. *Electronic Structure: Basic Theory and Practical Methods*. Cambridge University Press, 2004.
- [3] C. Kittel. *Introduction to Solid State Physics*. Wiley, New York, 6th ed. edition, 1986.
- [4] N. W. Ashcroft and N. D. Mermin. *Solid State Physics*. Thomson Learning, 1976.
- [5] E. Kaxiras. *Atomic and Electronic Structure of Solids*. Cambridge University Press, 2007.
- [6] J. C. Phillips and L. Kleinman. New Method for Calculating Wave Functions in Crystals and Molecules. *Phys. Rev.*, 116, 1959.
- [7] M. L. Cohen and J. R. Chelikowsky. *Electronic Structure and Optical Properties of Semiconductors*. Springer-Verlag, 1988.
- [8] C. Herring. A New Method for Calculating Wave Functions in Crystals. *Phys. Rev.*, 57:1169, 1940.
- [9] C. Fiolhais, F. Nogueira, and M. Marques. A Tutorial on Density Functional Theory. In *A Primer in Density Functional Theory*, page 218. Springer-Verlag, 2003.
- [10] D. R. Hamann, M. Schlüter, and C. Chiang. Norm-Conserving Pseudopotentials. *Phys. Rev. Lett*, 43:1494, 1979.
- [11] J. Ihm, A. Zunger, and M. L. Cohen. Momentum-space formalism for the total energy of solids. *J. Phys. C: Solid State Phys.*, 12:4409, 1979.
- [12] M. L. Cohen and T. K. Bergstresser. Band Structures and Pseudopotential Form Factors for Fourteen Semiconductors of the Diamond and Zinc-blende Structures. *Phys. Rev.*, 141:789, 1966.

- [13] J. R. Chelikowsky and M. L. Cohen. Nonlocal pseudopotential calculations for the electronic structure of eleven diamond and zinc-blende semiconductors. *Phys. Rev. B*, 14:556, 1976.
- [14] L. H. Thomas. The calculation of atomic fields. *Proc. Cambridge Phil. Soc.*, 23:542, 1927.
- [15] E. Fermi. A statistical method for the determination of some properties of atoms. II. Application to the periodic system of the elements. *Z. Phys.*, 48:73, 1928.
- [16] P. A. M. Dirac. Note on exchange phenomena in the Thomas atom. *Proc. Cambridge Phil. Soc.*, 26:376, 1930.
- [17] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136:B864, 1964.
- [18] W. Kohn, A. D. Becke, and R. G. Parr. Density Functional Theory of Electronic Structure. *J. Phys. Chem.*, 100:12974, 1996.
- [19] W. Kohn and L. J. Sham. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.*, 140:A1133, 1965.
- [20] W. Kohn. Nobel Lecture: Electronic structure of matter - wave functions and density functionals. *Rev. Mod. Phys.*, 71:1253, 1999.
- [21] R. G. Parr and Yang Weitao. *Density-Functional Theory of Atoms and Molecules*. Oxford University Press, 1994.
- [22] S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi. Phonons and related crystal properties from density-functional perturbation theory. *Rev. Mod. Phys.*, 73:515, 2001.
- [23] D. M. Ceperley and B. J. Alder. Ground State of the Electron Gas by a Stochastic Method. *Phys. Rev. Lett.*, 45:566, 1980.
- [24] J. P. Perdew and A. Zunger. Self-interaction correction to density-functional approximations for many-electron systems. *Phys. Rev. B*, 23:5048, 1981.
- [25] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais. Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation. *Phys. Rev. B*, 46:6671, 1992.
- [26] J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.*, 77:3865, 1996.

- [27] M. T. Yin and M. L. Cohen. Theory of static structural properties, crystal stability, and phase transformations: Application to Si and Ge. *Phys. Rev. B*, 26:5668, 1982.
- [28] M. T. Yin and M. L. Cohen. Microscopic Theory of the Phase Transformation and Lattice Dynamics of Si. *Phys. Rev. Lett.*, 45:1004, 1980.
- [29] M. L. Cohen. Pseudopotentials and Total Energy Calculations. *Phys. Scr.*, T1:5, 1982.
- [30] J. P. Perdew and M. Levy. Physical Content of the Exact Kohn-Sham Orbital Energies: Band Gaps and Derivative Discontinuities. *Phys. Rev. Lett.*, 51:1884, 1983.
- [31] L. J. Sham and M. Schluter. Density-Functional Theory of the Energy Gap. *Phys. Rev. Lett.*, 51:1888, 1983.
- [32] W. G. Aulbur, L. Jönsson, and J. W. Wilkins. Quasiparticle Calculations in Solids. In H. Ehrenreich and F. Spaepen, editors, *Solid State Physics*. Academic Press, 2000. Vol. 54.
- [33] M. Rohlfing and S. G. Louie. Electron-hole excitations and optical spectra from first principles. *Phys. Rev. B*, 62:4927, 2000.
- [34] X. Zhu and S. G. Louie. Quasiparticle band structure of thirteen semiconductors and insulators. *Phys. Rev. B*, 43:14142, 1991.
- [35] S. B. Zhang, D. Tománek, S. G. Louie, M. L. Cohen, and M. S. Hybertsen. Quasiparticle calculation of valence band offset of AlAs-GaAs(001). *Solid State Commun.*, 66:585, 1988.
- [36] C. G. Van de Walle and R. M. Martin. Theoretical study of band offsets at semiconductor interfaces. *Phys. Rev. B*, 35:8154, 1987.
- [37] S. Lany and A. Zunger. Assessment of correction methods for the band-gap problem and for finite-size effects in supercell defect calculations: Case studies for ZnO and GaAs. *Phys. Rev. B*, 78:235104, 2008.
- [38] C. G. van de Walle and J. Neugebauer. First-principles calculations for defects and impurities: Applications to III-nitrides. *J. Appl. Phys.*, 95:3851, 2004.
- [39] K. S. Thygesen and A. Rubio. Nonequilibrium GW approach to quantum transport in nano-scale contacts. *J. Chem. Phys.*, 126:091101, 2007.
- [40] L. Hedin and S. Lundqvist. Effects of Electron-Electron and Electron-Phonon Interactions on the One-Electron States of Solids. In F. Seitz, D. Turnbull, and H. Ehrenreich, editors, *Solid State Physics*. Academic Press, 1969. Vol. 23.

- [41] L. Hedin. New Method for Calculating the One-Particle Green's Function with Application to the Electron-Gas Problem. *Phys. Rev.*, 139:A796, 1965.
- [42] M. S. Hybertsen and S. G. Louie. Electron correlation in semiconductors and insulators: Band gaps and quasiparticle energies. *Phys. Rev. B*, 34:5390, 1986.
- [43] M. van Schilfgaarde, T. Kotani, and S. Faleev. Quasiparticle Self-Consistent GW Theory. *Phys. Rev. Lett.*, 96:226402, 2006.
- [44] M. Shishkin, M. Marsman, and G. Kresse. Accurate Quasiparticle Spectra from Self-Consistent GW Calculations with Vertex Corrections. *Phys. Rev. Lett.*, 99:246403, 2007.
- [45] R. Gómez-Abal, X. Li, and M. Scheffler. Influence of the Core-Valence Interaction and of the Pseudopotential Approximation on the Electron Self-Energy in Semiconductors. *Phys. Rev. Lett.*, 101:106404, 2008.
- [46] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch. QUANTUM ESPRESSO: a modular and open-source software project for quantum simulation of materials. *J. Phys.: Condens. Matter*, 21:395502, 2009.
- [47] M. J. T. Oliveira and F. Nogueira. Generating relativistic pseudo-potentials with explicit incorporation of semi-core states using APE, the Atomic Pseudo-potentials Engine. *Comp. Phys. Commun.*, 178:524, 2008.
- [48] M. L. Tiago, S. Ismail-Beigi, and S. G. Louie. Effect of semicore orbitals on the electronic band gaps of Si, Ge, and GaAs within the GW approximation. *Phys. Rev. B*, 69:125212, 2004.
- [49] M. Rohlfing, P. Krüger, and J. Pollmann. Role of semicore d electrons in quasiparticle band-structure calculations. *Phys. Rev. B*, 57:6485, 1998.
- [50] Z. Zanolli, F. Fuchs, J. Furthmüller, U. von Barth, and F. Bechstedt. Model GW band structure of InAs and GaAs in the wurtzite phase. *Phys. Rev. B*, 75:245121, 2007.
- [51] M. Rohlfing, P. Krüger, and J. Pollmann. Quasiparticle Band Structure of CdS. *Phys. Rev. Lett*, 75:3489, 1995.

- [52] C. L. Reis, J. M. Pacheco, and José Luís Martins. First-principles norm-conserving pseudopotential with explicit incorporation of semicore states. *Phys. Rev. B*, 68:155111, 2003.
- [53] O. Madelung, U. Rössler, and M. Schulz, editors. *The Landolt-Börnstein Database*, volume 41A1a, chapter Group IV Elements, IV-IV and III-V Compounds. Springer Materials, 2001.
- [54] O. Madelung, editor. *The Landolt-Börnstein Database*, volume 17a, chapter Numerical Data and Functional Relationships in Science and Technology - Crystal and Solid State Physics. Springer, 1984.
- [55] J. Deslippe, G. Samsonidze, D. A. Strubbe, M. Jain, M. L. Cohen, and S. G. Louie. BerkeleyGW: A massively parallel computer package for the calculation of the quasiparticle and optical properties of materials and nanostructures. *Comp. Phys. Commun.*, 183:1269, 2012.
- [56] J. Deslippe, G. Samsonidze, M. Jain, M. L. Cohen, and S. G. Louie. Coulomb-hole summations and energies from GW calculations with limited number of orbitals: a modified static remainder approach. (unpublished).
- [57] M. S. Hybertsen and S. G. Louie. Spin-orbit splitting in semiconductors and insulators from the *ab initio* pseudopotential. *Phys. Rev. B*, 34:2920, 1986.
- [58] A. A. Mostofi, J. R. Yates, Y.-S. Lee, I. Souza, D. Vanderbilt, and N. Marzari. **wannier90**: A tool for obtaining maximally-localised Wannier functions. *Comp. Phys. Commun.*, 178:685, 2008.
- [59] Emmanouil Kioupakis, Murilo L. Tiago, and S. G. Louie. Quasiparticle electronic structure of bismuth telluride in the *GW* approximation. *Phys. Rev. B*, 82:245203, 2010.
- [60] Emmanouil Kioupakis. *Quasiparticle calculations for solids and molecules*. PhD thesis, University of California, Berkeley, 2008.
- [61] O. Madelung, A. Goldmann, and E.-E. Koch, editors. *The Landolt-Börnstein Database*, volume 23a, chapter Electronic Structure of Solids. Springer Materials, 1989.
- [62] M. Rohlfing, P. Krüger, and J. Pollmann. Quasiparticle band-structure calculations for C, Si, Ge, GaAs, and SiC using Gaussian-orbital basis sets. *Phys. Rev. B*, 48:17791, 1993.
- [63] E. L. Shirley, X. Zhu, and S. G. Louie. Core polarization in solids: Formulation and application to semiconductors. *Phys. Rev. B*, 56:6648, 1997.

- [64] D. R. Mašović, F. R. Vukajlović, and S. Zeković. Local-pseudopotential calculation for optical properties and photoemission valence-band spectrum of silicon. *J. Phys. C: Solid State Phys.*, 16:6731, 1983.
- [65] D. E. Aspnes. Schottky-barrier electroreflectance of Ge: Nondegenerate and degenerate critical points. *Phys. Rev. B*, 12:2297, 1975.
- [66] R. W. Godby, M. Schlüter, and L. J. Sham. Self-energy operators and exchange-correlation potentials in semiconductors. *Phys. Rev. B*, 37:10159, 1988.
- [67] H. J. Lee, L. Y. Juravel, and J. C. Woolley. Electron transport and band structure of $\text{Ga}_{1-x}\text{Al}_x\text{As}$ alloys. *Phys. Rev. B*, 21:659, 1980.
- [68] P. Carrier and S.-H. Wei. Calculated spin-orbit splitting of all diamondlike and zinc-blende semiconductors: Effects of $p_{1/2}$ local orbitals and chemical trends. *Phys. Rev. B*, 70:035212, 2004.
- [69] R. Eppenga, M. F. H. Schuurmans, and H. W. A. M. Rompa. *Ab-initio* determinations of semiconductor spin-orbit splittings from ASW. *J. Phys. Chem. Solids*, 49:1119, 1988.
- [70] E. Matatagui, A. G. Thompson, and M. Cardona. Thermorefectance in Semiconductors. *Phys. Rev.*, 176:950, 1968.
- [71] G. G. Wepfer, T. C. Collins, and R. N. Euwema. Calculated Spin-Orbit Splittings of Some Group IV, III-V and II-VI Semiconductors. *Phys. Rev. B*, 4:1296, 1971.
- [72] B. Monemar. Fundamental Energy Gaps of AlAs and AlP from Photoluminescence Excitation Spectra. *Phys. Rev. B*, 8:5711, 1973.
- [73] S. Massidda, A. Continenza, and A. J. Freeman. Structure and electronic properties of narrow-band-gap semiconductors: InP, InAs, and InSb. *Phys. Rev. B*, 41:12079, 1990.
- [74] Z. Zanolli, F. Fuchs, J. Furthmüller, U. von Barth, and F. Bechstedt. Model GW band structure of InAs and GaAs in the wurtzite phase. *Phys. Rev. B*, 75:245121, 2007.
- [75] J. Crain, S. J. Clark, G. J. Ackland, M. C. Payne, V. Milman, P. D. Hatton, and B. J. Reid. Theoretical study of high-density phases of covalent semiconductors. I. *Ab initio* treatment. *Phys. Rev. B*, 49:5329, 1994.
- [76] B. G. Pfrommer, M. Côté, S. G. Louie, and M. L. Cohen. *Ab initio* study of silicon in the R8 phase. *Phys. Rev. B*, 56:6662, 1997.

- [77] S. J. Clark, G. J. Ackland, and J. Crain. Theoretical study of high-density phases of covalent semiconductors. II. Empirical treatment. *Phys. Rev. B*, 49:5341, 1994.
- [78] R. J. Needs and A. Mujica. First-principles pseudopotential study of the structural phases of silicon. *Phys. Rev. B*, 51:9652, 1995.
- [79] A. Mujica, A. Rubio, A. Muñoz, and R. J. Needs. High-pressure phases of group-IV, III-V, and II-VI compounds. *Rev. Mod. Phys.*, 75:863, 2003.
- [80] G. J. Ackland. High-pressure phases of group IV and III-V semiconductors. *Rep. Prog. Phys.*, 64:483, 2001.
- [81] J. M. Ziman. *Principles of the Theory of Solids*. Cambridge University Press, 2nd edition, 1972. p.228.
- [82] M. I. McMahon, R. J. Nelmes, N. G. Wright, and D. R. Allan. Pressure dependence of the *Imma* phase of silicon. *Phys. Rev. B*, 50:739, 1994.
- [83] A. Mujica, S. Radescu, A. Muñoz, and R. J. Needs. Comparative Study of Novel Structures in Silicon and Germanium. *Phys. Status Solidi B*, 223:379, 2001.
- [84] J. Crain, G. J. Ackland, J. R. Maclean, R. O. Piltz, P. D. Hatton, and G. S. Pawley. Reversible pressure-induced structural transitions between metastable phases of silicon. *Phys. Rev. B*, 50:13043, 1994.
- [85] J. D. Joannopoulos and M. L. Cohen. Theory of Short-Range Order and Disorder in Tetrahedrally Bonded Semiconductors. In H. Ehrenreich, F. Seitz, and D. Turnbull, editors, *Solid State Physics*. Academic Press, 1976. Vol. 31.
- [86] R. O. Piltz, J. R. Maclean, S. J. Clark, G. J. Ackland, P. D. Hatton, and J. Crain. Structure and properties of silicon XII: A complex tetrahedrally bonded phase. *Phys. Rev. B*, 52:4072, 1995.
- [87] C. Raffy, J. Furthmüller, and F. Bechstedt. Properties of hexagonal polytypes of group-IV elements from first-principles calculations. *Phys. Rev. B*, 66:75201, 2002.
- [88] S. Q. Wang and H. Q. Ye. First-principles study on the lonsdaleite phases of C, Si, and Ge. *J. Phys.: Condens. Matter*, 15:L197, 2003.
- [89] R. J. Nelmes, M. I. McMahon, N. G. Wright, D. R. Allan, and J. S. Loveday. Stability and crystal structure of BC8 germanium. *Phys. Rev. B*, 48:9883, 1993.
- [90] B. G. Pfrommer, M. Côté, S. G. Louie, and M. L. Cohen. Relaxation of Crystals with a Quasi-Newton Method. *J. Comput. Phys.*, 131:233, 1997.

- [91] A. Canning, D. Raczkowski, B. Pfrommer, J. Li, Y. Yoon, D. Roundy, F. Mauri, M. Côté, Jonathan Yates, Chris Packard, and Peter Haynes. <http://www.neresc.gov/projects/paratec/>.
- [92] J. R. Yates, X. Wang, D. Vanderbilt, and I. Souza. Spectral and Fermi surface properties from Wannier interpolation. *Phys. Rev. B*, 75:195121, 2007.
- [93] B. Pfrommer, D. Roundy, Y. Yoon, and D. Raczkowski. *Paratec User Guide*, 2003.
- [94] X. Zhu, S. Fahy, and S. G. Louie. *Ab initio* calculation of pressure coefficients of band gaps of silicon: Comparison of the local-density approximation and quasiparticle results. *Phys. Rev. B*, 39:7840, 1989.
- [95] M. Dragoman, A. Cismaru, H. Hartnagel, and R. Plana. Reversible metal-semiconductor transitions for microwave switching applications. *Appl. Phys. Lett.*, 88:73503, 2006.
- [96] M. Green. Intrinsic concentration, effective densities of states, and effective mass in silicon. *J. Appl. Phys.*, 67:2944, 1990.
- [97] J. M. Besson, E. H. Mokhtari, J. Gonzalez, and G. Weill. Electrical properties of semimetallic silicon III and semiconductive silicon IV at ambient pressure. *Phys. Rev. Lett.*, 59:473, 1987.
- [98] J. D. Joannopoulos and M. L. Cohen. Electronic Properties of Complex Crystalline and Amorphous Phases of Ge and Si. I. Density of States and Band Structures. *Phys. Rev. B.*, 7:2644, 1973.
- [99] J. S. Kasper and S. M. Richards. The crystal structures of new forms of silicon and germanium. *Acta Crystallogr.*, 17:752, 1964.
- [100] Y. Zhang, Z. Iqbal, S. Vijayalakshmi, and H. Grebel. Stable hexagonal-wurtzite silicon phase by laser ablation. *Appl. Phys. Lett.*, 75:2758, 1999.
- [101] R. Alben, D. Weaire, J. E. Smith, and M. H. Brodsky. Vibrational properties of amorphous Si and Ge. *Phys. Rev. B*, 11:2271, 1975.
- [102] E. Bourgeois and X. Blase. Superconductivity in doped cubic silicon: An *ab initio* study. *Appl. Phys. Lett.*, 90:142511, 2007.
- [103] W. L. McMillan. Transition Temperature of Strong-Coupled Superconductors. *Phys. Rev.*, 167:331, 1968.
- [104] L. Boeri, J. Kortus, and O. K. Andersen. Three-Dimensional MgB₂-Type Superconductivity. *Phys. Rev. Lett.*, 93:237002, 2004.

- [105] B. D. Malone, J. D. Sau, and M. L. Cohen. *Ab initio* survey of the electronic structure of tetrahedrally bonded phases of silicon. *Phys. Rev. B*, 78:035210, 2008.
- [106] S. Ruffell, J. E. Bradby, N. Fujisawa, and J. S. Williams. Identification of nanoindentation-induced phase changes in silicon by *in situ* electrical characterization. *J. Appl. Phys.*, 101:083531, 2007.
- [107] D. Ge, V. Domnich, and Y. Gogotsi. High-resolution transmission electron microscopy study of metastable silicon phases produced by nanoindentation. *J. Appl. Phys.*, 93:2418, 2003.
- [108] J. E. Bradby, J. S. Williams, J. Wong-Leung, M. V. Swain, and P. Munroe. Transmission electron microscopy observation of deformation microstructure under spherical indentation in silicon. *Appl. Phys. Lett.*, 77:3749, 2000.
- [109] D. Brust, M. L. Cohen, and J. C. Phillips. Reflectance and Photoemission from Si. *Phys. Rev. Lett.*, 9:389, 1962.
- [110] M. Rohlfing and S. G. Louie. Electron-Hole Excitations in Semiconductors and Insulators. *Phys. Rev. Lett*, 81:2312, 1998.
- [111] G. Strinati. Effects of dynamical screening on resonances at inner-shell thresholds in semiconductors. *Phys. Rev. B*, 29:5718, 1984.
- [112] L. X. Benedict and E. L. Shirley. *Ab initio* calculation of $\epsilon_2(\omega)$ including the electron-hole interaction: Application to GaN and CaF₂. *Phys. Rev. B*, 59:5441, 1999.
- [113] W. Hanke and L. J. Sham. Many-Particle effects in the Optical Excitations of a Semiconductor. *Phys. Rev. Lett.*, 43:387, 1979.
- [114] S. Albrecht, L. Reining, R. Del Sole, and G. Onida. *Ab Initio* Calculation of Excitonic Effects in the Optical Spectra of Semiconductors. *Phys. Rev. Lett.*, 80:4510, 1998.
- [115] L. X. Benedict, E. L. Shirley, and R. B. Bohn. Optical Absorption of Insulators and the Electron-Hole Interaction: An *Ab Initio* Calculation. *Phys. Rev. Lett.*, 80:4514, 1998.
- [116] M. A. Green. Crystalline and thin-film silicon solar cells: state of the art and future potential. *Solar Energy*, 74:181, 2003.
- [117] G. E. Jellison, M. F. Chisholm, and S. M. Gorbatskin. Optical functions of chemical vapor deposited thin-film silicon determined by spectroscopic ellipsometry. *Appl. Phys. Lett.*, 62:3348, 1993.

- [118] ASTM Standard G173, Standard Tables for Reference Solar Spectral Irradiances: Direct Normal and Hemispherical on 37 degree Tilted Surface, 2006.
- [119] M. G. Mauk. Silicon solar cells: Physical metallurgy principles. *JOM*, 55:38, 2003.
- [120] Y. Fujimoto, T. Koretsune, S. Saito, Y. Miyake, and A. Oshiyama. A new crystalline phase of four-fold coordinated silicon and germanium. *New J. Phys.*, 10:083001, 2008.
- [121] Y. Omata, Y. Yamagami, K. Tadano, T. Miyake, and S. Saito. Nanotube nanoscience: A molecular-dynamics study. *Physica E*, 29:454, 2005.
- [122] B. D. Malone, J. D. Sau, and M. L. Cohen. *Ab initio* study of the optical properties of Si-XII. *Phys. Rev. B*, 78:161202, 2008.
- [123] G. Onida, L. Reining, R. W. Godby, R. Del Sole, and W. Andreoni. *Ab initio* calculations of the quasiparticle and absorption spectra of clusters: the sodium tetramer. *Phys. Rev. Lett.*, 75:818, 1995.
- [124] S. Albrecht, L. Reining, R. Del Sole, and G. Onida. *Ab initio* calculation of excitonic effects in the optical spectra of semiconductors. *Phys. Rev. Lett.*, 80:4510, 1998.
- [125] L. X. Benedict, E. L. Shirley, and R. B. Bohn. Optical Absorption of Insulators and the Electron-Hole Interaction: An *ab initio* calculation. *Phys. Rev. Lett.*, 80:4514, 1998.
- [126] J. W. van der Horst, P.A. Bobbert, M. A. J. Michels, G. Brooks, and P. J. Kelly. *Ab initio* calculation of the electronic and optical excitations in polythiophene: effects of intra- and interchain screening. *Phys. Rev. Lett.*, 83:4413, 1999.
- [127] T. Kotani and M. van Schilfgaarde. All-electron GW approximation with the mixed basis expansion based on the full-potential LMTO method. *Solid State Commun.*, 121:461, 2002.
- [128] S. V. Faleev, M. van Schilfgaarde, and T. Kotani. All-Electron Self-Consistent GW Approximation: Application to Si, MnO, and NiO. *Phys. Rev. Lett.*, 93:126406, 2004.
- [129] G. E. Jellison, Jr. Optical functions of silicon determined by two-channel polarization modulation ellipsometry. *Opt. Mater.*, 1:41, 1992.
- [130] F. Dimroth and S. Kurtz. High-Efficiency Multijunction Solar Cells. *MRS Bull.*, 32:230, 2007.

- [131] D. E. Aspnes and A. A. Studna. Dielectric functions and optical parameters of Si, Ge, GaP, GaAs, GaSb, InP, InAs, and InSb from 1.5 to 6.0 eV. *Phys. Rev. B*, 27:985, 1983.
- [132] D. Straub, L. Ley, and F. J. Himpsel. Conduction-Band and Surface-State Critical Points in Si: An Inverse-Photoemission Study. *Phys. Rev. Lett.*, 54:142, 1985.
- [133] F. J. Ribeiro and M. L. Cohen. Theoretical prediction of the high-pressure phase Ge-Cmca. *Phys. Rev. B*, 62:11388, 2000.
- [134] R. J. Nelmes and M. I. McMahon. Chapter 3: Structural Transitions in the Group IV, III-V, and II-VI Semiconductors under Pressure. *Semicond. Semimetals*, 54:145, 1998.
- [135] A. G. Lyapin, V. V. Brazhkin, S. V. Popova, and A. V. Sapelkin. Nonequilibrium Phase Transitions in Diamond and Zincblende Semiconductors under High Pressure. *phys. stat. sol.(b)*, 198:481, 1996.
- [136] B. D. Malone and M. L. Cohen. Prediction of a metastable phase of silicon in the *Ibam* structure. *Phys. Rev. B*, 85:024116, 2012.
- [137] C. H. Bates, F. Dache, and R. Roy. High-Pressure Transitions of Germanium and a New High-Pressure Form of Germanium. *Science*, 147:860, 1965.
- [138] A. Mujica and R. J. Needs. First-principles calculations of the structural properties, stability, and band structure of complex tetrahedral phases of germanium: ST12 and BC8. *Phys. Rev. B*, 48:17010, 1993.
- [139] A. Mujica, S. Radescu, A. Muñoz, and R. J. Needs. Comparative Study of Novel Structures in Silicon and Germanium. *phys. stat. sol. (b)*, 223:379, 2001.
- [140] S. Ruffell, K. Sears, A. P. Knights, J. E. Bradby, and J. S. Williams. Experimental evidence for semiconducting behavior of Si-XII. *Phys. Rev. B*, 83:075316, 2011.
- [141] M. L. Cohen and B. D. Malone. Wave function engineering: Other phases of Si for photovoltaic applications. *J. Appl. Phys.*, 109:102402, 2011.
- [142] F. Coppari, J. C. Chervin, A. Congeduti, M. Lazzeri, A. Polian, E. Principi, and A. Di Cicco. Pressure-induced phase transitions in amorphous and metastable crystalline germanium by Raman scattering, x-ray spectroscopy, and *ab initio* calculations. *Phys. Rev. B*, 80:115213, 2009.
- [143] S. G. Louie, S. Froyen, and M. L. Cohen. Nonlinear ionic pseudopotentials in spin-density-functional calculations. *Phys. Rev. B*, 26:1738, 1982.

- [144] J. Donohue. The Structure of the Elements. In *The Structure of the Elements*. J. Wiley Sons, 1974.
- [145] L. J. Bruner and R. W. Keyes. Electronic Effect in the Elastic Constants of Germanium. *Phys. Rev. Lett.*, 7:55, 1961.
- [146] S.-Q. Xiao and P. Pirouz. On diamond-hexagonal germanium. *J. Mater. Res.*, 7:1406, 1992.
- [147] J. D. Joannopoulos and M. L. Cohen. Electronic Properties of Complex Crystalline and Amorphous Phases of Ge and Si. I. Density of States and Band Structures. *Phys. Rev. B*, 7:2644, 1973.
- [148] As a point of reference, our calculations find the band gap of cubic germanium to be 0.24 eV at Γ for the LDA-relaxed structure.
- [149] B. D. Malone and M. L. Cohen. First principles study of phosphorus and boron substitutional defects in Si-XII. *J. Phys.:Condens. Matter*, 24:055505, 2012.
- [150] D. Olego and M. Cardona. Pressure dependence of Raman phonons of Ge and 3C-SiC. *Phys. Rev. B*, 25:1151, 1982.
- [151] F. Birch. Finite Elastic Strain of Cubic Crystals. *Phys. Rev.*, 71:809, 1947.
- [152] A. Werner, J. A. Sanjurjo, and M. Cardona. X-rays investigation of the $\alpha \rightarrow \beta$ transition in the $\text{Ge}_x\text{Si}_{1-x}$ solid solutions at high pressure. *Solid State Commun.*, 44:155, 1982.
- [153] M. T. Yin and M. L. Cohen. Microscopic theory of the static structural properties and phase transformations of Ge. *Solid State Commun.*, 38:625, 1980.
- [154] E. Wigner. On the Interaction of Electrons in Metals. *Phys. Rev.*, 46:1002, 1934.
- [155] N. Moll, M. Bockstedte, M. Fuchs, E. Pehlke, and M. Scheffler. Application of generalized gradient approximations: The diamond- β -tin phase transition in Si and Ge. *Phys. Rev. B*, 52:2550, 1995.
- [156] R. G. Hennig, A. Wadehra, K. P. Driver, W. D. Parker, C. J. Umrigar, and J. W. Wilkins. Phase transformation in Si from semiconducting diamond to metallic β -Sn phase in QMC and DFT under hydrostatic and anisotropic stress. *Phys. Rev. B*, 82:014101, 2010.
- [157] S. J. Kim, O. K. Quy, L.-S. Chang, E. A. Stach, C. A. Handwerker, and A. Wei. Formation of the ST12 phase in nanocrystalline Ge at ambient pressure. *J. Mater. Chem.*, 20:331, 2010.

- [158] Specifically, we use the pseudopotential Ge.pw91-n-van.UPF provided in the Quantum-ESPRESSO pseudopotential library.
- [159] A. García, C. Elsässer, J. Zhu, S. G. Louie, and M. L. Cohen. Use of gradient-corrected functionals in total-energy calculations for solids. *Phys. Rev. B*, 46:9829, 1992.
- [160] M. Rohlfing, P. Krüger, and J. Pollmann. Quasiparticle calculations of semicore states in Si, Ge, and CdS. *Phys. Rev. B*, 56:7065, 1997.
- [161] B. C. Johnson, B. Haberl, S. Deshmukh, B. D. Malone, M. L. Cohen, J. C. McCallum, J. S. Williams, and J. E. Bradby. High pressure phases formed from amorphous germanium by indentation. (in preparation).
- [162] S. J. Duclos, Y. K. Vohra, and A. L. Ruoff. Experimental study of the crystal stability and equation of state of Si to 248 GPa. *Phys. Rev. B*, 41:12021, 1990.
- [163] M. T. Yin and M. L. Cohen. Theory of static structural properties, crystal stability, and phase transformations: Application to Si and Ge. *Phys. Rev. B*, 26:5668, 1982.
- [164] S. P. Lewis and M. L. Cohen. Theoretical study of high-pressure orthorhombic silicon. *Phys. Rev. B*, 48:16144, 1993.
- [165] M. Hanfland, U. Schwarz, K. Syassen, and K. Takemura. Crystal Structure of the High-Pressure Phase Silicon VI. *Phys. Rev. Lett.*, 82:1197, 1999.
- [166] N. E. Christensen, D. L. Novikov, and M. Methfessel. The intermediate high-pressure phase of silicon. *Solid State Commun.*, 110:615, 1999.
- [167] Y.-X. Zhao, F. Buehler, J. R. Sites, and I. L. Spain. New metastable phases of silicon. *Solid State Commun.*, 59:679, 1986.
- [168] D. Ge, V. Domnich, and Y. Gogotsi. Thermal stability of metastable silicon phases produced by nanoindentation. *J. Appl. Phys.*, 95:2725, 2004.
- [169] S. Ruffell, B. Haberl, J. E. Bradby, and J. S. Williams. Annealing of nanoindentation-induced high pressure crystalline phases created in crystalline and amorphous silicon. *J. Appl. Phys.*, 105:093513, 2009.
- [170] B. D. Malone, S. G. Louie, and M. L. Cohen. Electronic and optical properties of body-centered-tetragonal Si and Ge. *Phys. Rev. B*, 81:115201, 2010.
- [171] G. Samsonidze, S. G. Louie, and M. L. Cohen. Compensation-doped silicon for photovoltaic applications. *Phys. Rev. B*, 84:195201, 2011.

- [172] C. J. Pickard and R. J. Needs. High-Pressure Phases of Silane. *Phys. Rev. Lett.*, 97:045504, 2006.
- [173] G. Kresse and J. Hafner. *Ab initio* molecular dynamics for open-shell transition metals. *Phys. Rev. B*, 48:13115, 1993.
- [174] G. Kresse and J. Furthmüller. Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set. *Phys. Rev. B*, 54:11169, 1996.
- [175] G. Kresse and D. Joubert. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B*, 59:1758, 1999.
- [176] This lowest energy structure was obtained 4 times out of 5000 trial structures.
- [177] J. Peters and B. Krebs. Silicon disulphide and silicon diselenide: a reinvestigation. *Acta. Cryst. B*, 38:1270, 1982.
- [178] M. L. Cohen. Superconductivity in Many-Valley Semiconductors and in Semimetals. *Phys. Rev.*, 134:A511, 1964.
- [179] A. R. Oganov, A. O. Lyakhov, and M. Valle. How Evolutionary Crystal Structure Prediction Works - and Why. *Accts. Chem. Res.*, 44:227, 2011.
- [180] A. O. Lyakhov, A. R. Oganov, and M. Valle. How to predict very large and complex crystal structures. *Comp. Phys. Commun.*, 181:1623, 2010.
- [181] C. W. Glass, A. R. Oganov, and N. Hansen. USPEX-Evolutionary crystal structure prediction. *Comp. Phys. Commun.*, 175:713, 2006.
- [182] J. E. Bradby, J. S. Williams, and M. V. Swain. *In situ* electrical characterization of phase transformations in Si during indentation. *Phys. Rev. B*, 67:085205, 2003.
- [183] S. Ruffell, J. E. Bradby, and J. S. Williams. High pressure crystalline phase formation during nanoindentation: Amorphous versus crystalline silicon. *Appl. Phys. Lett.*, 89:091919, 2006.
- [184] S. Ruffell, K. Sears, J. E. Bradby, and J. S. Williams. Room temperature writing of electrical conductive and insulating zones in silicon by nanoindentation. *Appl. Phys. Lett.*, 98:052105, 2011.
- [185] W. E. Beadle, J. C. C. Tsai, and R. D. Plummer. *Quick Reference Manual for Silicon Integrated Circuit Technology*. Wiley, 1985.
- [186] J. Heyd, G. E. Scuseria, and M. Ernzerhof. Hybrid functionals based on a screened Coulomb potential. *J. Chem. Phys.*, 118:8207, 2003.

- [187] J. Heyd, G. E. Scuseria, and M. Ernzerhof. Erratum: Hybrid functionals based on a screened Coulomb potential. *J. Chem. Phys.*, 124:219906, 2006.
- [188] We take the values of μ_P and μ_B to be the total energies per atom for the black phosphorus phase and the B_{50} phase, respectively.
- [189] C. W. M. Castleton, A. Höglund, and S. Mirbt. Managing the supercell approximation for charged defects in semiconductors: Finite-size scaling, charge correction factors, the band-gap problem, and the *ab initio* dielectric constant. *Phys. Rev. B*, 73:035215, 2006.
- [190] M. L. Cohen, M. Schlüter, J. R. Chelikowsky, and S. G. Louie. Self-consistent pseudopotential method for localized configurations: Molecules. *Phys. Rev. B*, 12:5575, 1975.
- [191] C. Adamo and V. Barone. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.*, 110:6158, 1999.
- [192] A. V. Krukau, O. A. Vydrov, A. F. Izmaylov, and G. E. Scuseria. Influence of the exchange screening parameter on the performance of screened hybrid functionals. *J. Chem. Phys.*, 125:224106, 2006.
- [193] J. P. Perdew, M. Ernzerhof, and K. Burke. Rationale for mixing exact exchange with density functional approximations. *J. Chem. Phys.*, 105:9982, 1996.
- [194] J. Heyd, J. E. Peralta, G. E. Scuseria, and R. L. Martin. Energy band gaps and lattice parameters evaluated with the Heyd-Scuseria-Ernzerhof screened hybrid functional. *J. Chem. Phys.*, 123:174101, 2005.
- [195] E. R. Batista, J. Heyd, R. G. Hennig, B. P. Uberuaga, R. L. Martin, G. E. Scuseria, C. J. Umrigar, and J. W. Wilkins. Comparison of screened hybrid density functional theory to diffusion Monte Carlo in calculations of total energies of silicon phases and defects. *Phys. Rev. B*, 74:121102, 2006.
- [196] W.-K. Leung, R. J. Needs, G. Rajagopal, S. Itoh, and S. Ihara. Calculations of Silicon Self-Interstitial Defects. *Phys. Rev. Lett.*, 83:2351, 1999.
- [197] C. Freysoldt, J. Neugebauer, and C. G. van de Walle. Electrostatic interactions between charged defects in supercells. *Phys. Status Solidi B*, 248:1067, 2011.
- [198] R. Nieminen. Issues in first-principles calculations for defects in semiconductors and oxides. *Modell. Simul. Mater. Sci. Eng.*, 17:084001, 2009.
- [199] G. Makov and M. C. Payne. Periodic boundary conditions in *ab initio* calculations. *Phys. Rev. B*, 51:4014, 1995.

- [200] S. Lany and A. Zunger. Accurate prediction of defect properties in density functional supercell calculations. *Modelling Simul. Mater. Sci. Eng.*, 17:084002, 2009.
- [201] C. W. M. Castleton, A. Höglund, and S. Mirbt. Density functional theory calculations of defect energies using supercells. *Modelling Simul. Mater. Sci. Eng.*, 17:084003, 2009.
- [202] N. D. M. Hine, K. Frensch, W. M. C. Foulkes, and M. W. Finnis. Supercell size scaling of density functional theory formation energies of charged defects. *Phys. Rev. B*, 79:024112, 2009.
- [203] The complete structural description for Si-XII calculated with HSE is $a=5.73\text{\AA}$, $\alpha=109.84^\circ$, $u=0.2867$, $x=0.4782$, $y=-0.0199$, and $z=0.2826$. The corresponding parameters for the LDA calculation are $a=5.70\text{\AA}$, $\alpha=109.87^\circ$, $u=0.2847$, $x=0.4756$, $y=-0.0222$, and $z=0.2792$.
- [204] The experimental band gap of Si-XII has not yet been measured and so we take our prior GW results as the best guess for this value.
- [205] U. Gerstmann, P. Deák, R. Rurali, B. Aradi, Th. Frauenheim, and H. Overhof. Charge corrections for supercell calculations of defects in semiconductors. *Physica B*, 340:190, 2003.
- [206] S. Ossicini, E. Degoli, F. Iori, E. Luppi, R. Magri, G. Cantele, F. Trani, and D. Ninno. Simultaneously B- and P-doped silicon nanoclusters: Formation energies and electronic properties. *Appl. Phys. Lett.*, 87:173120, 2005.

Appendix A

Convergence data for GW calculations in Chapter 2

Contained in this Appendix is all of the convergence tests that were done in association with the work of Chapter 2. For each phase we have carried out convergence tests regarding the number of bands in the Coulomb-hole part of the self-energy (Eq. 2.2), the number of bands in the polarization (Eq. 2.7), as well as the size of the dielectric matrix. The steps taken in evaluating the convergence is given in detail in Section 2.2.2.

A.1 Si

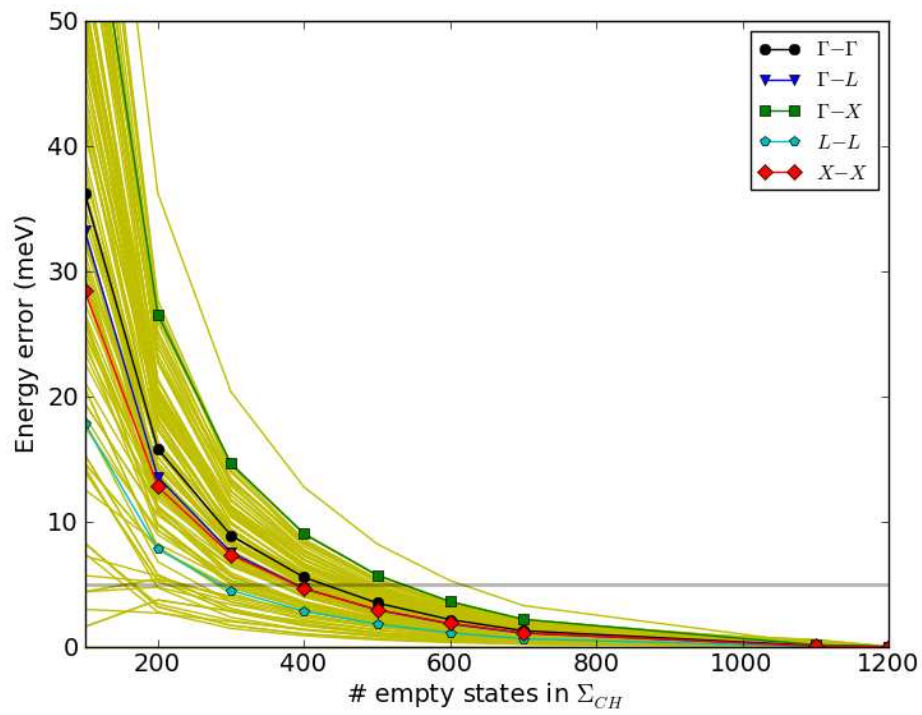


Figure A.1: Testing of the number of bands needed for the Σ_{CH} summation in Si. From this plot it is determined that 600 bands is enough to obtain an error less than 5 meV.

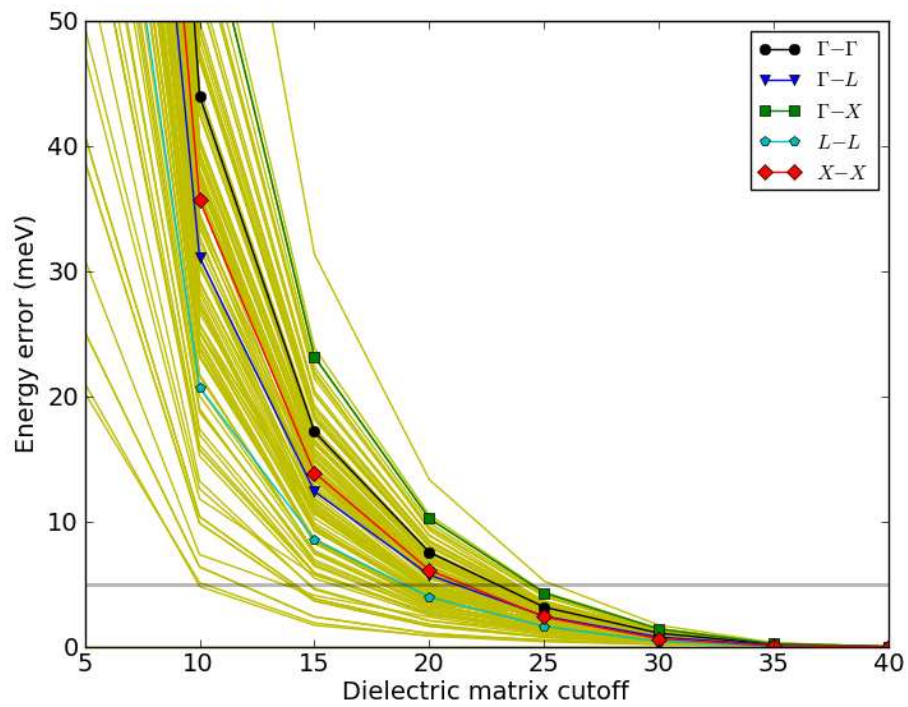


Figure A.2: Testing of the energy cutoff for the dielectric matrix in Si. From this plot it can be determined that 25 Ry is enough to converge the resulting quasiparticle energies to less than 5 meV.

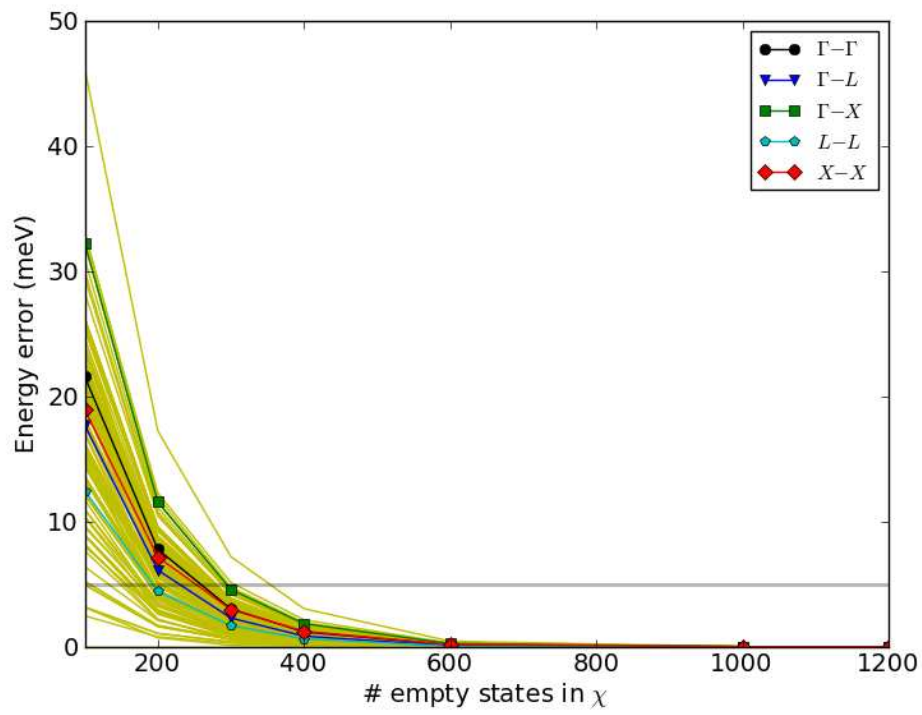


Figure A.3: Testing of the number of bands needed for the χ summation in Si. From this plot it is determined that 300 bands is sufficient to obtain an error less than 5 meV.

A.2 Ge

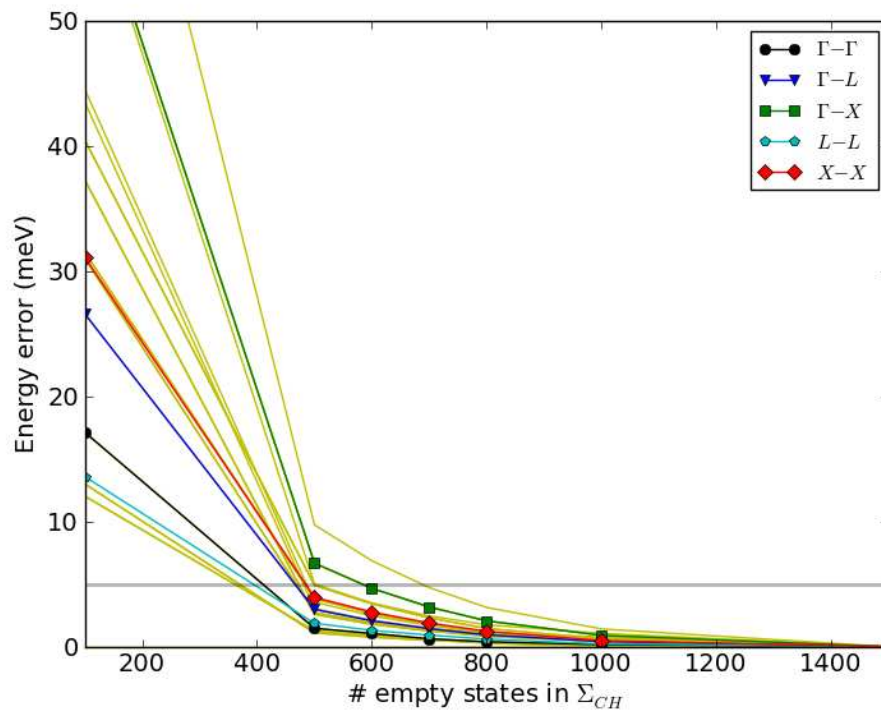


Figure A.4: Testing of the number of bands needed for the Σ_{CH} summation in Ge. From this plot it is determined that 700 bands is enough to obtain an error less than 5 meV.

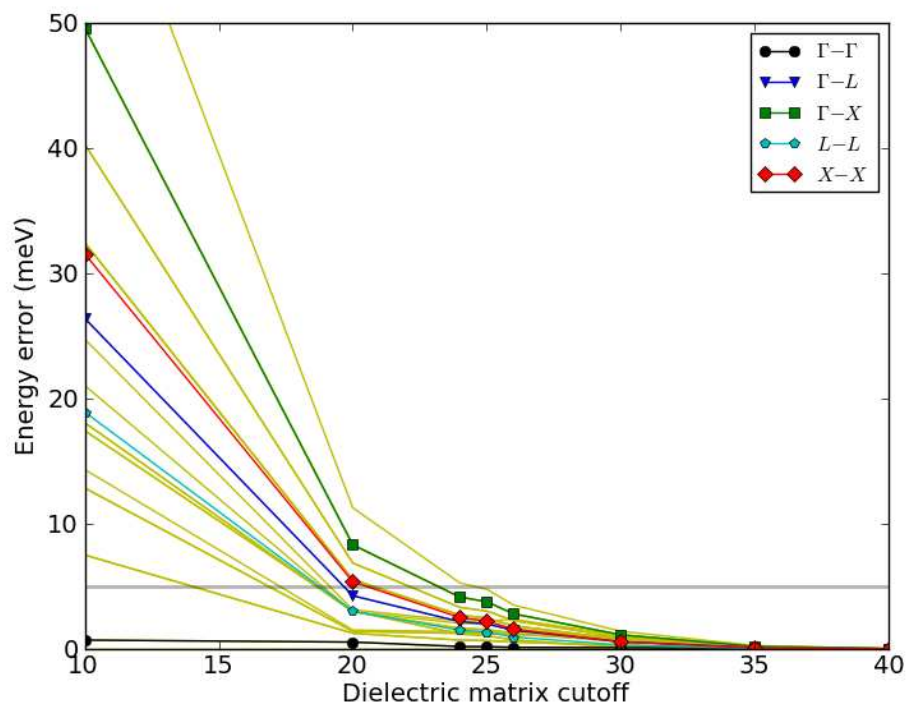


Figure A.5: Testing of the energy cutoff for the dielectric matrix in Ge. From this plot it can be determined that 25 Ry is enough to converge the resulting quasiparticle energies to less than 5 meV.

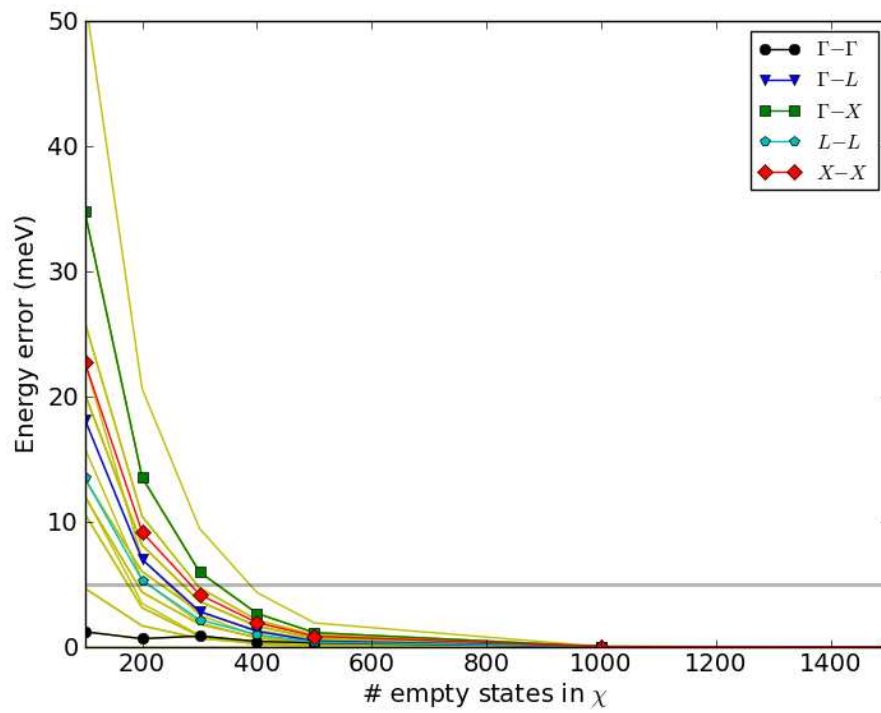


Figure A.6: Testing of the number of bands needed for the χ summation in Ge. From this plot it is determined that 400 bands is sufficient to obtain an error less than 5 meV.

A.3 AIP

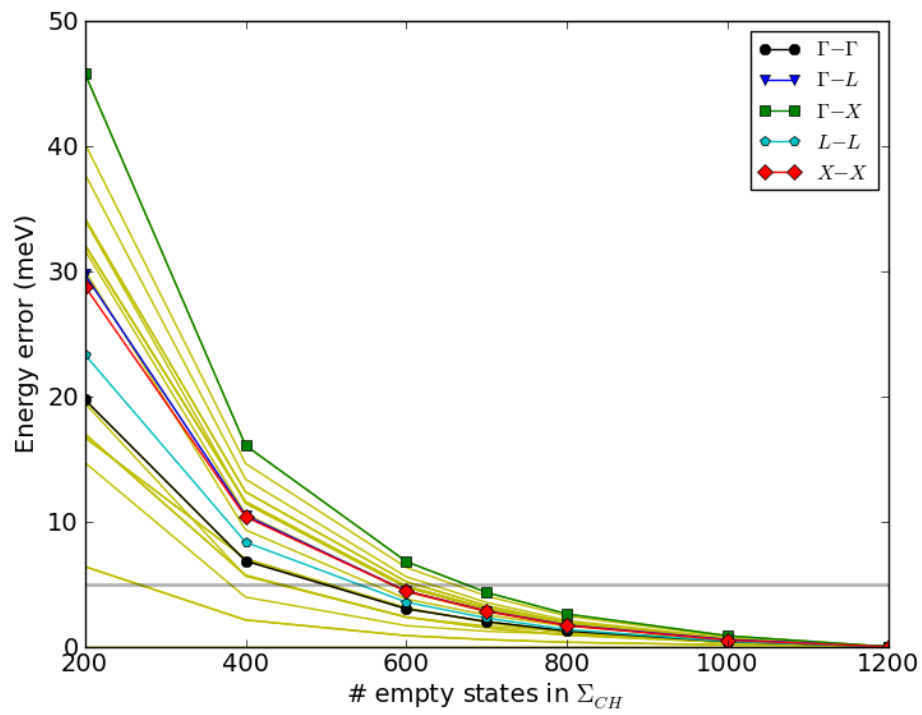


Figure A.7: Testing of the number of bands needed for the Σ_{CH} summation in AIP. From this plot it is determined that 700 bands is enough to obtain an error less than 5 meV.

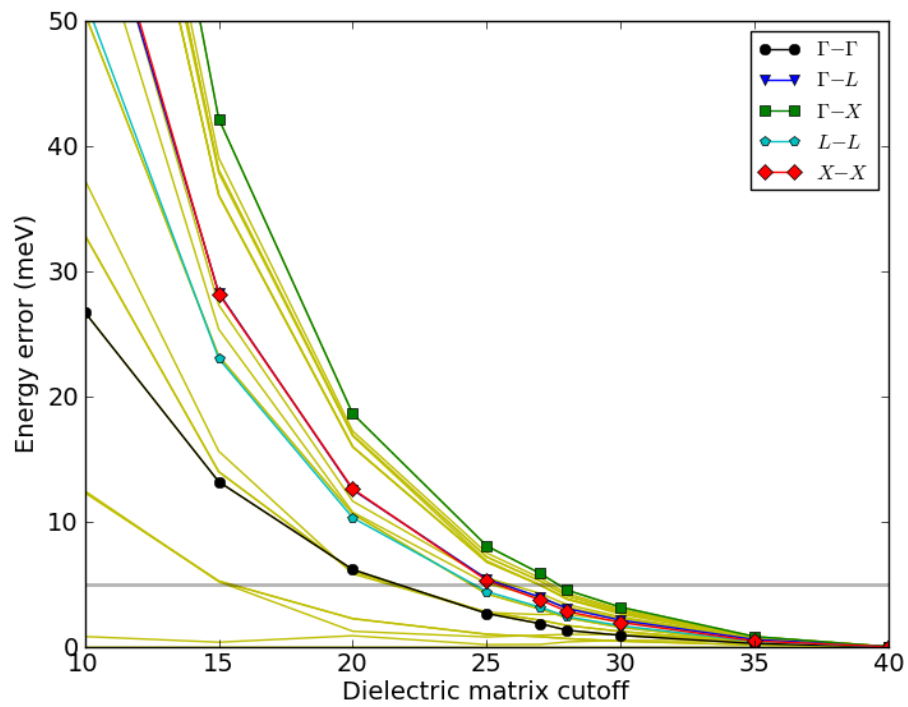


Figure A.8: Testing of the energy cutoff for the dielectric matrix in AIP. From this plot it can be determined that 28 Ry is enough to converge the resulting quasiparticle energies to less than 5 meV.

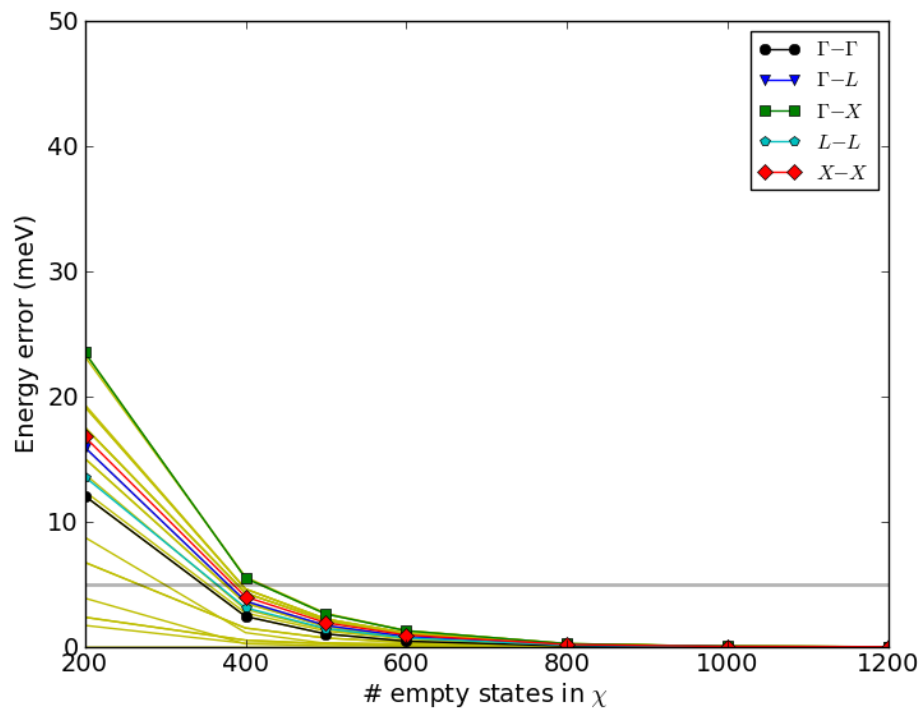


Figure A.9: Testing of the number of bands needed for the χ summation in AIP. From this plot it is determined that 500 bands is sufficient to obtain an error less than 5 meV.

A.4 AlAs

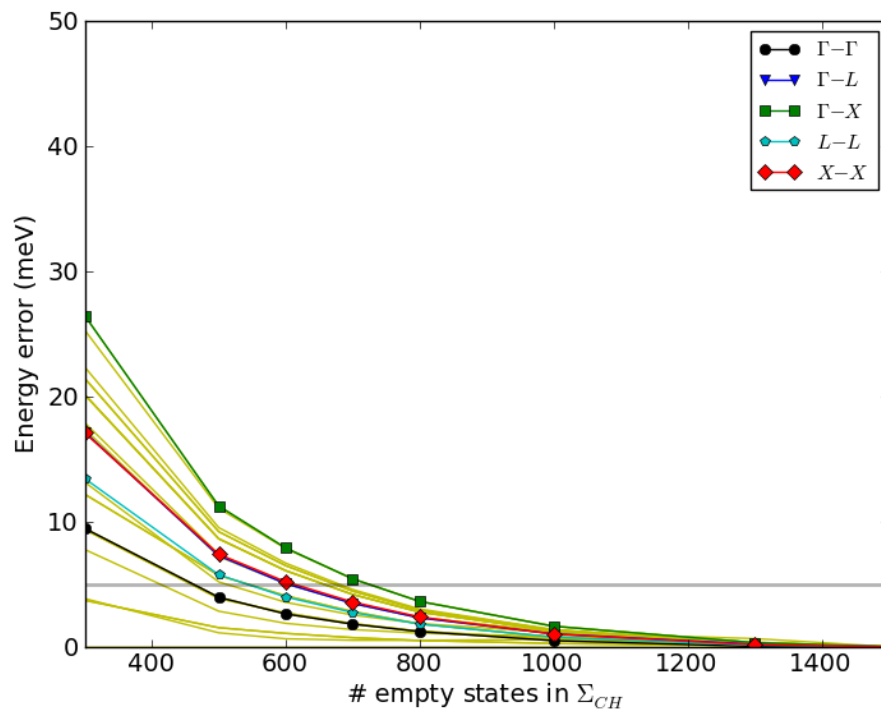


Figure A.10: Testing of the number of bands needed for the Σ_{CH} summation in AIs. From this plot it is determined that 800 bands is enough to obtain an error less than 5 meV.

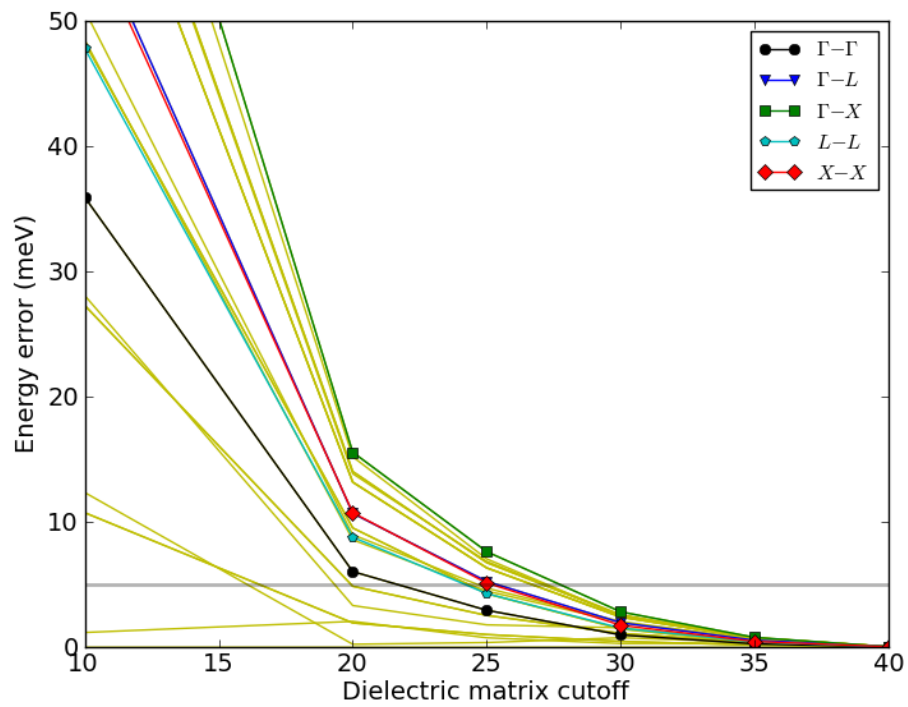


Figure A.11: Testing of the energy cutoff for the dielectric matrix in AlAs. From this plot it can be determined that 30 Ry is enough to converge the resulting quasiparticle energies to less than 5 meV.

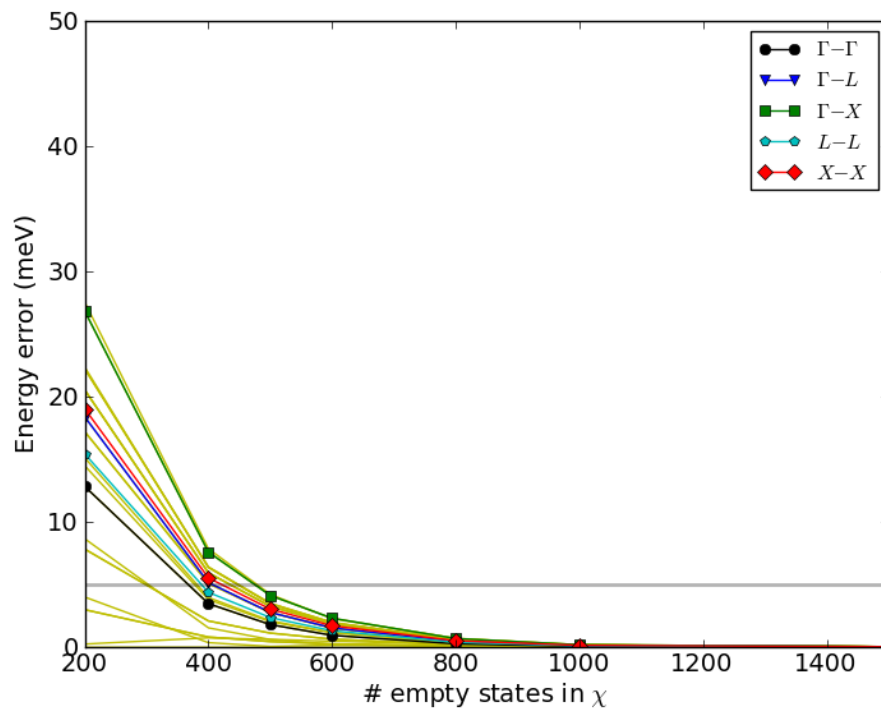


Figure A.12: Testing of the number of bands needed for the χ summation in AIs. From this plot it is determined that 500 bands is sufficient to obtain an error less than 5 meV.

A.5 AlSb

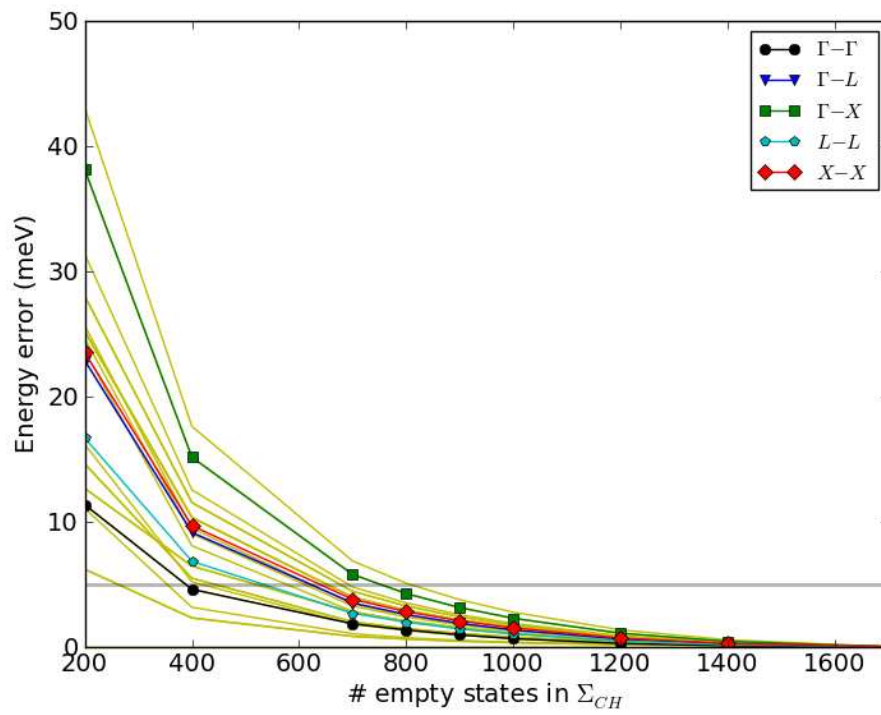


Figure A.13: Testing of the number of bands needed for the Σ_{CH} summation in AlSb. From this plot it is determined that 800 bands is enough to obtain an error less than 5 meV.

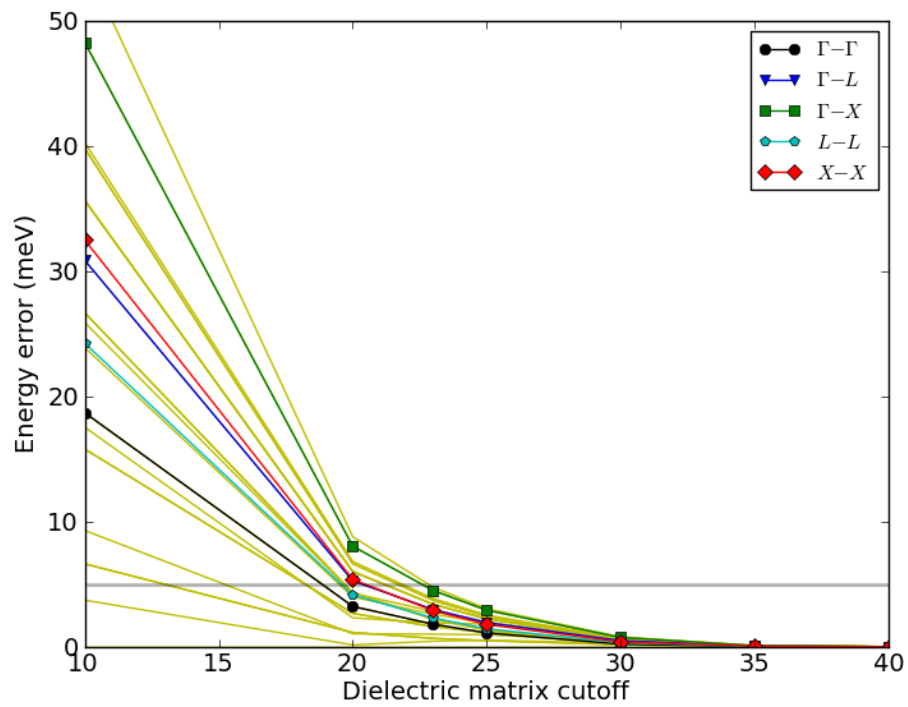


Figure A.14: Testing of the energy cutoff for the dielectric matrix in AlSb. From this plot it can be determined that 23 Ry is enough to converge the resulting quasiparticle energies to less than 5 meV.

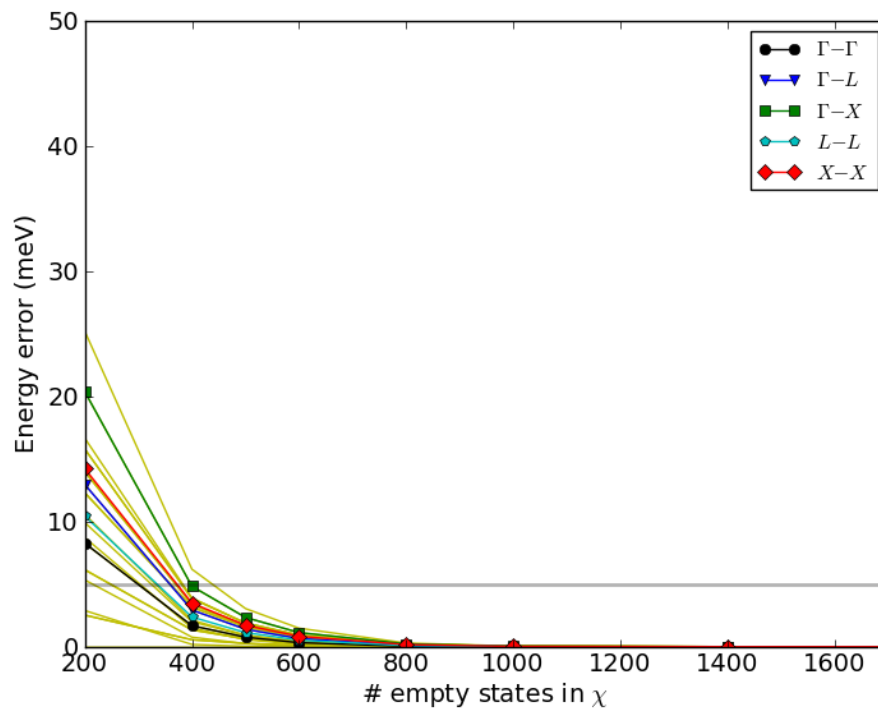


Figure A.15: Testing of the number of bands needed for the χ summation in AlSb. From this plot it is determined that 500 bands is sufficient to obtain an error less than 5 meV.

A.6 InP

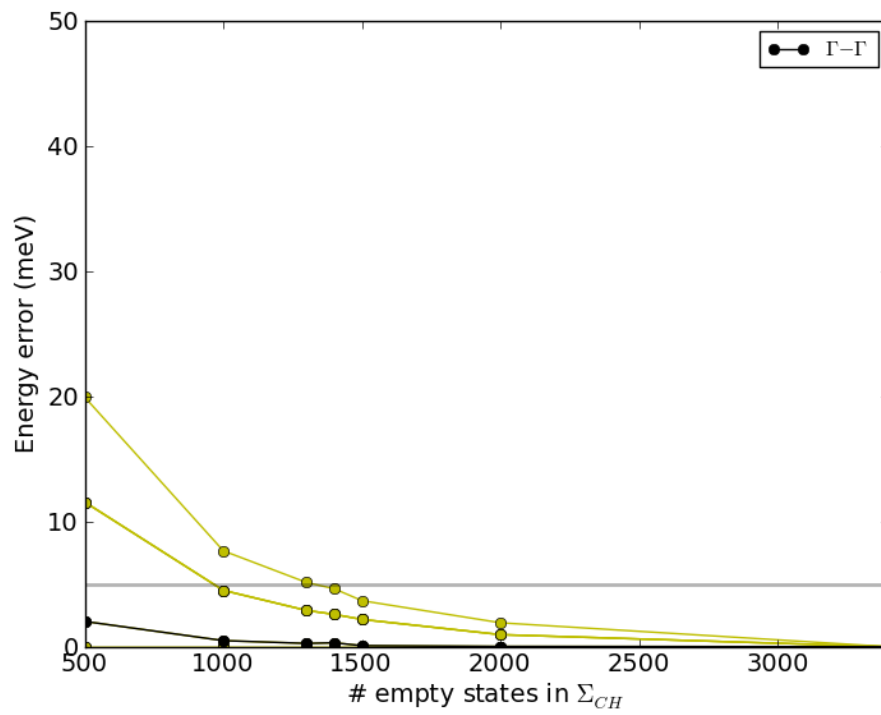


Figure A.16: Testing of the number of bands needed for the Σ_{CH} summation in InP. From this plot it is determined that 1400 bands is enough to obtain an error less than 5 meV.

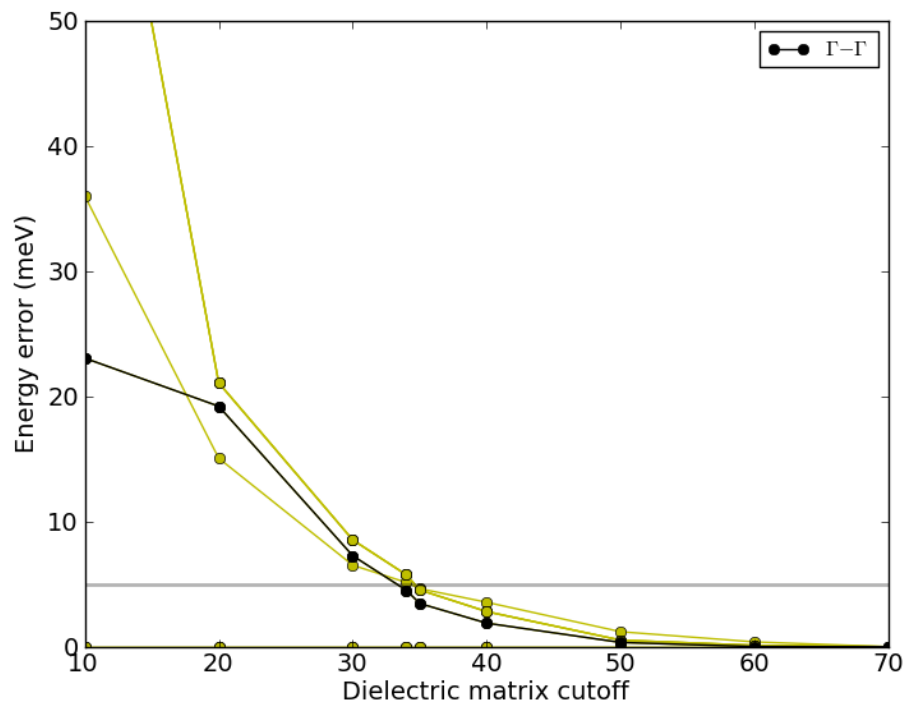


Figure A.17: Testing of the energy cutoff for the dielectric matrix in InP. From this plot it can be determined that 35 Ry is enough to converge the resulting quasiparticle energies to less than 5 meV.

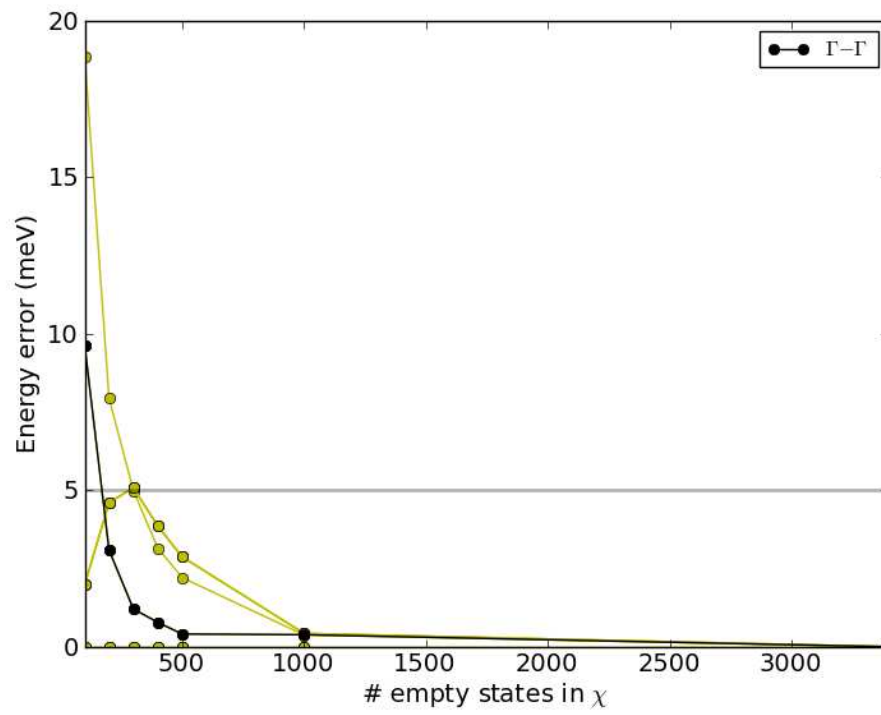


Figure A.18: Testing of the number of bands needed for the χ summation in InP. From this plot it is determined that 400 bands is sufficient to obtain an error less than 5 meV.

A.7 InAs

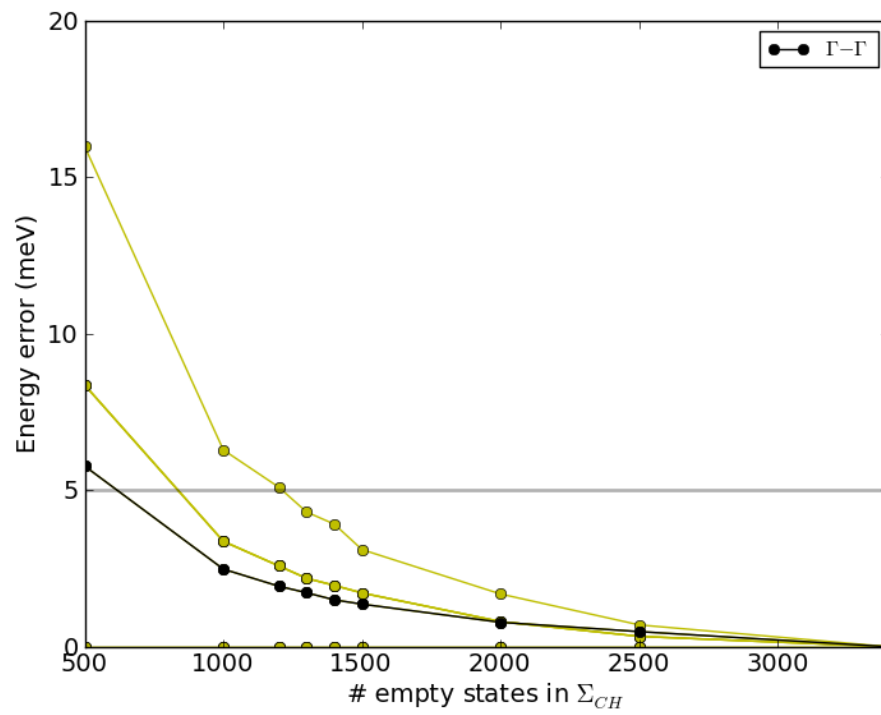


Figure A.19: Testing of the number of bands needed for the Σ_{CH} summation in InAs. From this plot it is determined that 1300 bands is enough to obtain an error less than 5 meV.

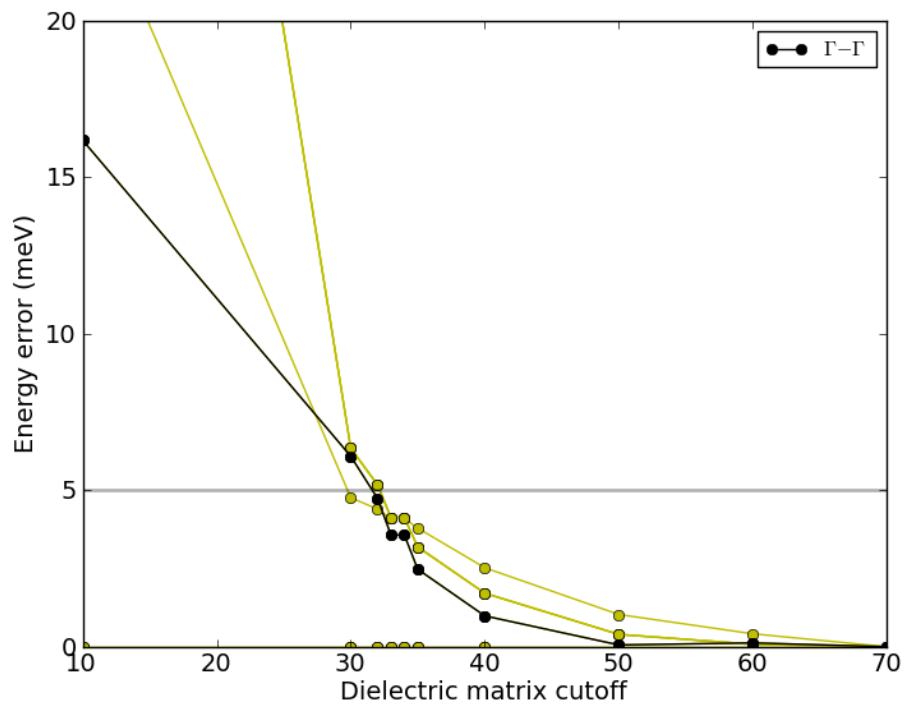


Figure A.20: Testing of the energy cutoff for the dielectric matrix in InAs. From this plot it can be determined that 33 Ry is enough to converge the resulting quasiparticle energies to less than 5 meV.

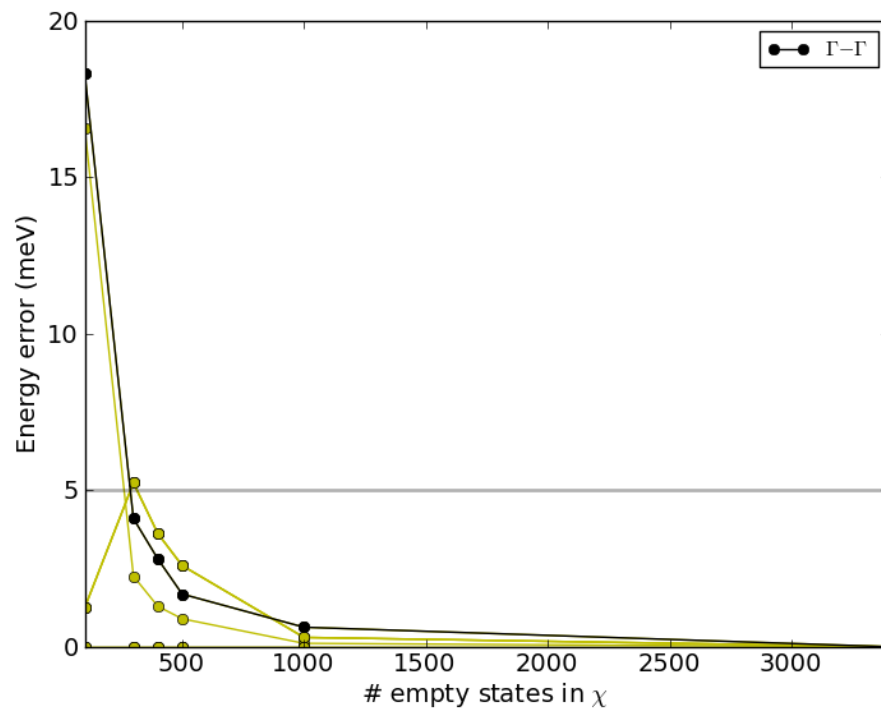


Figure A.21: Testing of the number of bands needed for the χ summation in InAs. From this plot it is determined that 400 bands is sufficient to obtain an error less than 5 meV.

A.8 InSb

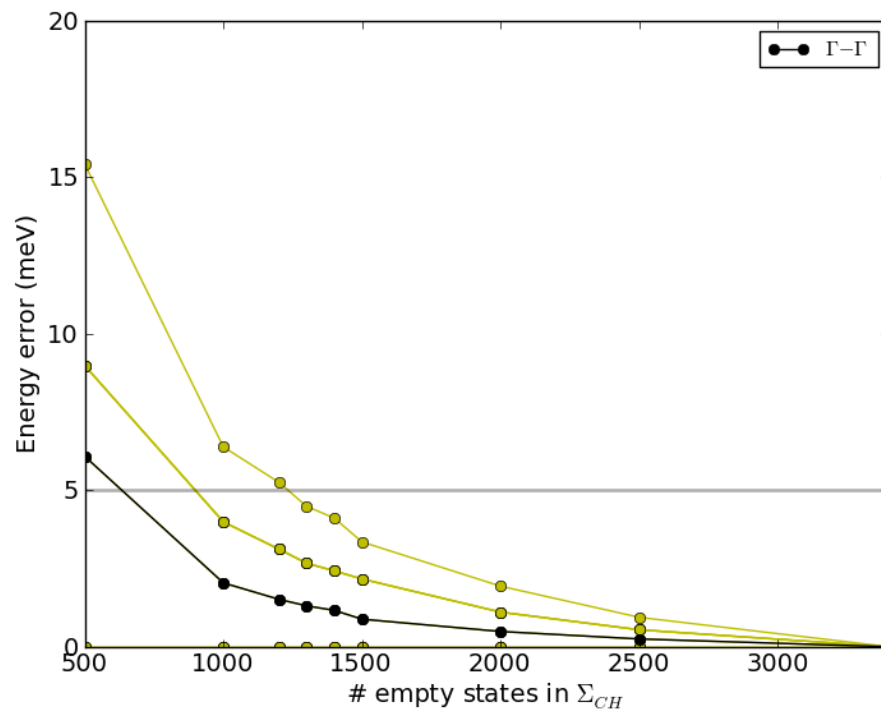


Figure A.22: Testing of the number of bands needed for the Σ_{CH} summation in InSB. From this plot it is determined that 1300 bands is enough to obtain an error less than 5 meV.

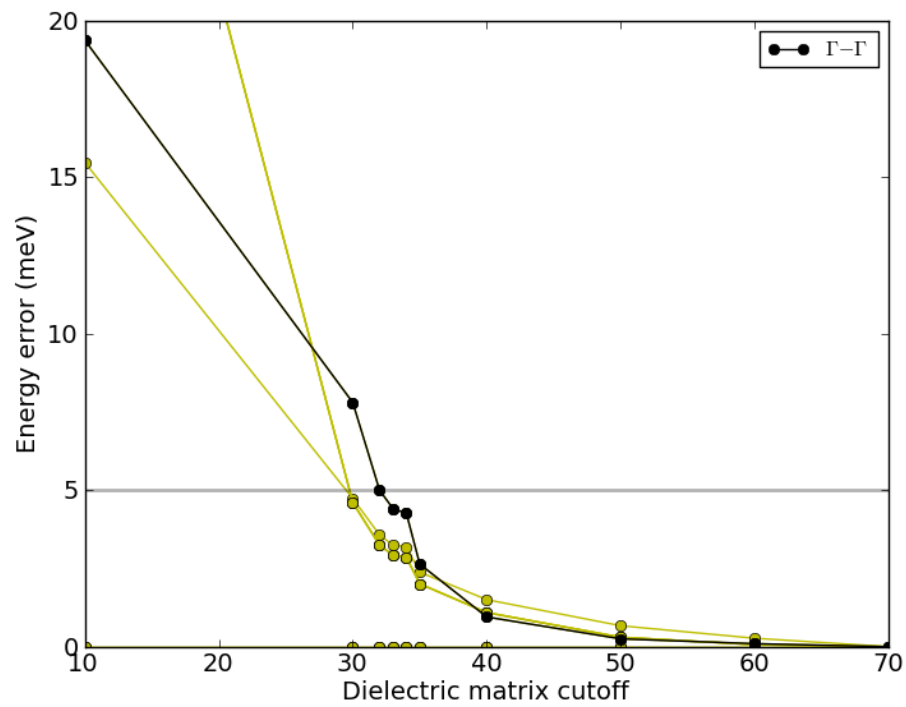


Figure A.23: Testing of the energy cutoff for the dielectric matrix in InSb. From this plot it can be determined that 33 Ry is enough to converge the resulting quasiparticle energies to less than 5 meV.

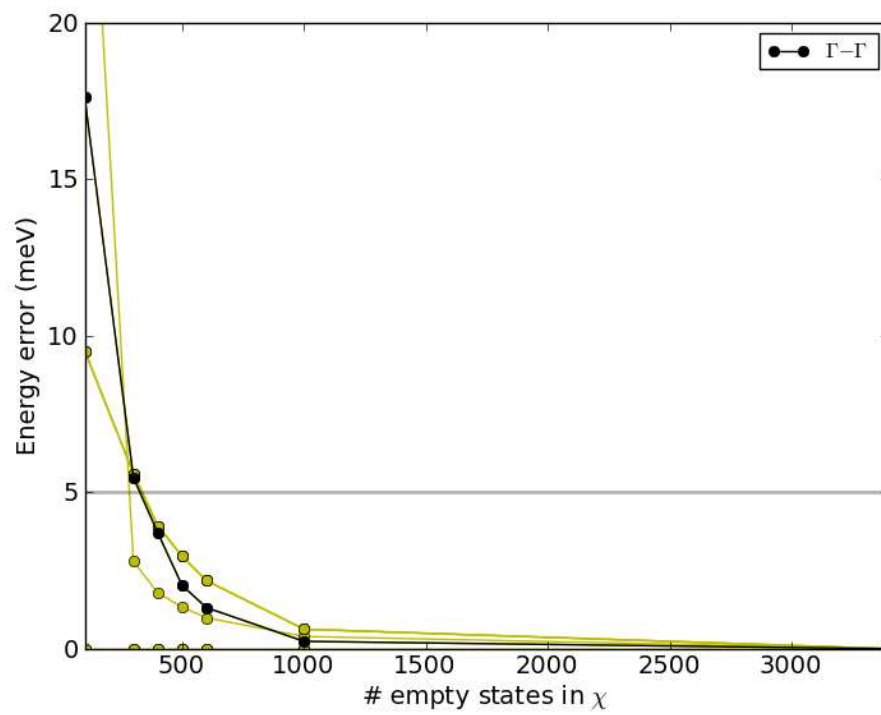


Figure A.24: Testing of the number of bands needed for the χ summation in InSb. From this plot it is determined that 400 bands is sufficient to obtain an error less than 5 meV.

A.9 GaP

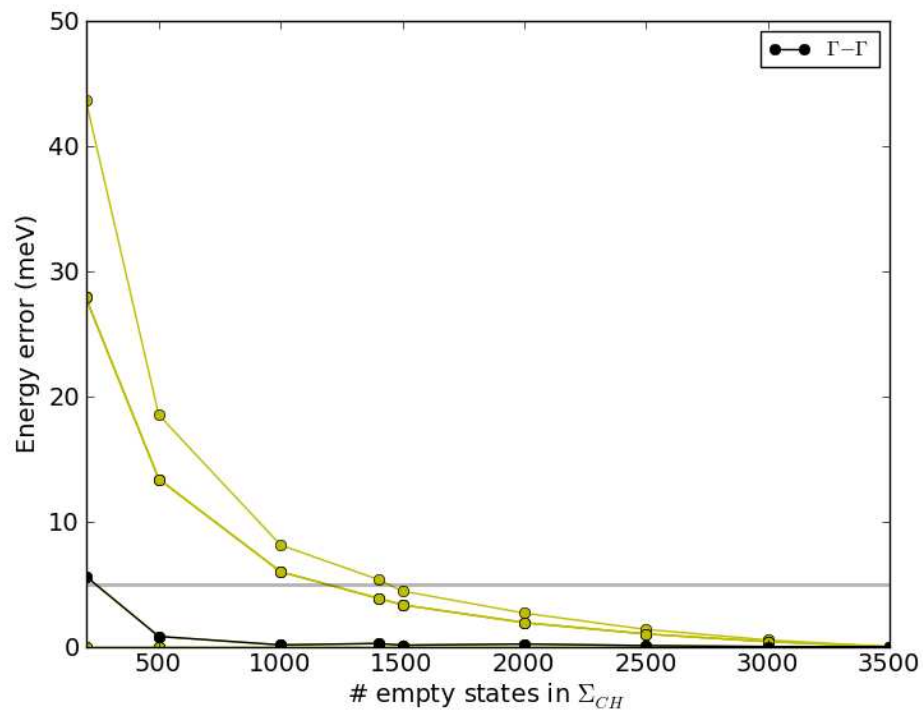


Figure A.25: Testing of the number of bands needed for the Σ_{CH} summation in GaP. From this plot it is determined that 1500 bands is enough to obtain an error less than 5 meV.

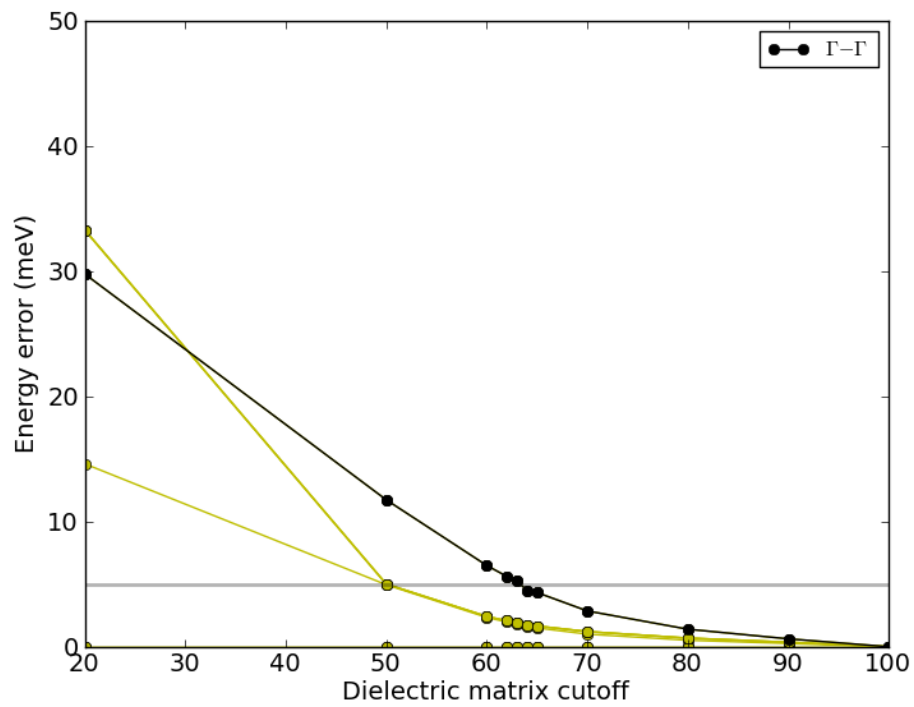


Figure A.26: Testing of the energy cutoff for the dielectric matrix in GaP. From this plot it can be determined that 64 Ry is enough to converge the resulting quasiparticle energies to less than 5 meV.

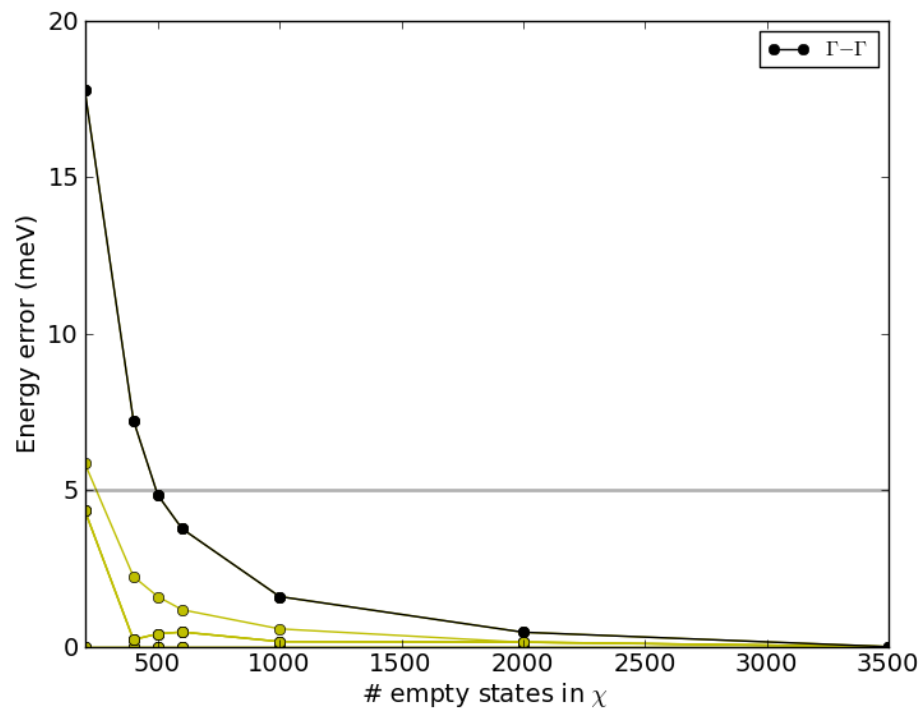


Figure A.27: Testing of the number of bands needed for the χ summation in GaP. From this plot it is determined that 500 bands is sufficient to obtain an error less than 5 meV.

A.10 GaAs

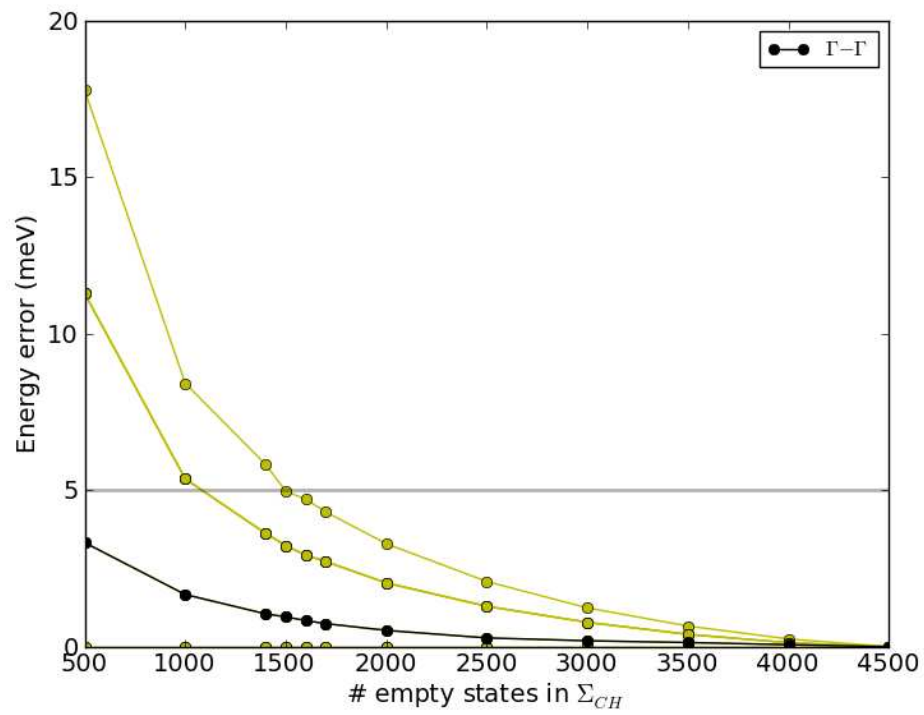


Figure A.28: Testing of the number of bands needed for the Σ_{CH} summation in GaAs. From this plot it is determined that 1600 bands is enough to obtain an error less than 5 meV.

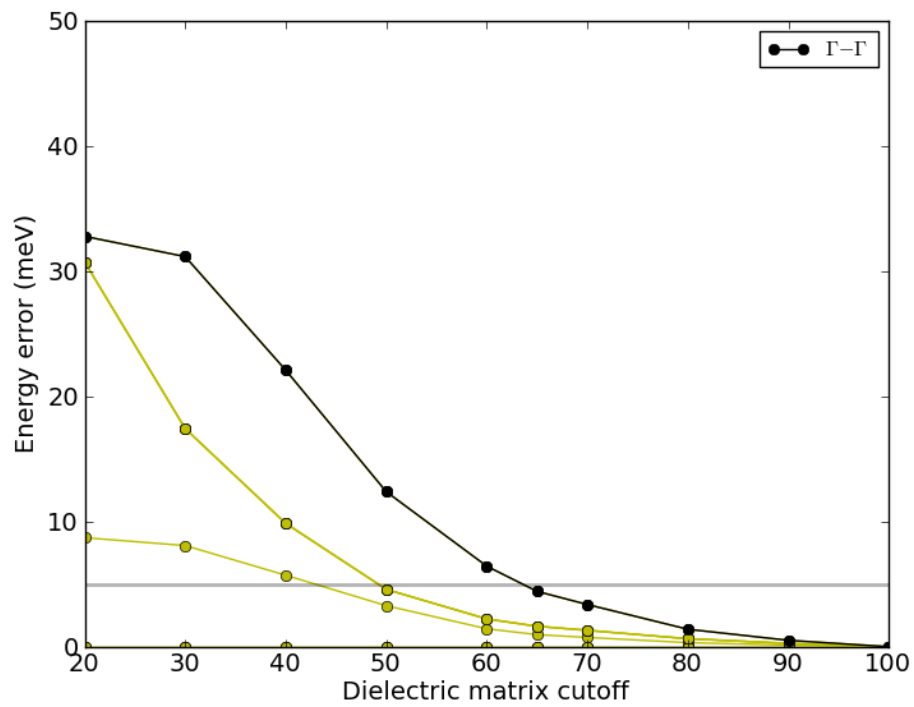


Figure A.29: Testing of the energy cutoff for the dielectric matrix in GaAs. From this plot it can be determined that 65 Ry is enough to converge the resulting quasiparticle energies to less than 5 meV.

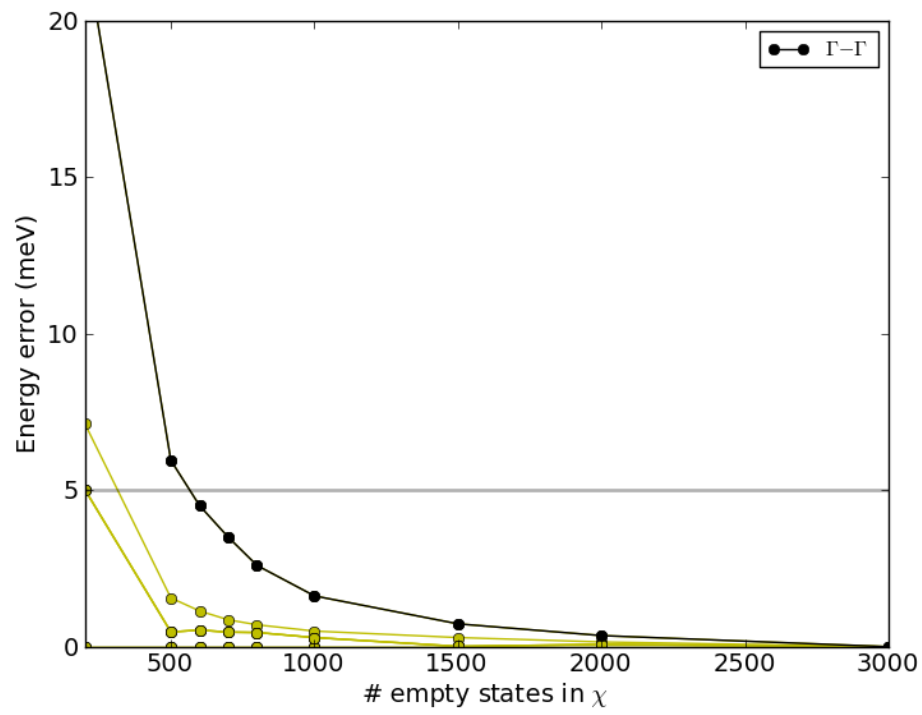


Figure A.30: Testing of the number of bands needed for the χ summation in GaAs. From this plot it is determined that 600 bands is sufficient to obtain an error less than 5 meV.

A.11 GaSb

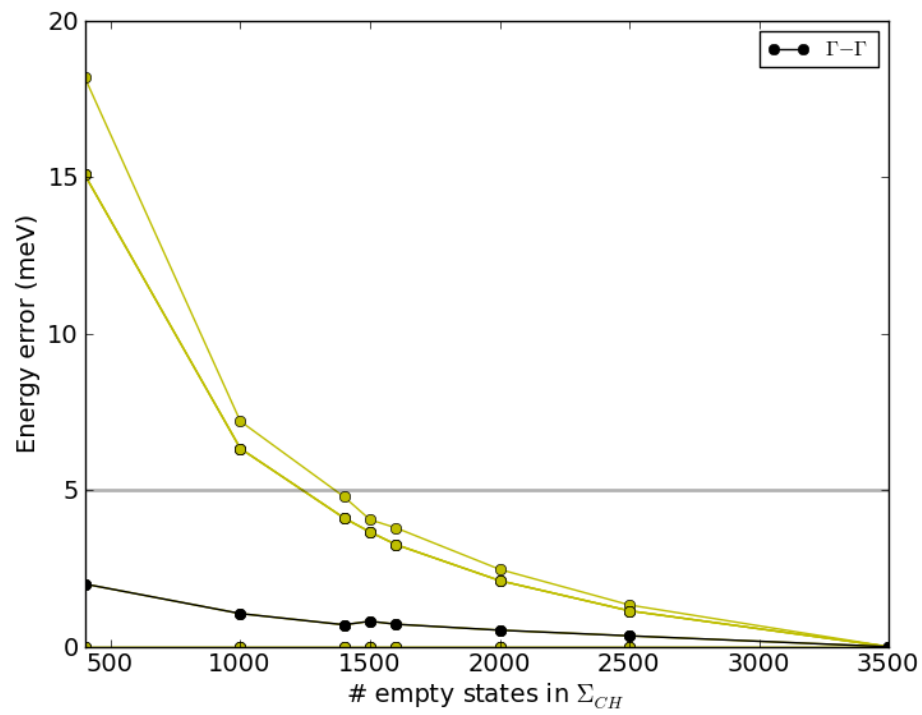


Figure A.31: Testing of the number of bands needed for the Σ_{CH} summation in GaSb. From this plot it is determined that 1500 bands is enough to obtain an error less than 5 meV.

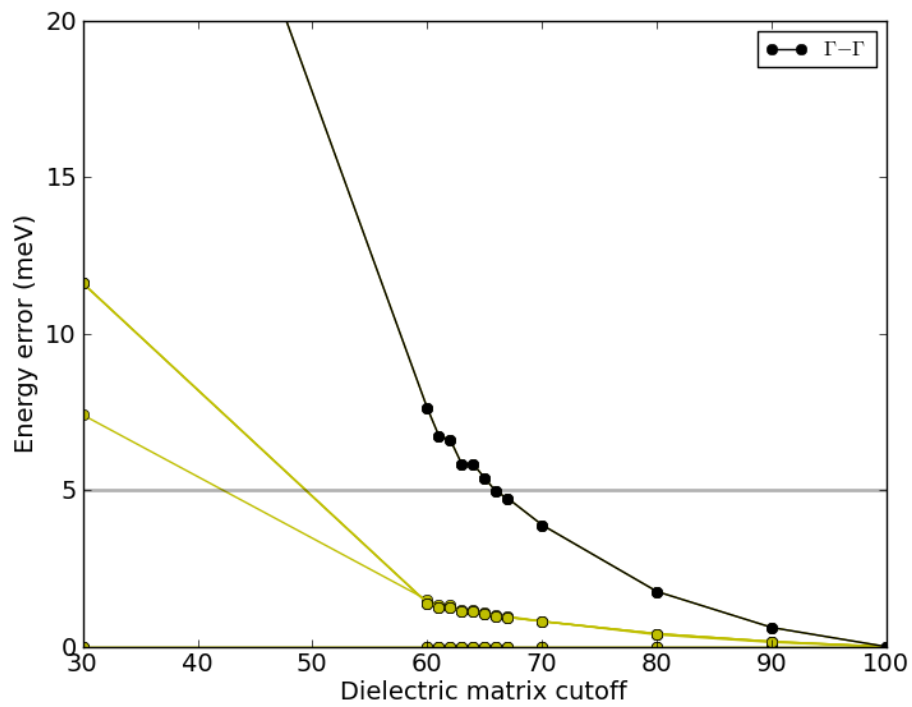


Figure A.32: Testing of the energy cutoff for the dielectric matrix in GaSb. From this plot it can be determined that 67 Ry is enough to converge the resulting quasiparticle energies to less than 5 meV.

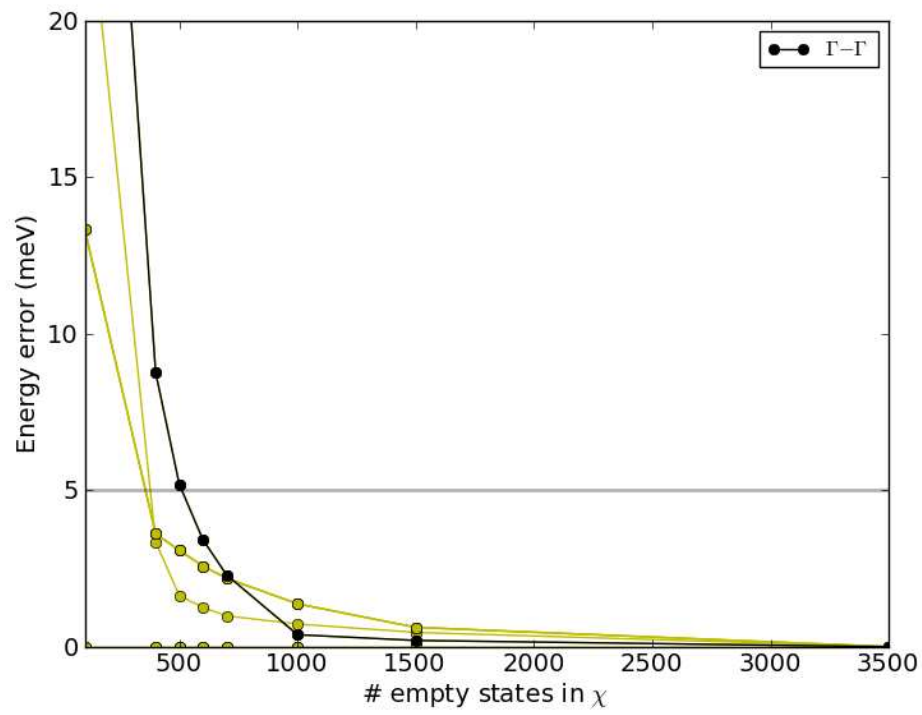


Figure A.33: Testing of the number of bands needed for the χ summation in GaSb. From this plot it is determined that 600 bands is sufficient to obtain an error less than 5 meV.

Appendix B

Structural descriptions of Si-VIII candidates

In this section we give Quantum ESPRESSO input files for the obtained 30 and 32 atom structural candidates for Si-VIII discussed in Chapter 7 in order to facilitate further study on these materials.

B.1 30 atom candidate

The 30 atom candidate is a structure with $P1$ symmetry (i.e., no symmetry). The input file for this material is shown below. Note that the Brillouin sampling is likely larger than required to converge most quantities of interest.

```
&control
  calculation='scf'
  restart_mode='from_scratch',
  prefix='si'
  pseudo_dir = './',
  outdir='./'
/
&system
 ibrav= 0
  celldm(1)= 1.8897268777743552
  nat= 30
  ntyp= 1
  ecutwfc=40.0
/
&electrons
  conv_thr = 1.0d-10
  mixing_beta = 0.7
```

```

    startingwfc='atomic+random'
/
CELL_PARAMETERS {cubic}
    8.560964823  -0.263476152  -0.000048179
   -0.257126924   9.187091473   0.000029574
   -0.000044769   0.000045295   7.511112458
ATOMIC_SPECIES
Si 28.086 Si.vbc.UPF
ATOMIC_POSITIONS {crystal}
Si      0.149018121  0.642213443  0.786314836
Si      0.076295408  0.392334496  0.184816730
Si      0.076355088  0.387558739  0.684808269
Si      0.138454806  0.646085019  0.288062478
Si      0.047540711  0.764958085  0.034639113
Si      0.041849524  0.762795904  0.541785498
Si      0.322626135  0.342502777  0.067279902
Si      0.110865212  0.017085847  0.041766723
Si      0.105171365  0.014923137  0.534634265
Si      0.320504229  0.338943653  0.555368495
Si      0.014228261  0.133796050  0.788046011
Si      0.003687432  0.137656513  0.286296660
Si      0.699491558  0.359251162  0.310491474
Si      0.453192976  0.420674494  0.810510446
Si      0.832161175  0.440979386  0.055337765
Si      0.830026884  0.437432901  0.567277052
Si      0.450780446  0.442116257  0.312626788
Si      0.701889878  0.337779507  0.812609673
Si      0.356617563  0.099365829  0.461605346
Si      0.418510135  0.685603097  0.298609260
Si      0.732356378  0.109043826  0.298386509
Si      0.598166383  0.014516897  0.546718337
Si      0.554717274  0.764458073  0.551306151
Si      0.796076066  0.680573028  0.961596203
Si      0.554523207  0.765417637  0.046718324
Si      0.597970872  0.015476696  0.051340425
Si      0.794924613  0.679116776  0.638468836
Si      0.734172945  0.094293900  0.798626076
Si      0.357766306  0.100826359  0.138470568
Si      0.420353891  0.670885791  0.798368214
K_POINTS automatic
6 6 6 1 1 1

```


B.2 32 atom candidate

As mentioned in the text of Chapter 7 the 32 atom candidate is actually a supercell of an 8 atom per unit cell structure with space group $P222_1$. The input file for this material is shown below. Note that the Brillouin sampling is likely larger than required to converge most quantities of interest.

```
&control
  calculation='scf'
  restart_mode='from_scratch',
  prefix='si'
  pseudo_dir = './',
  outdir='./'
  tstress= .true.
  tprnfor = .true.
/
&system
 ibrav= 0
  celldm(1)= 1.8897268777743552
  nat= 8
  ntyp= 1
  ecutwfc=40.0
/
&electrons
  conv_thr = 1.0d-10
  mixing_beta = 0.7
/
CELL_PARAMETERS {cubic}
  3.957165976   0.000000000   0.000000000
  0.000000000   8.678004394   0.000000000
  0.000000000   0.000000000   4.085544736
ATOMIC_SPECIES
  Si 28.086 Si.vbc.UPF
ATOMIC_POSITIONS {crystal}
Si      0.000000001   0.190072541   0.250000000
Si      0.000000001  -0.190072541   0.750000000
Si      0.251554178   0.373063077   0.594746662
Si     -0.251554180  -0.373063076   1.094746658
Si     -0.251554180   0.373063076  -0.094746658
Si      0.251554178  -0.373063077  -0.594746662
Si      0.339135261   0.000000000   0.000000000
Si     -0.339135260   0.000000000   0.500000000
K_POINTS automatic
```

12 6 12 1 1 1