THEORY OF LOCALIZED ELECTRON STATES AND NOVEL STRUCTURAL MODELING OF AMORPHOUS SILICON

A dissertation presented to

the faculty of

the College of Arts and Sciences of Ohio University

In partial fulfillment

of the requirements for the degree

Doctor of Philosophy

Raymond Atta-Fynn

November 2005

This dissertation entitled

THEORY OF LOCALIZED ELECTRON STATES AND NOVEL STRUCTURAL MODELING OF AMORPHOUS SILICON

BY

RAYMOND ATTA-FYNN

has been approved

for the Department of Physics and Astronomy

and the College of Arts and Sciences by

David A. Drabold

Distinguished Professor of Physics

Benjamin M. Ogles

Interim Dean, College of Arts and Sciences

ATTA-FYNN, RAYMOND Ph.D. November 2005. Physics

Theory of Localized Electron States and Novel Structural Modeling of Amorphous Silicon (198 pp.)

Director of dissertation: David A. Drabold

Traditionally, there is no scheme for modeling continuous random network (CRN) configurations of amorphous silicon (*a*-Si) that was not based upon the use of inter-atomic potentials, and molecular dynamics or Monte Carlo. Such schemes are known as *forward* modeling schemes. In this dissertation, we demonstrate how a *reverse* modeling scheme, known as reverse Monte Carlo (RMC), which employs information from experiments and appropriate *a priori* network topology constraints, can be used to model *realistic* atomistic configurations of *a*-Si without necessarily using inter-atomic potentials. The RMC approach is flexible and we have used it to form new models of *a*-Si that are consistent with fluctuation electron microscopy (FEM) experiments (which measures the amount of medium range order in a material), which the conventional CRN model of *a*-Si fail to satisfy.

Defects in covalent materials display a host novel properties and fully understanding their behavior is important from a fundamental and technological point of view. We systematically studied the localization of dangling bond defect electron states in silicon by performing *ab initio* static lattice calculations. Using defected models of amorphous and crystalline silicon, and a localized basis density functional Hamiltonian, we studied the dependence of wave function and spin localization on exchange-correlation functionals and localized basis sets. We observed that the minimal basis set tends to overestimate measures of localization, and we came to the conclusion that to accurately represent of the localization dangling bond defect electron states, a larger basis set is nec-

essary. Then, we added thermal disorder to the underlying topological disorder of the lattice and, assuming the harmonic approximation, showed that the electron-phonon coupling is large for localized defect electron states. We deduced analytic expressions connecting a *static property*, that is, wave function localization [gauged by the inverse participation ratio (IPR)] and a *dynamic property*, that is, mean square thermal fluctuations of the electronic eigenvalues, for localized electron states. In particular, both the variance of the electronic eigenvalues and the square of the electron-phonon coupling are linearly related to the IPR of the localized states. We verified this from first principles thermal molecular dynamics simulations.

Approved:_____

Distinguished Professor of Physics

Acknowledgments

Glory be to the holy trinity: God the Father, Jesus Christ, and the Holy Ghost for continuous guidance and protection. Rastafari every time.

I am highly indebted to my advisor Professor David Drabold, and to him, I give maximum thanks. Besides the immense knowledge that he imparted on me, I learned a great deal from him, particularly how to write quality scientific articles and report. He was extremely generous and kind, and thanks to him, I am leaving Ohio University to pursue further research with extreme confidence. Maximum respect is due to Professor Drabold.

Besides my advisor, I have had a great opportunity to work with wonderful and good physicists whiles pursuing my doctoral research. I give heartfelt thanks to Dr. Parthapratim Biswas, a former postdoctoral research fellow of Professor Drabold's group and a collaborator. Apart from the stimulating discussions that I had with Dr. Biswas, he taught me a great deal, especially how to write complex and efficient computer codes. I want to thank Professor Pablo Ordejón for a wonderful collaboration during his visit here at Ohio University in the fall of 2003. I also thank Professor Normand Mousseau and Professor Gerard Barkema for providing models and codes used for part of this work.

Special thanks go to the following: Professor David Ingram, Professor Surrender Jain, and Dr. Nancy Sandler for serving on my dissertation committee; Professor John Abelson, Professor Mike Treacy, and Dr. Paul Voyles for providing experimental data and code that I used for part of the work on FEM. I am grateful to the following for useful discussions and help: Dr. Serge Nakhmanson, Dr. De Nyago Tafen, Dr. Daniel Asare, Tesfaye Abtew, Dr. Archibald Peters, and Don Roth. I am also grateful to Ohio University, especially the Physics department, and the National Science Foundation for the financial support during the period of my graduate studies and research.

Had it not been for the love, support, and encouragement of my better half, Autumn, whom I met whiles studying here at Ohio University, I would not have easily made it alone. To Autumn, I express my deepest gratitude. I would like to thank Autumn's family: her parents Angela and Steve, her siblings Summer and Winter, and her grandparents Margie and Helen for their love and support.

Finally, I say a big thank you to my wonderful and exceptionally loving mother, Theresa, for her constant love, support, and care, I thank my father, Sebastian Sr., for being a great friend, teacher and advisor, and last but not the least I thank my siblings Gertrude, Sebastian Jr., Pius, Michael, Veronica and Laverne, and my grandmother Sophia for their love and prayers.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Understanding the properties of disordered or amorphous systems is one of the major challenges in condensed matter physics. Amorphous semiconductors in general and amorphous silicon (a-Si) in particular have vast existing and promising future applications as electronic devices and large area flat displays. Today, the photostructural properties of the chalcogenide glass GeSbTe is used in CD and DVD technology. Amorphous silica a-SiO_2, major dielectric material in electronic circuits, is also used in fiber optic cables [1].

According to the US Department of Energy forecasts, domestic use of energy will grow by about 25% in the year 2025. Oil, no longer cheap, may soon decline in availability. One of the several adverse effects of the process of oil consumption is the massive emissions of carbon dioxide, which is primarily responsible for global warming. It is therefore imperative to generate alternative sources of energy, particularly energy from renewable sources like water, wind, sunlight, and heat from the ground to supplement the increasing energy demand by an increasing population and also by industries [2].

So what role can amorphous silicon play to alleviate this energy crisis? First of all, silicon is a major material for manufacturing photo-voltaic solar devices. Typical example of these devices are wrist watches, calculators, and solar cell panels. For solar cell panels, *a*-Si is preferred over crystalline silicon (*c*-Si) because it absorbs solar radiation more efficiently and so only thin films (and therefore a relatively small amount) are needed. Also, there is an economic advantage for using *a*-Si because of the low cost of production and its deposition on inexpensive substrates. However, as promising as this may sound, there are some major hurdles that need to be cleared. The bad news is that prolonged exposure of *a*-Si to light degrades the material, and hence the performance of the solar cells, because of the creation of defects. The creation of defects in amorphous silicon upon exposure to light is known as the Staebler-Wronski effect [3]. Defects degrade the performance of devices as they act as centers for electron-hole recombination and thereby limit the flow of current. This is one of the several reasons why we have to understand defects in amorphous silicon. Most semiconductors, be it amorphous or crystalline, have some form of defect. There are several kinds of defects and the study of defects in semiconductors is a broad field.

1.1 Atomic Ordering

The nature of atomic arrangements in a material characterizes the type of order possessed by the material. In solids and liquids, atomic ordering may be short, medium, or long. Short range order (SRO) describes the order that involves atoms and their first nearest neighbors. Typical examples of SRO parameters are atomic bond lengths and bond angles which involve two and three atoms respectively. Medium range order (MRO) describes atomic order beyond the first nearest neighbor shell. MRO corresponds to the 5-20 Å length scale. The concept of MRO in a network

is not easily defined. Here, we give some examples using the definition by Elliott [4]. Elliott [4] categorizes MRO into 3 groups: short MRO (SMRO) which typically spans a length scale of 5 Å, intermediate MRO (IMRO) which corresponds a length scale of 5-10 Å, and long range MRO (LRMRO) which corresponds to 10-20 Å length scale. An example of SMRO parameter is the dihedral angle which involves 4 atoms[1]. An example of IMRO parameters in a network is the presence of 'super-structural' units (consisting of several basic SRO units). Such structural units may be clusters of atoms or rings[2] of particular shapes and sizes, which occur in a considerably higher proportion than would be expected on a purely random basis. Finally, long range order (LRO) defines order on an infinitely long length scale. Only crystals possess such an order.

1.2 Structure of Amorphous Materials

Amorphous materials do not possess LRO and consequently lack the periodicity characteristic of a crystal. The absence of LRO or periodicity in amorphous materials has been proved in diffraction experiments [5]. In X-ray diffraction experiments for amorphous solids, broad diffuse halos are observed instead of the periodic array of sharp Bragg spots or rings produced by a single crystal or polycrystalline material [4]. Hence, the structure of crystals can be unambiguously defined using a unit cell and a set of symmetry operations. The structure of amorphous materials cannot be described this way [1]. Nevertheless, to say that amorphous materials have no well-defined structural order would be incorrect. Even though amorphous materials do not possess LRO,

[1] To define the dihedral angle, consider 4 atoms $\{A, B, C, D\}$ such that A is bonded to B, B is bonded to C, C is bonded to D, and A is not bonded to D. The dihedral angle is defined as the angle between the plane passing through the atoms A,B,C and the plane passing through the atoms B,C,D.

[2] An n-fold ring is the shortest closed path of atomic bonds involving n-atoms. The closed path begins on a particular atom and ends on the same atom.

they have an appreciable amount of SRO and to some degree an amount of MRO, which is quite controversial and a subject of intense investigation.

Disorder in real materials is manifested in different forms. These include topological or geometric disorder, magnetic disorder, and compositional disorder. Topological disorder is a direct consequence of lack of LRO or absence of translational periodicity. In this work, we will consider topologically disordered networks of amorphous silicon. To interpret experiments, the theorist must make models whose properties agree with experiments. The theorist can further use the models to make predictions that are not yet realized in experiments. One successful and widely used theoretical model for topologically disordered networks is the *continuous random network* (CRN) model proposed by Zachariasen [6] in 1932. The CRN model has the simplicity that each of the atoms should satisfy its local bonding requirements and should have as small strain as possible in the network, which is generally characterized by having a narrow bond angle as well as bond length distribution. For *a*-Si in particular, the CRN model should satisfy the following conditions:

1. LRO should be absent from the network.

2. Every atom should be fully coordinated[3] to meet its ideal bonding requirements. For covalently bonded materials, the general rule for determining coordination is Mott's $8-N$ rule [7], where N is number of valence electrons. This translates to coordination number of 4 for Si.

3. Disorder should be uniform in the network. In particular, ordered crystalline zones and voids should be absent.

The structural modeling of tetrahedral amorphous semiconductors appears to be quite difficult. The first CRN model of *a*-Si was hand-built using balls and sticks [8]. This process was quite cumber-

[3]The coordination number of atom is the number of atoms it is bonded to.

some and time-consuming. Since 1976, CRN models were built via computers. The first computer CRN model was generated by Guttman [9]. From then on, models were designed using computers, and enormous progress has been made in that direction. CRN models have successfully described the electronic and structural properties realized in experiments. We will discuss the computer modeling methods for the continuous random network in Chapter 3.

In Fig. 1.1, we show pictures of a 1000-atom CRN model of amorphous silicon (top picture) and a 1000-atom model crystalline silicon (bottom picture). The LRO absent in the amorphous matrix is obvious. At zero temperature, the bond length distribution of the crystal is idealized to be a Dirac delta function centered on its length whereas the distribution for the CRN would be much broader. It can clearly be seen that the variations in the bond lengths and bond angles in the CRN results in the lack of order over long range distances.

1.3 Mathematical Description of Structural Properties

Several important properties of amorphous materials are determined by the structure of the material on the microscopic level. Mathematically, the structural properties can be described in real space and reciprocal space. As stated earlier, the structure of crystals can be uniquely defined using a unit cell and its periodicity. The structure of amorphous materials cannot be described this way because of the random variations in its structure. Rather, statistical analysis is used to describe their structural properties. The analysis are performed using atomic distribution functions [1]. We briefly discuss below, the distributions of structural properties that are frequently used to characterize disordered systems. For the derivations of the associated formulas, the reader may refer to Ref. [10].

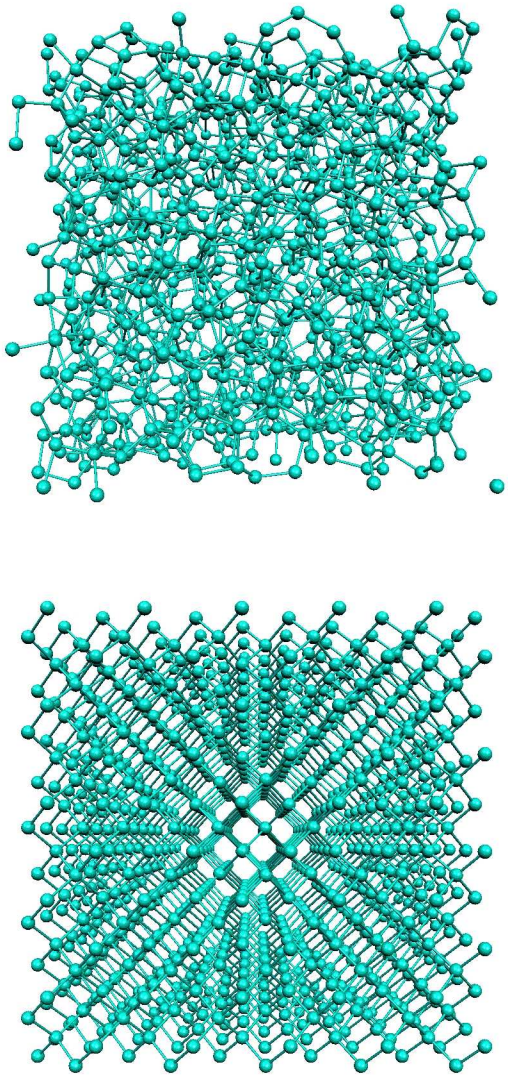

Figure 1.1: Top: topological disorder exhibited by a 1000-atom model of amorphous silicon. Bottom: LRO and periodicity exhibited by a 1000-atom model of diamond crystalline silicon.

1.3.1 Real Space

One of the most important and simplest real space distributions is the pair correlation function (PCF). PCF measures the probability of finding an atom at position $\vec{r}$ relative to an atom located at the origin [10]. It is written as

$$g(\vec{r}) = \frac{1}{\rho N} \sum_i \sum_{j \neq i} \delta(\vec{r} - \vec{r}_i)\delta(\vec{r} - \vec{r}_j), \tag{1.1}$$

where ρ is the average number density. Physically, PCF is the simplest atomic distribution function that tells us how, on the average, atoms are distributed around each other. If we assume that a disordered system is isotropic, then we can integrate out the angular dependence θ and ϕ to obtain a purely radially dependent function $g(r)$ called the radial distribution function (RDF). The RDF is given by [10]

$$g(r) = \frac{1}{4\pi\rho r^2 N} \sum_{i \neq j} \delta(r - r_{ij}), \tag{1.2}$$

where $r_{ij} = |\vec{r}_i - \vec{r}_j|$. It follows that the distribution

$$J(r) = 4\pi\rho r^2 g(r) = \frac{1}{N} \sum_{i \neq j} \delta(r - r_{ij}) \tag{1.3}$$

is the probability of atoms that are at a distance r from each other [4]. It can clearly be observed that $J(r)$ grows as the surface area of the sphere. The total number of atoms n in a spherical shell beginning at $r = r_1$ and ending at $r = r_2$ is given by

$$n = \int_{r_1}^{r_2} J(r)dr. \tag{1.4}$$

$g(r)$ (and subsequently $J(r)$) cannot be measured directly from experiment. They are obtained indirectly from the Fourier transform of the static structure factor which is obtained experimentally from X-ray diffraction, neutron diffraction or electron diffraction. We will briefly discuss the static structure factor in the next subsection.

In Fig. 1.2 we show the RDF of amorphous and crystalline silicon. The RDF of the crystalline structure, shown in the bottom figure, has well-defined sharp peaks corresponding to the atomic coordination shells. For the crystal, the value of the inter-atomic separation r at the n^{th} peak is the n^{th} nearest neighbor distance. In the top figure, we show the RDF of amorphous silicon obtained from X-ray diffraction experiments using ion implantation [11]. As r increases beyond the first shell, the peaks broaden and ultimately vanish at around $r = 8$ Å. This makes the extraction of information about MRO particularly difficult. In short, the RDF gives us rich information about SRO but not MRO. The first peak in RDF in each figure is sparse sampling of the bond length distribution. Other important material properties that can be directly computed from real space as far as theoretical models are concerned are the distribution of bond angles, the distribution of dihedral angles, and the distribution of rings.

1.3.2 Reciprocal Space

Experimentally, some of the properties of amorphous silicon can be measured using diffraction techniques like X-ray, neutron, and electron scattering. For more information on the experimental techniques, the reader may refer to Ref. [4]. The most widely measured experimental quantity in reciprocal space is the static structure factor (SSF). The SSF, which gives information about the scattering properties and internal arrangement of a material, is defined as [4]

$$S(\vec{k}) = \frac{1}{N}\left\langle \sum_i \sum_j e^{-i\vec{k}\cdot(\vec{r}_i - \vec{r}_j)} \right\rangle, \tag{1.5}$$

where $\vec{k}$ is the scattering vector. Again, if we assume that amorphous materials are isotropic i.e., $S(\vec{k})$ is independent of the direction of the scattering vector $\vec{k}$ but depends only on its magnitude k,

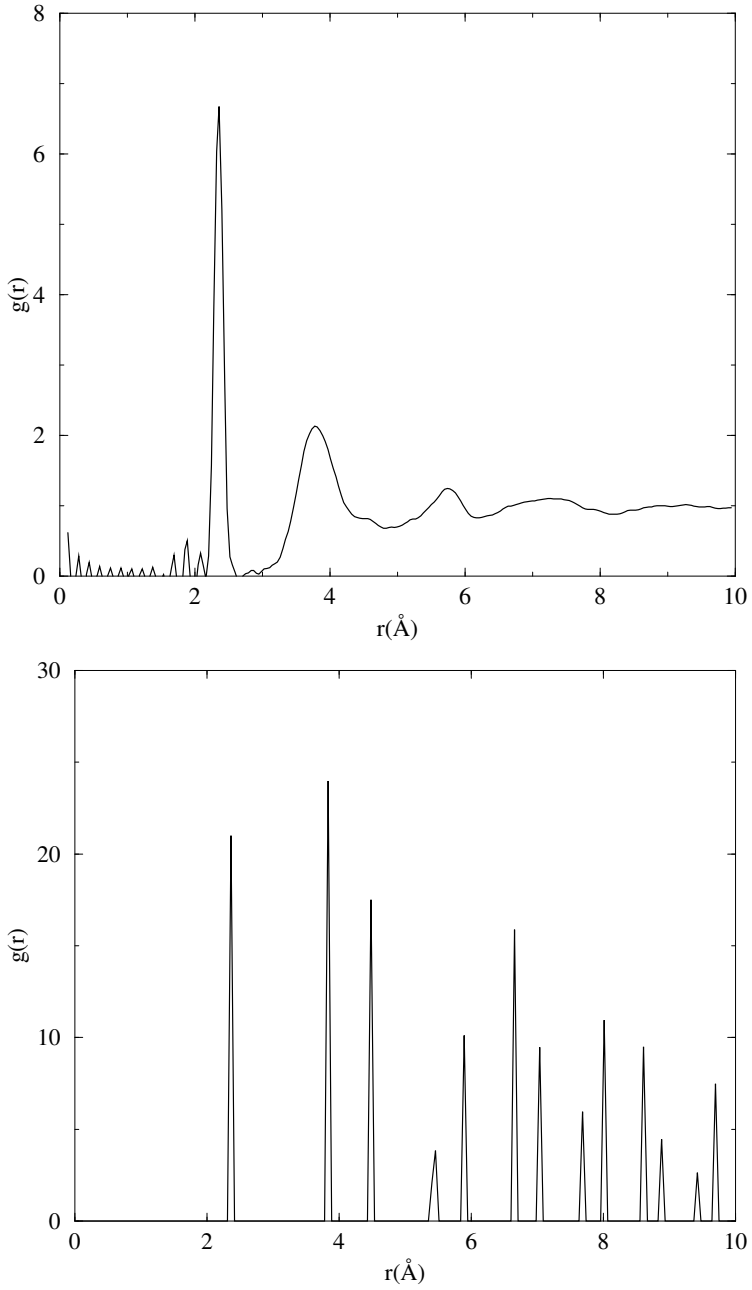

Figure 1.2: Top: RDF of amorphous silicon obtained via Fourier transform of experimental static structure factor data due to Laaziri and co-workers [11]. Bottom: RDF of computer generated model crystalline silicon. The distribution in the $r < 2$ region are artifacts of the data analysis.

then it can be shown that

$$S(k) = 1 + \frac{4\pi\rho_0}{k}\int_0^\infty r(g(r)-1)sin(kr)dr. \tag{1.6}$$

To extract the RDF, $g(r)$ from the experimentally measured $S(k)$, one inverts equation 1.6:

$$g(r) = 1 + \frac{1}{2\pi^2 r\rho_0}\int_0^\infty k(S(k)-1)sin(kr)dk. \tag{1.7}$$

We must mention that the RDF or the SSF are necessary but not sufficient to fully characterize the information inherent to the structure of a material. In practice, one can come up with many hypothetical (unrealistic models) that have RDF and SSF matching closely with experiments. One always has to go a step further to probe the electronic and vibrational spectra of the model.

1.4 Vibrational Properties

Vibrational excitations in amorphous materials and crystals are quite important as they determine the thermal properties of a material. Lattice vibrations occur in amorphous materials just as in crystals. By exploiting the periodicity of the lattice, vibrational modes of a crystal are well described using the Bloch wave vector. Dispersion relations are therefore used to express the phonon frequencies in terms the wave-vector [13].

For amorphous systems, vibrational bands cannot be described this way since the wave vector is no longer a good quantum number. However, the modes in amorphous systems are well defined. The vibrational bands can be described using the vibrational density of states (VDOS). If ω is the frequency of vibration, then $g(\omega)d\omega$ is the number of modes in the interval $[\omega, \omega + d\omega]$, where $g(\omega)$ denotes the VDOS. For a system with N atoms, the VDOS is defined as

$$g(\omega) = \frac{1}{3N}\sum_{i=1}\delta(\omega - \omega_i), \tag{1.8}$$

where ω_i are the vibrational eigenvalues. Experimentally, the VDOS is measured using inelastic neutron scattering and it is connected to the Raman spectroscopy, but they are different in detail as the Raman spectrum involves transition matrix elements. The Raman spectra tells us a lot about the scattering properties of a material and it has been used to experimentally estimate the width of the bond angle distribution $\Delta\theta$ of *a*-Si [12].

The eigenstates of dynamical matrix $\mathcal{D}$ are the vibrational modes. To compute $\mathcal{D}$, the model is well relaxed structurally to obtain the ground state configuration. Next, the harmonic approximation to the inter-atomic potential is used to compute the dynamical matrix elements. To construct a matrix element of $\mathcal{D}$, begin by first displacing a single atom I by an infinitesimal displacement $\Delta x_{I\alpha}$ along a coordinate direction $\alpha = (x, y, z)$ and then compute the forces $F_{J\beta}$ on all atoms J (including I) along a given direction β. The dynamical matrix element between atoms I and J having masses M_i and M_j respectively is the defined as [13, 14]

$$\mathcal{D}_{I\alpha,J\beta} = \frac{1}{\sqrt{M_i M_j}} \frac{F_{J\beta}}{\Delta x_{I\alpha}}. \tag{1.9}$$

For a given atom, there are $3N$ matrix elements, where N is the total number of atoms. This implies that the dimension of $\mathcal{D}$ is $3N \times 3N$. The eigenvalues of $\mathcal{D}$ are $\{\omega_i^2 : i = 1, 2, \ldots, 3N\}$ and there are $3N$ vibrational modes, each with dimension $3N$. If the inter-atomic forces are obtained from direct diagonalization of a first principles Hamiltonian, then computing $\mathcal{D}$ scales as $\mathcal{O}(N^4)$ since the force computation scales as $\mathcal{O}(N^3)$ and the construction the matrix elements in equation 1.9 scales as $\mathcal{O}(N)$. Therefore computing $\mathcal{D}$ for large systems from first principles can be quite demanding. However, Ordejón *et al*. [15], and quite recently Biswas [14], have used the decay of the dynamical matrix elements to develop first principles ($\mathcal{O}(N)$) methods for computing $\mathcal{D}$. For forces

derived from empirical and semi-empirical potentials, the dynamical matrix computation is far less demanding.

In Fig. 1.3, we depict the experimental VDOS of polycrystalline silicon and amorphous silicon obtained inelastic neutron scattering [16]. The top figure depicts the VDOS of *a*-Si which is compared with a 216-atom CRN model (open circles). The VDOS of polycrystalline silicon, shown in the lower figure, consists of well-defined sharp peaks and has marked discontinuities in the first derivatives known as Van Hove singularities [17] due the the periodicity of the lattice. For *a*-Si, the sharp peaks become broadened, smoothing out the singularities. However, there is a significant overlap between the transverse acoustic and transverse optical peaks of the amorphous and crystalline structure. This implies that the vibrational features of both crystalline and amorphous materials depends to a large extent on SRO. The disorder in the amorphous network usually leads to localized vibrational modes whereas crystalline vibrational modes are extended in real space.

1.5 Electronic Properties

Electrons in a crystal move in a periodic potential and their wave-functions are extended Bloch states. The periodicity results in gaps in the well-defined electronic band structure. The band structure tells us whether a crystal is a conductor or an insulator. In an amorphous network, the absence of periodicity does not allow the wave-functions to be expressed in terms of Bloch waves due to the fact that the translation operator does not commute with the Hamiltonian. However, the eigenenergy E of electron states is a good quantum number and the electronic density of states (EDOS) remains a valid description of electronic eigenstates for both crystalline and amorphous systems [4]. We define $n(E)dE$ as the number of electrons in the energy interval $[E, E + dE]$,

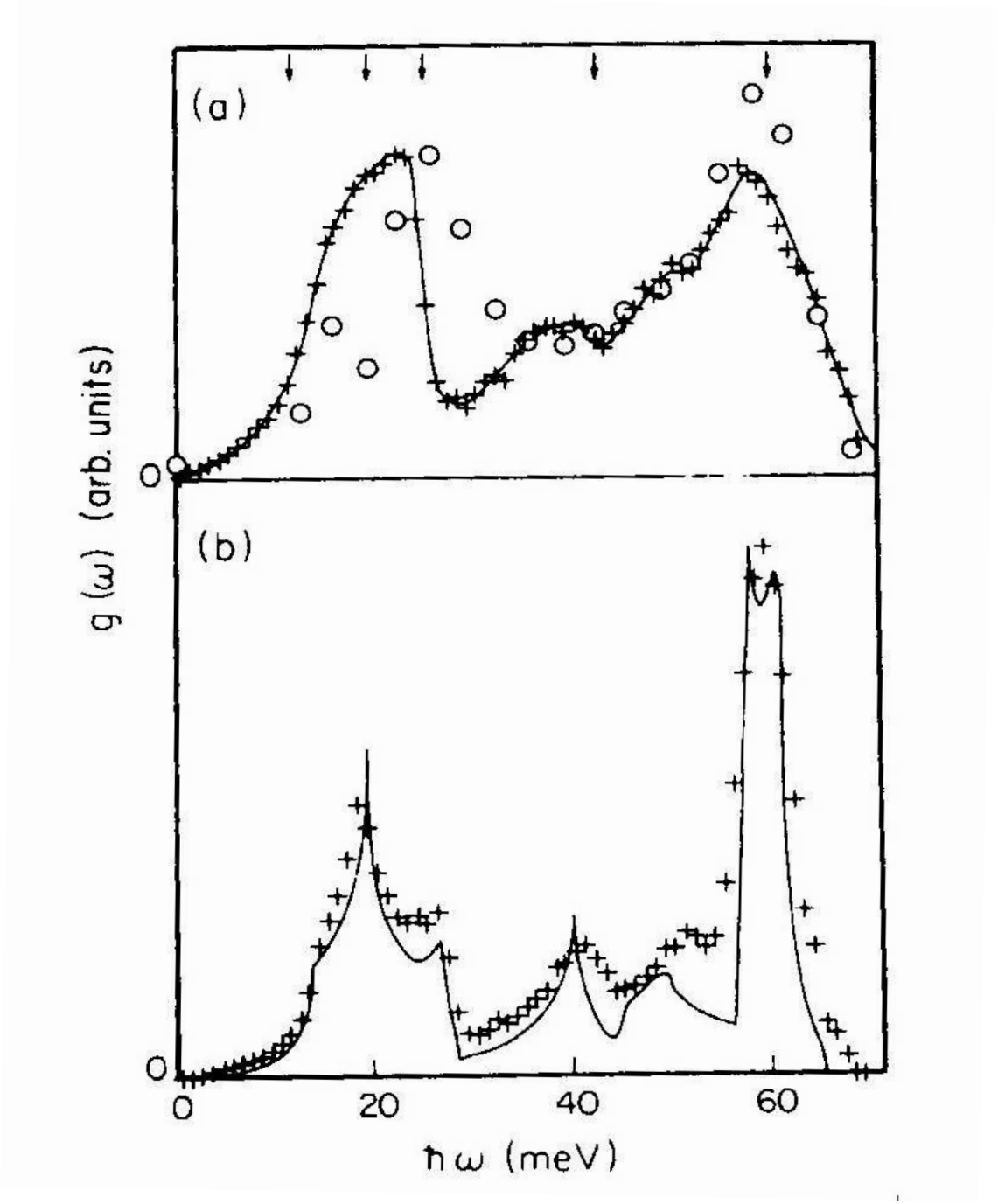

Figure 1.3: (a) Experimental VDOS of *a*-Si (solid line with crosses) due to Kamitakahara *et al*. [16]. Open circles correspond computer generated 216-atom CRN model of *a*-Si. (b) VDOS for *c*-Si. The picture was take from Ref. [16].

where the EDOS, $n(E)$, is defined as

$$n(E) = \frac{1}{N}\sum_{i=1}^{N_f}\delta(E - E_i), \tag{1.10}$$

where N_f is total number of occupied states, $\{E_i\}$ are the energy eigenvalues of the electronic Hamiltonian, and N is the number of atoms in the material. This implies that $\int_{-\infty}^{E_f} n(E)dE$ is the total number of electrons per atom, where E_f is the Fermi level. Structural disorder in the form of topological irregularities and defects lead to spatially localized electron wave-functions in the tails of the band gap and sometimes inside the band gap. Experimentally, the electronic spectrum can be probed using photo-emission spectroscopy, electron spin resonance, and other techniques [4].

1.6 Defects

Many properties of amorphous and crystalline semiconductors can be controlled by defects. These include magnetic, opto-electronic, vibrational and transport properties [4]. Unlike crystals, the disordered nature of amorphous systems makes the concept of defect not a clearly defined one. One can generally make *ad hoc* provisions to define a defect. Drabold [19] defines a defect as departure from the typical disorder of an amorphous network. The most recognizable defect in single component amorphous materials is a coordination defect. A coordination defect is either under-coordination, i.e. an atom with a missing neighbor, or over-coordination, i.e. an atom with more than enough neighbors. In binary glasses with stoichiometric compositions, defects may exist in the form of homopolar bonds but they do not necessarily impact the optical gap.

In tetrahedral covalently bonded amorphous networks like *a*-Si, the predominant coordination defect is the under-coordination defect known as a dangling bond. A dangling bond defect is an atom with three nearest neighbors instead of four or equivalently, a dangling bond state is a

valence electron that cannot pair up with another electron to form a covalent bond. In a crystal, the analogous defect is called a vacancy. To identify the dangling bond defect, one can analyze the RDF.

It is well known that the dangling bond defect introduces states in the electronic band gap (optical gap) whose wave-functions are spatially localized. Identifying defects using the electronic spectrum is more reliable than just computing atomic coordination numbers. In the case of thermal disorder, the EDOS will tell more about the nature of disorder that exists in the network. Other structural irregularities in the network (e.g. bond angles far from the tetrahedral angle) can also lead to exponential localized states in the tails of the gap even if coordination is perfect. This implies that there is not a one-to-one correspondence between coordination defects and electronic defects (localized gap and tail states to be precise). The presence of unpaired electrons spins detected in electron spin resonance experiments, is the most informative measure of dangling bond defects and possibly floating bonds.

So why study defects in semiconductors? Defects are responsible for many of the important properties of semiconductors. For example, silicon would be useless for transistors and microchips without defects in the form of dopant impurities having shallow donor and acceptor energy levels [18]. Defects exist naturally in crystalline and amorphous semiconductors alike, some desired and some unwanted. To understand and optimize the performance of semiconductors, it is essential to understand the behavior of defects in them [18]. The Staebler-Wronski effect (SWE) [3], i.e. the creation of metastable dangling bond defect when *a*-Si:H is exposed to light, degrades the performance of photo-voltaic devices made of *a*-Si:H. The atomistic origin of SWE is still not exactly known. This is a challenge to experimentalists and theorists.

1.7 Outline of the Dissertation

The outline of the dissertation is as follows. Chapter 2 reviews approximate first principles electronic structure methods with particular emphasis on the Kohn-Sham formalism of density functional theory. In chapter 3, we first briefly examine existing computer modeling techniques for generating periodic CRN supercell models of *a*-Si. Then, we devote the rest of the chapter to the best modeling scheme for generating CRN models of *a*-Si, the WWW modeling method. The structural, electronic, and vibrational properties of the WWW-generated CRN models are briefly discussed and compared to experiments.

In chapter 4, we propose a new reverse Monte Carlo (RMC) approach to modeling *a*-Si. Unlike old RMC techniques in which atomistic configuration are modeled using static structure factor data (obtained from X-ray and neutron diffraction experiments), resulting in unrealistic models, we stress on the use of key topological constraints in addition to experimental data as a necessary factor to generate realistic atomistic models.

A recent electron microscopy experiment known as fluctuation electron microscopy (FEM) suggests that the conventional CRN model of silicon has little or no medium range atomic order. To interpret FEM experiments, the para-crystalline (PC) model, which consists of nanometer-sized crystalline grains embedded in a CRN was proposed. Besides satisfying FEM experiments, the PC model has all the structural, vibrational, and electronic properties of amorphous silicon. In chapter 5, we use RMC and the WWW scheme to model new CRN models of silicon having experimental FEM signature that is otherwise absent in the conventional model.

In chapter 6, we systematically perform calculations to examine the localization properties of dangling bond defect electron states in a realistic model of *a*-Si for a static (zero temperature)

lattice. The calculations are performed using an *ab initio* DFT code SIESTA. We examine the dependence of localization of wave-function and spin on exchange correlation functionals and localized basis functions. Then, in chapter 7, thermal disorder (lattice motion) is added to the underlying topological disorder of the lattice to demonstrate that the electron-phonon coupling is large for localized states. We also show within the harmonic approximation that a static property (wave-function localization) and a dynamic property (thermal fluctuations of the electronic eigenvalues) are linearly related for eigenstates near the electronic band gap. Concluding remarks are presented in chapter 8.

Chapter 2

Electronic Structure from First Principles

A large fraction of the current problems in condensed matter physics would be solved if the electronic structure of atoms, molecules, and compounds were to be exactly computed. This amounts to solving the many-body problem in quantum mechanics to obtain the many-body wave-function describing the system. Currently, obtaining such a solution is almost impossible. Practical solution to the many-body problem relies on several approximations. In this chapter, we will review practical approximate methods for computing *ab initio*[1] or first principles electronic structure.

[1] *ab initio* is a Latin phrase which translates as "from the beginning."

2.1 The Hamiltonian for Interacting Electrons and Nuclei

Consider a system with N electrons and N_n nuclei. The Hamiltonian (non-relativistic) for the system is given by [20]:

$$\hat{\mathcal{H}} = \hat{T}_e(\mathbf{r}^N) + \hat{T}_n(\mathbf{R}^{N_n}) + \hat{V}_{ee}(\mathbf{r}^N) + \hat{V}_{en}(\mathbf{r}^N, \mathbf{R}^{N_n}) + \hat{V}_{nn}(\mathbf{R}^{N_n}), \tag{2.1}$$

where

$$\hat{T}_e(\mathbf{r}^N) = -\frac{1}{2}\sum_{i=1}^{N} \nabla_{\mathbf{r}_i}^2, \tag{2.2}$$

is the kinetic energy operator for the electrons,

$$\hat{T}_e(\mathbf{R}^{N_n}) = -\frac{1}{2M}\sum_{i=1}^{N_n} \nabla_{\mathbf{R}_i}^2, \tag{2.3}$$

is the kinetic energy operator for the nuclei,

$$\hat{V}_{ee}(\mathbf{r}^N) = \sum_{i=1}^{N}\sum_{j>i}^{N} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \tag{2.4}$$

is the Coulomb electrostatic repulsion between electrons[2],

$$\hat{V}_{en}(\mathbf{r}^N, \mathbf{R}^{N_n}) = -\sum_{i=1}^{N}\sum_{j=1}^{N_n} \frac{Z_j}{|\mathbf{r}_i - \mathbf{R}_j|}, \tag{2.5}$$

is the electron-nucleus Coulomb attraction, and

$$\hat{V}_{nn}(\mathbf{R}^{N_n}) = \sum_{i=1}^{N_n}\sum_{j>i}^{N_n} \frac{Z_i Z_j}{|\mathbf{R}_i - \mathbf{R}_j|}, \tag{2.6}$$

is the nucleus-nucleus Coulomb repulsion. In writing down the Hamiltonian, we have used atomic units (a.u.), that is

$$m_e = 1, \;\; e = 1, \;\; \hbar = 1, \;\; \frac{1}{4\pi\epsilon_0} = 1, \tag{2.7}$$

[2]For we will consider Coulomb electrostatic repulsion as the interaction between electrons. Later we will include additional terms to the electrostatic repulsion and explain why.

where m_e, e respectively denote the mass and charge of an electron and other symbols have their usual meaning.

$$\mathbf{r}^N = \{\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N\} \tag{2.8}$$

are the electron coordinates (spatial coordinates plus spin). M, Z, and

$$\mathbf{R}^{N_n} = \{\mathbf{R}_1, \mathbf{R}_2, \cdots, \mathbf{R}_{N_n}\} \tag{2.9}$$

are respectively the nucleus mass, nuclear charge, and nuclei coordinates. The many-body wave-function Ψ for the system is a function of the positions of the nuclei, position and spin of the electrons and it is given by

$$\Psi = \Psi(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N, \mathbf{R}_1, \mathbf{R}_2, \cdots, \mathbf{R}_{N_n}) \tag{2.10}$$

Solving the many body problem for a system amounts to solving time dependent and time independent Schrödinger's equation. We will mostly focus on the time independent Schrödinger equation:

$$\hat{\mathcal{H}}\Psi(\mathbf{r}^N, \mathbf{R}^{N_n}) = E\Psi(\mathbf{r}^N, \mathbf{R}^{N_n}), \tag{2.11}$$

where E is the total energy of system.

2.1.1 The Born-Oppenheimer Approximation

Certain important approximations can be made in the Hamiltonian given by equation (2.1) to reduce the difficulty in solving the many-body problem. The first approximation is the Born-Oppenheimer (BO) approximation [21]. The approximation states that the nuclei are much heavier compared to the electrons and so the electrons instantaneously adjust their positions to the motion of the nuclei. As an analogy, consider lots of flies (electrons) around a gigantic wedding cake

(nucleus) [22]. If you move the cake, the flies instantaneously adjust their positions relative to the cake. The approximation allows the positions of the nuclei to be made stationary relative to the motion of the electrons and hence, the kinetic energy operator for the nuclei motion given by equation (2.3) can be omitted from the many-body Hamiltonian. The nuclei can now be considered as a fixed external potential $\hat{V}_{ext}$ for the electronic motion:

$$\hat{V}_{ext} = \hat{V}_{en}(\mathbf{r}^N, \mathbf{R}^{N_n}) = \sum_{i=1}^{N} v_{ext}(\mathbf{r}_i), \tag{2.12}$$

where

$$v_{ext}(\mathbf{r}_i) = -\sum_{j=1}^{N_n} \frac{Z_j}{|\mathbf{r}_i - \mathbf{R}_j|}. \tag{2.13}$$

The term given by equation 2.6 is the nucleus-nucleus classical electrostatic repulsion which is simply added to the Hamiltonian. It will be ignored for now and will be added to the total electronic energy at the end of the calculation (that is, after the solution to the Schrödinger's equation for electrons have been found). Using the BO approximation and ignoring the classical nucleus-nucleus repulsion term, the many-body Hamiltonian given by equation (2.1) reduces to the many-body *electronic* Hamiltonian

$$\hat{\mathcal{H}}_e = \hat{T}_e(\mathbf{r}^N) + \hat{V}_{ee}(\mathbf{r}^N) + \hat{V}_{ext}(\mathbf{r}^N, \mathbf{R}^{N_n}). \tag{2.14}$$

The BO approximation now allows us to decouple the electronic degrees of freedom in the many-body wave-function from the nuclear degrees of freedom:

$$\Psi(\mathbf{r}^N, \mathbf{R}^{N_n}) = \psi(\mathbf{r}^N, \mathbf{R}^{N_n})\xi(\mathbf{R}^{N_n}). \tag{2.15}$$

where $\psi(\mathbf{r}^N, \mathbf{R}^{N_n})$ and $\xi(\mathbf{R}^{N_n})$ respectively denote the electronic and nuclear parts of the many-body wave-function $\Psi(\mathbf{r}^N, \mathbf{R}^{N_n})$. The Schrödinger equation for the electrons is given by

$$\hat{\mathcal{H}}_e\psi(\mathbf{r}^N, \mathbf{R}^{N_n}) = E_e\psi(\mathbf{r}^N, \mathbf{R}^{N_n}), \tag{2.16}$$

where E_e is the total electronic energy. Equation 2.16 is the starting point for most *ab initio* predictions of electronic properties of atoms, molecules, and solids [20]. It should be noted that the BO approximation is not always valid in electron structure calculations. There are cases in which the BO approximation is known to fail. The reader may refer to Refs. [23, 24, 25, 26] for details.

2.1.2 Physical Quantities of Interest

We will now define physical quantities that play key roles in the theory of electronic structure. We will adopt the scheme in Parr and Yang [27]. For the rest of this chapter, we will assume that the wave-function describing a given state is normalized, that is

$$\langle\psi|\psi\rangle = \int_{\mathbb{R}^{3N}} \prod_{i=1}^{N} d^3\mathbf{r}_i \psi^*(\mathbf{r}^N, \mathbf{R}^{N_n})\, \psi(\mathbf{r}^N, \mathbf{R}^{N_n}) \tag{2.17}$$

$$= \int_{\mathbb{R}^{3N}} \prod_{i=1}^{N} d^3\mathbf{r}_i |\psi(\mathbf{r}^N, \mathbf{R}^{N_n})|^2 = 1. \tag{2.18}$$

For a system in the state $\psi = \psi(\mathbf{r}^N, \mathbf{R}^{N_n})$ (which is not necessarily a solution to equation 2.16), the *expected value* of an observable corresponding to the Hermitian operator $\hat{O}$ is given by

$$\langle O\rangle = \langle\psi|\hat{O}|\psi\rangle = \int_{\mathbb{R}^{3N}} \prod_{i=1}^{N} d^3\mathbf{r}_i \psi^* \hat{O}\psi. \tag{2.19}$$

Another important physical quantity in electronic structure is the electron density. It is defined as the expected value of the electron density operator $\hat{\rho}(\mathbf{r}) = \sum_{i=1}^{N} \delta(\mathbf{r} - \mathbf{r}_i)$ [20],

$$\rho(\mathbf{r}) = \langle\psi|\hat{\rho}(\mathbf{r})|\psi\rangle = \langle\psi| \sum_{i=1}^{N} \delta(\mathbf{r} - \mathbf{r}_i)|\psi\rangle = \sum_{i=1}^{N} \langle\psi|\delta(\mathbf{r} - \mathbf{r}_i)|\psi\rangle. \tag{2.20}$$

Using the definition of the delta function and the symmetry of the wave-function, equation 2.20 simplifies to

$$\rho(\mathbf{r}) = N \int_{\mathbb{R}^{3N-3}} \prod_{i=2}^{N} d^3\mathbf{r}_i |\psi(\mathbf{r}, \mathbf{r}_2, \cdots, \mathbf{r}_N, \mathbf{R}^{N_n})|^2. \tag{2.21}$$

Integrating each side of equation 2.21 over $\mathbb{R}^3$, we obtain electron number conservation:

$$\int_{\mathbb{R}^3} \rho(\mathbf{r}) d^3\mathbf{r} = N. \tag{2.22}$$

2.1.3 Variational Principle for the Ground State

Before explicitly stating the variational principle for the ground state, let us briefly define a *functional*. In simple terms, a functional is a function of a function but not in the usual sense.

A function $y : A \longrightarrow B$ is defined by an element $y(x) \in B$ for a given element $x \in A$ but a functional $F : B \longrightarrow C$ is defined by an element $F[y(x)] \in C$ (square brackets are used to distinguish a functional from a function) *for all* values of the function $y(x) \in B$ defined at each point $x \in A$. In short, a function maps a point in a domain to a point in a co-domain but a functional maps a function defined over an entire domain to a point in a co-domain.

The differential of the functional δF is given by

$$\begin{aligned} \delta F &= F[y(x) + \delta y(x)] - F[y(x)] \\ &= \int_{x \in \widetilde{A}} \frac{\delta F}{\delta y(x)} \delta y(x)\, dx, \end{aligned} \tag{2.23}$$

where the quantity $\frac{\delta F}{\delta y(x)}$ is defined as the *functional derivative* of F with respect to y. The functional derivative of a functional F can be easily derived by performing a first order binomial expansion of $F[y(x) + \delta y(x)] - F[y(x)]$ in terms of δy [27].

From the definition of the expectation given by equation (2.20), we can define the expected value of the total energy as a functional of the wave-function for a given state:

$$E[\psi] = \langle \psi | \hat{\mathcal{H}}_e | \psi \rangle = \int_{\mathbb{R}^{3N}} \prod_{i=1}^{N} d^3\mathbf{r}_i \, \psi^* \hat{\mathcal{H}}_e \psi. \tag{2.24}$$

The variational principle for the energy:

If ψ is a trial wave-function for the true ground state wave-function ψ_0, then the ground state energy E_0 satisfies the condition

$$E_0 \leq E_e[\psi], \tag{2.25}$$

$$\tag{2.26}$$

with equality holding if and only if

$$\psi_o = \psi. \tag{2.27}$$

For the proof of this principle see Ref [28]. The variational method consists of minimizing $E_e[\psi]$ with respect parameters on which ψ depends. We can use the method of Lagrange multipliers [29] to show that the variational principle is in fact equivalent to

$$\delta[\langle\psi|\hat{\mathcal{H}}_e|\psi\rangle - E_e(\langle\psi|\psi\rangle - 1)] = 0. \tag{2.28}$$

The solution to equation (2.28) yields a supremum of the ground state energy and the corresponding wave-function is the best estimate for the ground state. In addition, the wave-function must exhibit all fermionic properties.

2.2 Approximate First Principles Electronic Structure Methods

By first principles or *ab initio* electronic structure, we mean a solution to the many-electron Schrödinger's equation without any assumptions. As stated earlier, it is not practical to find an exact eigenstate for the many-electron Hamiltonian nor is it necessary. Virtually all the information in many-body wave-functions, if we had them, is irrelevant. Here, we will consider

two major approximations to first principles electronic structure methods; the Hartree-Fock (HF) approach [30, 31] and methods based density functional theory (DFT). HF methods are popular in the Quantum Chemistry community because it affords quantum chemists the luxury of computing the properties of small systems, usually molecules, with a high degree of accuracy [32]. For most modern electronic structure work, DFT-based methods are preferred over HF methods because (1) it is relatively easier to code as a computer program, (2) it can handle much larger systems compared to the HF method, (3) it is computationally less demanding, and (4) it is flexible with respect to its applications to different systems [32]. In the following sections, we will review the HF approach and for the rest of this chapter, we will focus on Density Functional methods and its applications.

2.2.1 Hartree-Fock Methods

The HF method was an extension of the single-particle approximation proposed by Hartree [33]. In the Hartree approximation, the many-electron wave-function $\psi(\mathbf{r})$ is written as a product of orthogonal single-electron orbitals $\{\psi_i(\mathbf{r}_i)\}$,

$$\psi(\mathbf{r}) = \psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2)\cdots\psi_N(\mathbf{r}_N). \tag{2.29}$$

However, this form of the wave-function within the Hartree approximation has a serious flaw. Electrons are fermionic particles with a non-integer spin of $\frac{1}{2}$. Fermions obey Pauli's exclusion principle which states that: *In a system of identical fermions, no two particles can be in the same state*. For a system of electrons, the principle implies that no two electrons have the same set of quantum numbers (n, l, m). As a consequence of the principle, the permutation operator should have an eigenvalue of -1 when it acts on the many-electron wave-function, that is, the many-electron wave-function should be antisymmetric (change sign) when two electrons are interchanged (the exchange constitute spatial coordinate as well as spin). It is obvious that the form of the wave-function in

equation 2.29 does not change sign when two electrons are exchanged. The antisymmetry condition that the wave-function must satisfy results in a contribution to the total electronic energy known as the *electron exchange energy*. The HF method corrects the lack of antisymmetry in Hartree's approximation by expressing the many-electron wave function $\psi(\mathbf{r})$ as an antisymmetric product of orthogonal single-electron wave-functions using the Slater determinant

$$\psi(\mathbf{r}) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\mathbf{r}_1) & \psi_2(\mathbf{r}_1) & \dots & \psi_N(\mathbf{r}_1) \\ \psi_1(\mathbf{r}_2) & \psi_2(\mathbf{r}_2) & \dots & \psi_N(\mathbf{r}_2) \\ \vdots & \vdots & \vdots & \vdots \\ \psi_1(\mathbf{r}_N) & \psi_2(\mathbf{r}_N) & \dots & \psi_N(\mathbf{r}_N) \end{vmatrix}. \tag{2.30}$$

This form of the wave-function in the HF theory contains all the possible antisymmetric combinations of the single electron wave-functions and it exactly treats the electron-exchange energy that is missing in the Hartree total energy. The electron *exchange* energy is therefore defined as the difference between the Hartree and Hartree-Fock energies.

This dissertation does not use HF approach so we will not give detailed derivation of the important equations in the theory. However, it necessary to state the problem with HF theory and the need for an alternate and more reliable method. The problem with the HF method (and also the Hartree method) is the way in which the electron-electron interactions are treated. Besides the BO approximation which assumes that electrons moves in an external potential due to the nuclei, the HF methods treats interaction between electrons by assuming that each electron moves in an *mean potential field* due to the other $N-1$ electrons. This is where the problem with theory stems from. The fact is that electrons do not really move in an average potential. Rather, electrons monitor or *correlate* their respective motions avoiding each other so that there is a minimum amount of Coulomb electrostatic repulsion and thereby lower the total electronic energy [22]. The shortfall

of the HF theory is its inability to account for electron correlations. The difference between the exact total energy (corresponding to the correlated electron system) and the Hartree-Fock energy (corresponding uncorrelated electron system) known as the electron *correlation energy*.

The electron correlation energy is usually about one percent of the total (electronic plus nuclear) energy. In terms of absolute value, this is a fairly small fraction but it is crucial, especially for calculations involving light atoms. The amount of energy that is usually required for chemical processes like bond formation and bond dissociation are quite small when compared to the correlation energy. Hence to make accurate predictions of system properties via electronic structure calculations it is necessary to treat the correlation energy[3]. There are post Hartree-Fock methods that seek to include electron correlations in the original HF theory. These include Configuration Interaction (CI), Moller-Plesset (MP) perturbation theory, and Coupled cluster (CC) theory. These methods have huge computational demands so there is a serious limitation on the size of the systems that these methods can handle. For example CI scales as $N!$. HF methods and post HF methods are described in some landmark publications [34, 35] and in detail many in quantum chemistry and electronic structure textbooks [20, 27, 28, 36, 37].

For the rest of this chapter we will focus an alternative formulation, density functional theory, which also adopts the independent electron approximation to treat many-electron systems and exactly treats electron exchange just like the HF method, but treats in addition, the electron-electron correlation, albeit in an approximate manner.

[3] A well known example of the lack of the HF theory's to include electron correlations is the prediction by the theory that jellium is an insulator instead of a metal [32].

2.2.2 Density Functional Theory

Density functional theory (DFT) methods are the most widely used approximate first principles approach to computational material science today. DFT is in principle an exact formulation for the ground state of many-electron systems and it expresses ground state properties– such as equilibrium positions, total energies, and magnetic moments as functionals of the electron density $\rho(\mathbf{r})$ (see equations (2.21) and (2.22)) [38]. The formulation avoids the computation of many-electron wave-function. So why do most modern electronic structure Practitioners prefer DFT-based methods? The answer lies in the extreme reduction in the difficulty of the many-body problem. The electron density is a scalar defined on $\mathbb{R}^3$ whiles the many-body function is a scalar defined on $\mathbb{R}^{3N}$, so by avoiding the many-electron wave-function, DFT reduces the original problem by $3N - 3$ degrees of freedom.

The first attempt to formulate a density functionals was based on the Thomas-Fermi model due to Thomas [39] and Fermi [40]. Within the model, the total energy is expressed as a functional of the electron density. However, the model suffered from inaccuracies stemming from the crude formulation of the kinetic energy functional and the mean field approximation to the electron-electron interaction, which failed to include exchange and correlation effects [38]. Later on, Dirac included electron exchange contributions to the Thomas-Fermi functional but then, it still suffered from electron correlation effects [41]. Extended discussions on the Thomas-Fermi and Thomas-Fermi-Dirac models can be found in Refs. [20] and [27].

Two elegant theorems by Hohenberg and Kohn [42], which made the Thomas-Fermi model and similarly the Hartree theory exact, brought DFT theory to the forefront of electronic structure theory. The theorems and associated corollaries clearly show how the ground state energy and other ground state properties can be exactly expressed as functionals of *only* the ground state

electron density $\rho(\mathbf{r})$. Before we proceed to state the theorems, let us note that the expression for the total electronic energy given by equation 2.24 in conjunction with equation 2.20 simplifies to

$$\begin{aligned} E[\psi] &= \langle\psi|\hat{T}_e + \hat{V}_{ee}|\psi\rangle + \langle\psi|\sum_{i=1}^{N} v_{ext}(\mathbf{r}_i)|\psi\rangle \\ &= \langle\psi|\hat{T}_e + \hat{V}_{ee}|\psi\rangle + \langle\psi|\sum_{i=1}^{N}\left(\int_{\mathbb{R}^3} v_{ext}(\mathbf{r})\delta(\mathbf{r}-\mathbf{r}_i)d^3\mathbf{r}\right)|\psi\rangle \\ &= \langle\psi|\hat{T}_e + \hat{V}_{ee}|\psi\rangle + \int_{\mathbb{R}^3} v_{ext}(\mathbf{r})d^3\mathbf{r}\langle\psi|\sum_{i=1}^{N}\delta(\mathbf{r}-\mathbf{r}_i)|\psi\rangle \\ &= \langle\psi|\hat{T}_e + \hat{V}_{ee}|\psi\rangle + \int_{\mathbb{R}^3} v_{ext}(\mathbf{r})\rho(\mathbf{r})d^3\mathbf{r} \end{aligned} \tag{2.31}$$

We now state the two Hohenberg-Kohn theorems for a system of N interacting electrons in a non-degenerate ground state associated with an external potential $v_{ext}(\mathbf{r})$.

Theorem I: *The external potential $v_{ext}(\mathbf{r})$ is (to within a trivial additive constant) uniquely determined by the ground state electron density $\rho(\mathbf{r})$.*

The proof of this theorem is disarmingly simple. The reader can refer to the original paper by Hohenberg and Kohn [42] for the proof. Notice that if a solution to Schrödinger's equation is known, then the density can be easily constructed using the many-body wave-function. This is the converse of the first Hohenberg-Kohn theorem. The first Hohenberg-Kohn theorem and its converse guarantees the existence of a bijection from $\rho(\mathbf{r})$ to $v_{ext}(\mathbf{r})$. Since $v_{ext}(\mathbf{r})$ completely yields $\hat{\mathcal{H}}_e$, a consequence of the theorem is that $\rho(\mathbf{r})$ completely determines all ground state properties of the system [43].

Theorem II: *A universal functional of the density $\rho(\mathbf{r})$ for the energy $E[\rho]$ can be defined for all electron systems. Furthermore, for any trial electron density $\tilde{\rho}(\mathbf{r})$ associated with some external potential v_{ext}, the ground state electron density $\rho(\mathbf{r})$ minimizes $E[\tilde{\rho}]$ and the corresponding minimum energy is the ground state energy.*

Again, the proof for the theorem can be found in Ref. [42]. The second Hohenberg-Kohn theorem suggests that the energy functional $E[\rho]$ is sufficient to determine all ground state properties of the system.

From equation 2.31 we re-write electronic energy E as unique functional of $\rho(\mathbf{r})$,

$$E[\rho] = F[\rho] + \int_{\mathbb{R}^3} v_{ext}(\mathbf{r})\rho(\mathbf{r})d^3\mathbf{r} \tag{2.32}$$

where

$$F[\rho] = T_e[\rho] + V_{ee}[\rho] = \langle\psi|\hat{T}_e + \hat{V}_{ee}|\psi\rangle \tag{2.33}$$

is the *Hohenberg-Kohn* (HK) functional.

So far, both the exact ground state density ρ and the density functional $F[\rho]$ are unknown. So we cannot use the HK theorems to compute material properties. We now decompose the electron-electron interaction $V_{ee}[\rho]$ in equation (2.33) into the classical electrostatic repulsion between the electrons $E_H[\rho]$ (*Hartree* energy) and the *non-classical* contribution $E_{NCL}[\rho]$ (which is a major part of the electron exchange and correlation),

$$V_{ee}[\rho] = E_H[\rho] + E_{NCL}[\rho], \tag{2.34}$$

where the Hartree energy $E_H[\rho]$ is defined as

$$E_H[\rho] = \frac{1}{2}\int_{\mathbb{R}^3} d^3\mathbf{r}\rho(\mathbf{r})V_H(\mathbf{r}), \tag{2.35}$$

and the Hartree potential $V_H(\mathbf{r})$

$$V_H(\mathbf{r}) = \int_{\mathbb{R}^3} d^3\mathbf{r}' \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}. \tag{2.36}$$

$$F[\rho] = T_e[\rho] + E_H[\rho] + E_{NCL}[\rho], \tag{2.37}$$

At this point, $E_H[\rho]$ is known, $T_e[\rho]$ is unknown, and $E_{NCL}[\rho]$ is unknown.

2.3 The Kohn-Sham Formulation of Density Functional Theory

In a paper published by Kohn and Sham (KS) [44] in 1965, they proposed a formulation for mapping the full interacting system with the real potential, onto a fictitious non-interacting system whereby the electrons move within an "effective" KS single-particle potential $V_{KS}(\mathbf{r})$. The KS method is still exact since it yields the same ground state density as the real system, but greatly facilitates the calculation by partially treating the unknown kinetic energy functional $T_e[\rho]$. Starting with a set of independent electron orbitals,$\{\psi_i(\mathbf{r})\}$,[4] the Kohn-Sham (KS) kinetic energy $T_{KS}[\rho]$ of a *non-interacting* system of electrons is defined as

$$T_{KS}[\rho] = \sum_i f_i \langle \psi_i | \frac{1}{2}\nabla^2 | \psi_i \rangle = -\frac{1}{2} \sum_i f_i \int_{\mathbb{R}^3} d^3\mathbf{r} \psi_i(\mathbf{r}) \nabla^2 \psi_i^*(\mathbf{r}) \tag{2.38}$$

and electron charge density $\rho(\mathbf{r})$ as

$$\rho(\mathbf{r}) = \sum_i f_i |\psi_i(\mathbf{r})|^2 \tag{2.39}$$

where f_i is number of electrons occupying each orbital. f_i = 1 for energies below the Fermi level, f_i = 0 for energies above the Fermi level, and $\sum_i f_i = N$.

[4]The Kohn-Sham orbitals are constructed to be orthogonal, that is, $\langle \psi_i | \psi_j \rangle = \delta_{ij}$. In fact the orbitals are not orthogonal in our calculations and the normalization condition is $\langle \psi_i | S | \psi_j \rangle = \delta_{ij}$, where S is the overlap matrix.

Using equation (2.38), we can re-write equation (2.37) in terms of $T_{KS}[\rho]$,

$$\begin{aligned} F[\rho] &= T_{KS}[\rho] + E_H[\rho] + (E_{NCL}[\rho] + T_e[\rho] - T_{KS}[\rho]) \\ &= T_{KS}[\rho] + E_H[\rho] + E_{XC}[\rho], \end{aligned} \tag{2.40}$$

where

$$E_{XC}[\rho] = E_{NCL}[\rho] + T_e[\rho] - T_{KS}[\rho] \tag{2.41}$$

is the *exchange-correlation* energy functional which is unknown so far.

The expression for the energy functional given by equation (2.32) now becomes

$$E[\rho] = T_{KS}[\rho] + E_H[\rho] + E_{XC}[\rho] + \int_{\mathbb{R}^3} v_{ext}(\mathbf{r})\rho(\mathbf{r})d^3\mathbf{r}. \tag{2.42}$$

Now, $\rho(\mathbf{r})$ depends on Kohn-Sham orbitals $\{\psi_i\}$ so using the second HK theorem which is essentially a variational principle for the ground state, we will minimize $E[\rho]$ with respect to $\{\psi_i\}$ subject to the constraint that

$$\langle \psi_i | \psi_i \rangle = 1 \tag{2.43}$$

In particular, we require that

$$\frac{\delta \mathcal{L}[\rho(\{\psi_i\})]}{\delta \psi_j^*} = 0, \tag{2.44}$$

where

$$\begin{aligned} \mathcal{L}[\rho(\{\psi_i\})] &= E[\rho(\{\psi_i\}] - \sum_i \lambda_i \{\langle \psi_i | \psi_i \rangle\} = T_{KS}[\rho] + E_H[\rho] + E_{XC}[\rho] \\ &+ \int_{\mathbb{R}^3} v_{ext}(\mathbf{r})\rho(\mathbf{r})d^3\mathbf{r} - \sum_i \lambda_i \left(\int_{\mathbb{R}^3} \psi_i^*(\mathbf{r})\psi_i(\mathbf{r})d^3\mathbf{r} - 1 \right) \end{aligned} \tag{2.45}$$

and $\{\lambda_i\}$ are Lagrange multipliers.

This implies that

$$\begin{aligned}\frac{\delta\mathcal{L}}{\delta\psi_j^*(\mathbf{r})} &= \frac{\delta T_{KS}}{\delta\psi_j^*(\mathbf{r})} + \frac{\delta\rho(\mathbf{r})}{\delta\psi_j^*(\mathbf{r})}\left[\frac{\delta E_H}{\delta\rho(\mathbf{r})} + \frac{\delta E_{XC}}{\delta\rho(\mathbf{r})} + \frac{\delta}{\delta\rho(\mathbf{r})}\int_{\mathbb{R}^3} v_{ext}(\mathbf{r})\rho(\mathbf{r})d^3\mathbf{r}\right] \\ &- \lambda_j\psi_j \end{aligned} \tag{2.46}$$

Apart from the unknown $\delta E_{XC}/\delta\rho(\mathbf{r})$, we can simplify all the individual functional derivatives on the right hand side of equation (2.46) using the definition of the functional derivative in equation (2.23). Now from equation (2.38)

$$\frac{\delta T_{KS}}{\delta\psi_j^*} = -\frac{1}{2}\sum_{i=1} f_i\delta_{ij}\nabla^2\psi_i(\mathbf{r}) = -\frac{1}{2}f_j\nabla^2\psi_j(\mathbf{r}). \tag{2.47}$$

From equation (2.39)

$$\frac{\delta\rho(\mathbf{r})}{\delta\psi_j^*(\mathbf{r})} = \sum_{i=1} f_i\delta_{ij}\psi_i(\mathbf{r}) = f_j\psi_j(\mathbf{r}). \tag{2.48}$$

From equation (2.35,

$$\frac{\delta E_H}{\delta\rho(\mathbf{r})} = V_H[\rho] \tag{2.49}$$

$$\frac{\delta}{\delta\rho(\mathbf{r})}\int_{\mathbb{R}^3} v_{ext}(\mathbf{r})\rho(\mathbf{r})d^3\mathbf{r} = v_{ext}(\mathbf{r}). \tag{2.50}$$

Substituting equations (2.47), (2.48), (2.49), and (2.50) into equation (2.46) and using equation (2.44), we obtain the single particle *Kohn-Sham* equations

$$(-\frac{1}{2}\nabla^2 + V_{KS})\psi_j(\mathbf{r}) = \varepsilon_j\psi_j(\mathbf{r}), \tag{2.51}$$

where the Kohn-Sham potential

$$V_{KS}[\rho] = v_{ext}(\mathbf{r}) + V_H[\rho] + V_{XC}[\rho] \tag{2.52}$$

with $V_{XC}[\rho] = \frac{\delta E_{XC}}{\delta\rho(\mathbf{r})}$ and $\varepsilon_j = \frac{\lambda_j}{f_j}$. Thus the KS formulation reduces the N-electron problem to N single-electron problem. The solution of the system of KS equations yields the ground state

electronic charge density and hence the ground state energy and other ground state quantities which can be obtained. Most modern electronic structure calculations are based on the Kohn-Sham formulation of DFT. However, we still do not know the exact form of the exchange-correlation potential $V_{XC}[\rho]$ (since the exchange correlation functional $E_{XC}[\rho]$ is unknown). $E_{XC}[\rho]$ plays a crucial role in the exactness of the Kohn-Sham formulation of density functional theory. So far the exact analytical form for $E_{XC}[\rho]$ is unknown. However, approximations to $E_{XC}[\rho]$ exist. We briefly discuss below, the well known approximations.

2.4 Approximations to the Exchange-Correlation Functional

Finding an exact form for the exchange and correlation functional $E_{XC}[\rho]$ is the unsolved "holy grail" of electronic structure and, the success of DFT depends on it to a great extent. The most commonly used approximations for $E_{XC}[\rho]$ are the *local density approximation* (LDA) (or the *local spin density approximation* (LSDA) if it includes spin polarization) and the *generalized gradient approximation* (GGA).

2.4.1 The Local Density Approximation

Within the LDA, $E_{XC}[\rho]$ is constructed based on the assumption that exchange-correlation energy per electron $\varepsilon_{xc}(\rho)$ at point $\mathbf{r}$ in the electron gas is the same as that of an electron gas with uniform density (a uniform electron gas). $\varepsilon_{xc}(\rho)$ can be further broken into two parts:

$$\varepsilon_{xc}(\rho) = \varepsilon_x(\rho) + \varepsilon_c(\rho), \tag{2.53}$$

where $\varepsilon_x(\rho)$ is the exchange part and $\varepsilon_c(\rho)$ is the correlation part. For a homogeneous electron gas, $\varepsilon_x(\rho)$ has an analytic form [41] given by

$$\varepsilon_x^{LDA}(\rho) = -C_x\rho^{1/3}(\mathbf{r}), \quad C_x = \frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3} \tag{2.54}$$

Within LSDA, it can be shown that

$$\varepsilon_x^{LSDA}(\rho\uparrow, \rho\downarrow) = 2^{1/3}C_x((\rho\uparrow(\mathbf{r}))^{4/3} + (\rho\downarrow(\mathbf{r}))^{4/3}), \tag{2.55}$$

where $\rho\uparrow$, $\rho\downarrow$ respectively denote the electron up spin and down spin densities [27]. LSDA is particularly useful for systems with non-zero net spin and non-collinear spins.

The correlation part $\varepsilon_c(\rho)$ cannot be derived analytically. However, it has been calculated numerically with high accuracy using quantum Monte Carlo simulations by Ceperly and Alder [47]. These numerical values have been fitted to derive analytic forms for $\varepsilon_c(\rho)$ by Vosko *et al* [48], and by Perdew and Zunger [49]. Once $\varepsilon_{xc}(\rho)$ is known, the exchange-correlation functional is given by

$$E_{XC}^{LDA}[\rho] = \int_{\mathbb{R}^3} \varepsilon_{xc}^{LDA}(\rho(\mathbf{r}))\rho(\mathbf{r})d^3\mathbf{r} \tag{2.56}$$

$$E_{XC}^{LSDA}[\rho] = \int_{\mathbb{R}^3} \varepsilon_{xc}^{LSDA}(\rho\uparrow(\mathbf{r}), \rho\downarrow(\mathbf{r}))\rho(\mathbf{r})d^3\mathbf{r} \tag{2.57}$$

2.4.2 Beyond the Local Density Approximation

Besides the LDA, a number of non-local approximations for the exchange-correlation functional has been proposed. The first step beyond LDA is the use of exchange and correlation functionals involving the magnitude of the gradient of the density, known as the generalized gradient approximation (GGA). Within the GGA, E_{XC} is constructed as functional of both $\rho(\mathbf{r})$ and $|\nabla\rho(\mathbf{r})|^n$ $(n \geq 1)$ at each point $\mathbf{r}$ in the electron gas,

$$E_{XC}^{GGA}[\rho] = \int_{\mathbb{R}^3} \varepsilon_{xc}(\rho\uparrow(\mathbf{r}), \rho\downarrow(\mathbf{r}), |\nabla\rho\uparrow(\mathbf{r})|, |\nabla\rho\downarrow(\mathbf{r})|, \cdots)\rho(\mathbf{r})d^3\mathbf{r} \tag{2.58}$$

GGA functionals generally give more accurate results compared to LSDA although not in a consistent manner [20]. For example; LDA methods generally underestimate the electronic band gap compared to experiments and GGA generally does better. We will not talk about the constructions of GGA functionals. The reader may refer to Refs. [20] and [27] for detailed discussion. The most widely used GGA functional are parameterizations of Becke [50], Perdew and Wang [51], and Perdew Burke and Enzerhof [52].

It is worth making brief mention of other well known approximations. There is a perturbation Green's function approach called the GW[5] approximation which treats dynamic correlation effects due to the screened Coulomb interaction between electrons [38]. GW methods generally improves LSDA, GGA total energies and treats the excitation spectrum as well. Then, there is the LDA-type (or GGA-type) calculations coupled with an additional orbital-dependent interaction known as LDA+U i.e., LDA plus a Hubbard term which attempts to correct errors known to be large in the usual LDA or GGA calculations [20, 54, 55].

2.5 Solution of the Kohn-Sham Equations

The solution of the system of KS equations given by equation (2.51) yields the ground state electronic charge density, from which ground state energy[6] and all other ground state quantities can be derived. We note that the Kohn-Sham potential V_{KS} depends on $\rho[\mathbf{r}]$ which in turn depends on the $\psi(\mathbf{r})$. Thus the equations can be solved either self-consistently or non-self-consistently.

[5] G stands for a the representation of the self-energy as a convolution of the self energy and W stands for the screened Coulomb interaction function

[6] The total energy will now be the sum of the the total KS electronic energy and the classical Coulomb repulsion between the nuclei given by equation (2.6).

2.5.1 Self-Consistent Solution

We will briefly outline the self consistent algorithm for the solution to the KS equations:

1. Begin with some input trial density ρ_{in} which is usually the superposition of neutral pseudo-atomic charge densities.

2. Construct the Kohn-Sham potential $V_{KS}[\rho]$ according to equation (2.52) by computing the individual components. v_{ext} is simple to compute; $V_H[\rho_{\text{in}}(\mathbf{r})]$ is computed using Poisson's equation $\nabla^2 V_H(\rho_{in}\mathbf{r}) = -4\pi\rho_{\text{in}}$; V_{XC} is computed from known approximations (LSDA, GGA, e.t.c., discussed above).

3. Solve the KS equations (2.51) to obtain the eigenstates $\{\psi_i^{\text{out}}(\mathbf{r})\}$, occupation numbers f_i^{out}, and eigenvalues $\varepsilon_i^{\text{out}}$.

4. Use $\{\psi_i^{\text{out}}(\mathbf{r})\}$ from step 3 to construct the new output density $\rho_{\text{out}} = \sum_i f_i^{\text{out}} |\psi_i^{\text{out}}(\mathbf{r})|^2$.

5. Check convergence criterion: if $|\rho_{\text{out}} - \rho_{\text{in}}| < \epsilon$, where ϵ is some convergence parameter, then compute ground state properties of the system. Otherwise, replace ρ_{in} with ρ_{out}, go to step 2 and repeat the process until convergence is achieved.

6. After convergence is achieved, compute the total KS energy, which is given by

$$\begin{aligned} E[\rho_{\text{out}}(\mathbf{r})] &= \sum_i f_i^{\text{out}} \varepsilon_i^{\text{out}} - E_H[\rho_{\text{out}}(\mathbf{r})] + E_{XC}[\rho_{\text{out}}(\mathbf{r})] \\ &- \int_{\mathbb{R}^3} V_{XC}[\rho_{\text{out}}(\mathbf{r})]\rho_{\text{out}}(\mathbf{r}) d^3\mathbf{r} \\ &+ \int_{\mathbb{R}^3} V_{KS}[\rho_{\text{out}}(\mathbf{r})] - V_{KS}[\rho_{\text{in}}(\mathbf{r})]\rho_{\text{out}}(\mathbf{r}) d^3\mathbf{r} \end{aligned} \tag{2.59}$$

2.5.2 The Non-Self-Consistent Harris Approximation

Self consistent solutions to the KS equations described above can be time-consuming. Alternative approaches that seek to reduce the time-consuming self-consistent cycles and yet give reasonable solutions are of great importance in practice. Harris proposed a scheme which is suitable for covalent bonded systems where charge transfer between atomic orbitals is not too large [53]. We briefly outline Harris' method.

1. Begin with some input electron density for the $\rho_{\text{in}}(\mathbf{r})$ the ground state. An obvious choice input density in step 1 of the self-consistent solution above. Similarly, One can also start with some optimized input density.

2. Proceed just as in steps 2 and 3 in the self consistent case but in step 3, compute only $\varepsilon_i^{\text{out}}$ and f_i^{out}. The Harris energy functional does not depend on ρ_{out} and so there is no need for its construction using $\{\psi_i^{\text{out}}(\mathbf{r})\}$.

3. The expression for the total energy due Harris energy E_{Harris}, simply replace ρ_{out} with $\rho_{\text{in}}(\mathbf{r})$ in equation (2.59).

The beauty about the Harris approximation is that convergence is usually achieved in a single step and this greatly enhances the computational efficiency, especially for large systems. Also note that Poisson's equation must be solved again to obtain the Hartree potential $V_H(\rho_{\text{out}})$ in order to obtain the self-consistent total energy in equation (2.59) but the Harris energy does not require this. The Harris functional is particularly suitable and works well for methods that use localized basis sets (which we have exclusively used for this work). It has been shown in principle that, the Harris energy is an approximation to the Kohn-Sham energy, accurate to first order in $\Delta\rho$, where $\Delta\rho$ is the difference between the Kohn-Sham self-consistent field density and the Harris density [56].

2.5.3 Interpretation of the Kohn-Sham Eigenvalues

We now wish to stress that unlike the Hartree-Fock orbitals, the KS orbitals do not have a clear interpretation as single particle functions. In principle, they are mere mathematical entities designed to make the kinetic energy of the non-interacting system easy to evaluate. Similarly, the individual KS energies corresponding to the Kohn-Sham orbitals have no physical interpretation (they do not correspond to true electron addition and removal energies) except for the highest occupied orbital of a finite system, whose eigenvalue has been shown to be the negation of the ionization energy [45]. However, the eigenvalues have a well-defined meaning within the theory and they can be used to construct physically meaningful quantities. Chapter 7 of Richard Martin's book [20] discusses this in detail. Physical interpretation of the KS eigenenergies is a subject of intense research [46].

2.6 Basis Set Expansion of the Kohn-Sham Orbitals

To solve the KS eigenvalue problem, the electron wave-function can be expanded using a given basis set. The choice of an appropriate basis set is crucial to the quality of the results from electronic structures calculations and predictions.

Consider a basis set $\{\phi_\mu\}$ with N_{basis} functions in the Hilbert space. The basis functions may or may not be orthogonal. We define the overlap matrix element $S_{\mu\nu}$ between two basis functions ϕ_μ and ϕ_ν as

$$S_{\mu\nu} = \langle \phi_\mu | \phi_\nu \rangle = \int \phi_\mu^* \phi_\nu d^3\mathbf{r}. \tag{2.60}$$

We define the Hamiltonian matrix element as

$$H_{\mu\nu} = \langle \phi_\mu | \hat{\mathcal{H}}^{\mathrm{KS}} | \phi_\nu \rangle = \int \phi_\mu^* \hat{\mathcal{H}}^{\mathrm{KS}} \phi_\nu d^3\mathbf{r}, \tag{2.61}$$

where $\hat{\mathcal{H}}^{\mathrm{KS}}$ is the Kohn-Sham Hamiltonian.

If we expand a Kohn-Sham orbital ψ_i using the basis functions

$$\psi_i = \sum_{\nu=1}^{N_{basis}} a_{\nu i} \phi_\nu, \tag{2.62}$$

the Kohn-Sham equation $\hat{\mathcal{H}}^{\mathrm{KS}} \psi_i = \varepsilon_i \psi_i$ reduces to

$$\sum_{\nu=1}^{N_{basis}} (H_{\mu\nu} - \varepsilon_i S_{\mu\nu}) a_{\nu i}, \tag{2.63}$$

where $\{a_{\nu i}\}$ are the expansion coefficients.

The basis sets can be either extended, localized, or mixed (extended plus localized). Most DFT implementations utilize extended basis sets consisting of plane-wave basis or localized basis set consisting of numeric orbitals. In this section will briefly review the two basis set types and discuss their advantages and disadvantages.

2.7 Plane Wave Basis

For periodic systems, Bloch's theorem reduces the many-electron problem by restricting the problem to electrons in a unit cell and then determine the wave-functions of the other electron states using translational symmetry.

Bloch's theorem [5] states that the wave-function of an electron state in a periodic potential $V(\mathbf{r})$ is given by

$$\psi_{n\mathbf{k}} = \phi_n(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{r}} \tag{2.64}$$

where $\phi_n(\mathbf{r})$ is function that possesses the periodicity of the potential i.e. $\phi_n(\mathbf{r}) = \phi_n(\mathbf{r}+\mathbf{L})$, L is the length of the unit cell, n is the band index, and $\mathbf{k}$ is the wave vector in the first Brillouin zone. The periodicity of $\phi_n(\mathbf{r})$ allows us to express it as a Fourier series

$$\phi_n(\mathbf{r}) = \sum_{\mathrm{G}} c_{n,\mathrm{G}}\, e^{i\mathrm{G}\cdot\mathbf{r}}, \tag{2.65}$$

where the reciprocal lattice vector G and the real space lattice vector R satisfy the relation $\mathrm{R}\cdot\mathrm{G} = 2\pi s$, s is an integer and $c_{n,\mathrm{G}}$ are Fourier expansion coefficients. The wave-function now becomes

$$\psi_{\mathbf{k},n}(\mathbf{r}) = \sum_{\mathrm{G}} c_{n,\mathrm{G}+\mathbf{k}}\, e^{i(\mathrm{G}+\mathbf{k})\cdot\mathbf{r}}. \tag{2.66}$$

The reciprocal space representation of the Kohn-Sham equations 2.51 and 2.52 for Kohn-Sham orbital expressed in terms of plane-wave basis above is

$$\sum_{\mathrm{G}} \frac{1}{2}|\mathrm{G}+\mathbf{k}|^2\delta_{\mathrm{GG}'} + \bar{V}_{KS}(\mathrm{G}-\mathrm{G}')c_{j,\mathrm{G}} = \varepsilon_n c_{n,\mathrm{G}} \tag{2.67}$$

where $V_{KS}(\mathbf{r}) = \sum_{\mathrm{G}} \bar{V}_{KS}(\mathrm{G})e^{i\mathrm{G}\cdot\mathbf{r}}$. For an exact solution, infinitely many plane wave basis elements would be needed but fortunately, plane waves having high kinetic energies $\frac{1}{2}|\mathrm{G}+\mathbf{k}|^2$ are not of much importance so we have the freedom to truncate the kinetic energy at some cut-off E_{cut}, that is,

$$\frac{1}{2}|\mathrm{G}+\mathbf{k}|^2 \leq E_{cut} \tag{2.68}$$

2.7.1 Advantages of Plane Waves Basis

1. Plane wave basis are orthogonal and are independent of atomic positions. This makes it quite convenient to code in computer programs.

2. A single parameter, that is, the plane wave energy cut-off, E_{cut}, is used to determine the completeness of the basis set. Increasing E_{cut} improves the accuracy of the results.

2.7.2 Disadvantages of Plane Waves Basis

1. Plane waves are not suitable for non-periodic systems like molecules, clusters and surfaces. Application of plane waves to such systems require them to be placed in periodic supercells, leaving empty an space around the system. Since the entire supercell must filled with plane waves, it makes their application to such system computationally costly.

2. Despite their completeness, planes waves calculations are costly for large systems as several basis elements are needed to obtain well-converged results.

3. Because of the extended nature of basis functions, implementing algorithms that scales linearly the system size, the so-called $\mathcal{O}(N)$ methods, where N is the size of the system, is not possible.

2.8 Localized Basis

Localized basis sets are spatially compact atom-centered functions. They strictly confined real space i.e. orbitals that vanish in real space beyond a certain radius, and this increases the number of sparse matrix elements. The basis is obtained from a linear combination of the atomic orbitals (LCAO). If we expand the Kohn-Sham wave-function $\psi_i(\mathbf{r})$ in terms of a local basis set $\{\phi_j^\nu\}$ for a system with N_{atoms} atoms and N_j basis centered on each atom j, we get

$$\psi_i(\mathbf{r}) = \sum_{j=1}^{N_{atoms}} \sum_{\nu=1}^{N_j} a_{ij}^\nu \phi_j^\nu(\mathbf{r} - \mathbf{R}_j) \tag{2.69}$$

where j labels the index of the atomic site, ν labels the index of an orbital centered on a given atom. There are different kinds of localized basis orbitals. Some do have direct analytic forms (for example; Slater-type orbitals and Gaussian orbitals) and some are obtained using numerical

methods. The basis set that we have used for our work were generated using numerical atomic orbitals. We will discuss how these orbitals were obtained in detail in section 2.10.

We briefly list the advantages and disadvantages of localized basis functions.

2.8.1 Advantages of Localized Basis Orbitals

1. Generally, fewer localized basis orbitals are needed to expand a Kohn-Sham wave-function compared to plane wave basis. For example, solving DFT electronic structure calculations for amorphous silicon using plane-wave basis would require $N_{basis} \approx \mathcal{O}(10^3)$ plane-wave basis functions which reduces drastically to $\approx \mathcal{O}(10)$ localized basis functions if a localized basis set is used [57].

2. The strict confinement of the basis orbitals in real space generally leads to sparse Hamiltonian and overlap matrices. Sparse matrix elements can be used to achieve $\mathcal{O}(N)$ calculations i.e. linear scaling system size electronic structure calculations. $\mathcal{O}(N)$ techniques are almost impossible to implement with plane waves because of their extended nature.

3. Localized basis set offer a natural way of quantifying electron properties such as atomic charge, orbital population, bond charge, charge transfer, e.t.c through the Mulliken population analysis [58]. They are also particularly suitable for describing topological and chemical defects in a network [57].

2.8.2 Disadvantages of Localized Basis Orbitals

1. There is a difficulty in choice of localized basis set for a given problem since results of calculation are basis set dependent. For example; we will demonstrate later that accurate descrip-

tion of dangling bond defects in amorphous silicon using localized basis methods is basis set dependent.

2. Unlike plane wave calculations in which the accuracy of results is controlled by a single parameter (the kinetic energy cut-off E_{cut}), there are many parameters that have to adjusted to optimize localized basis calculations. This usually ends up in a very cumbersome and non-systematic trial and error process. In short, systematically improving plane wave basis is easier than devising a localized basis [57].

2.9 Ab Initio Pseudopotentials

Electrons in a solid or molecule are divided into core electrons and valence electrons. Core electrons occupy the filled inner shells and valence electrons occupy the outermost shell. Most chemical properties of solids depend mostly on valence electrons and slightly on the core electrons that are tightly bound to the nucleus. What core electrons mainly do is to screen the nuclear charges from the valence electrons. It is, therefore, tempting to eliminate them from electronic structure calculations and to replace them instead with an effective potential acting the valence electrons [38]. All-electron computations are prohibitively costly and so if core electrons can be cleverly removed from the calculations, it would dramatically reduce the enormous computational cost. Also, valence electrons states oscillate rapidly in the region the core electrons due to the orthogonality between valence electron states and core electron states [38]. A lot of basis functions are needed to expand the valence wave-function in the region of the core electrons and this makes the calculations extremely costly.

The expensive nature of all electron calculations be reduced in the following way:

1. Freeze the core electrons in electronic structure calculations and replace the Coulomb potential by an effective weak potential acting only the valence electrons. The weak potential should completely agree with the effective Coulomb potential $-Z_{val}/r$ outside the core region.

2. because core electrons are no longer considered, the orthogonality requirement for the valence states is relaxed and so the nodes can be removed by replacing with a pseudo wave-function that is smoothly varying (nodeless) inside the core region. We also require the pseudo and true wave-functions to completely agree outside the core region [38].

3. The weak potential in (1) will act on the pseudo-valence electron states.

This is the *pseudopotential* [59, 60] approximation in electronic structure. What it seeks to do is to replace the deep Coulomb potential of the nucleus and the effects of the tightly bound core electrons by a weak effective potential acting on a nodeless pseudo valence electron state. This greatly reduces the computational effort as fewer basis function are needed to expand the pseudo valence wave-function. A good pseudopotential should reproduce basic properties of an element and also have the ability to work well in different chemical environments. This implies that a pseudopotential so constructed must satisfy certain conditions to make it useful. The conditions will be given in section 2.9.3.

2.9.1 A Simple Qualitative Picture

In Fig. 2.1, we show a simplified picture for neutral Si atom ($1s^2\ 2s^2 2p^6\ 3s^2 3p^2$). Within the pseudopotential approximation, we remove all the core electrons that are bound to the nucleus leaving an effective ionic potential $-4e/r$ acting on the 4 valence electrons. Next replace with

effective ionic potential with a a weak pseudopotential acting on the 4 valence states. In constructing a pseudopotential, we require the pseudopotential to agree with the ionic potential outside the core region i.e. $r \geq r_c$. This can clearly be seen in the upper panels of the graph in Fig. 2.2, where we have plotted the pseudopotential for the $3s$ and $3p$ valence electrons states for the angular momenta $l = 0, 1$. Inside the core region, the pseudopotential is smooth and well-defined at the origin (it does not diverge as $r \to 0$). Similarly, in the lower panels of the graphs, we see a match between the radial components of the all-electron (AE) wave-function u_{nl}^{AE} and the pseudo wave-function u_l^{PS} outside the core. Due to the orthogonality between valence and core states and the strong nuclear potential within the core, the all-electron valence electron states oscillate rapidly, and as stated earlier, a lot basis functions are needed to capture the states within these regions. Clearly, we see that the valence pseudo wave-function u_l^{PS} varies smoothly inside the core region and its description in terms of basis size reduces significantly.

2.9.2 A Simple Quantitative Picture

In this section, we will show that it is possible to construct a pseudopotential $\hat{V}_{PS}$ such that the pseudo valence wave-function has the same eigenvalue as the true valence wave-function [38].

Consider an arbitrary valence wave-function $|\psi_v\rangle$. Decompose $|\psi_v\rangle$ into a smoothly varying pseudo valence wave-function $|\psi_{PS}\rangle$ and the rapidly varying orthogonal core states $\{|\psi_c\rangle\}$ with energy eigenvalues $\{E_c\}$

$$|\psi_v\rangle = |\psi_{PS}\rangle + \sum_{c \in core} \lambda_c |\psi_c\rangle \tag{2.70}$$

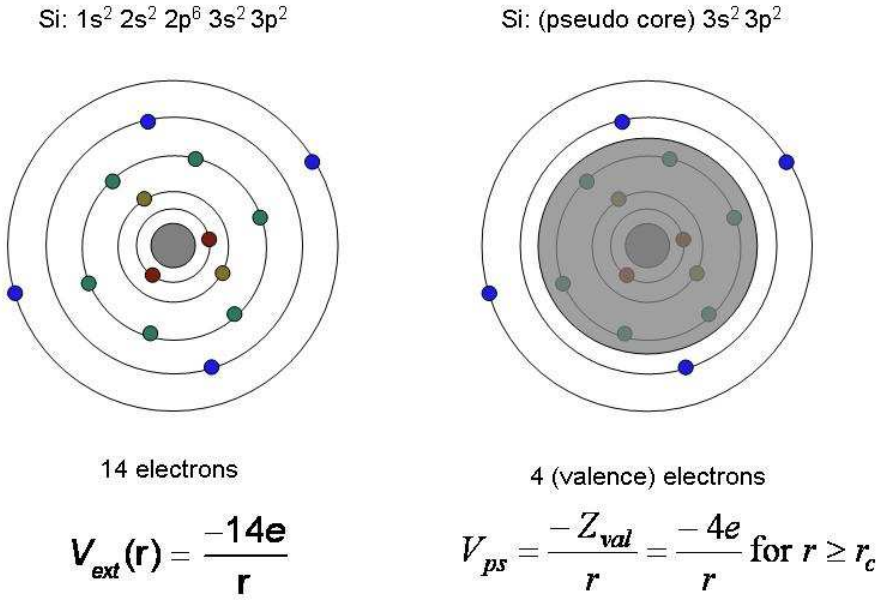

Figure 2.1: A picture for a neutral silicon atom. The tightly bound core electrons and the nucleus are replaced an effective ionic core, leaving the 4 valence electrons that mainly contribute to bonding and other properties in solids. $-4e/r$ should completely match the weak pseudopotential outside the core. The figure was taken from Ref. [61].

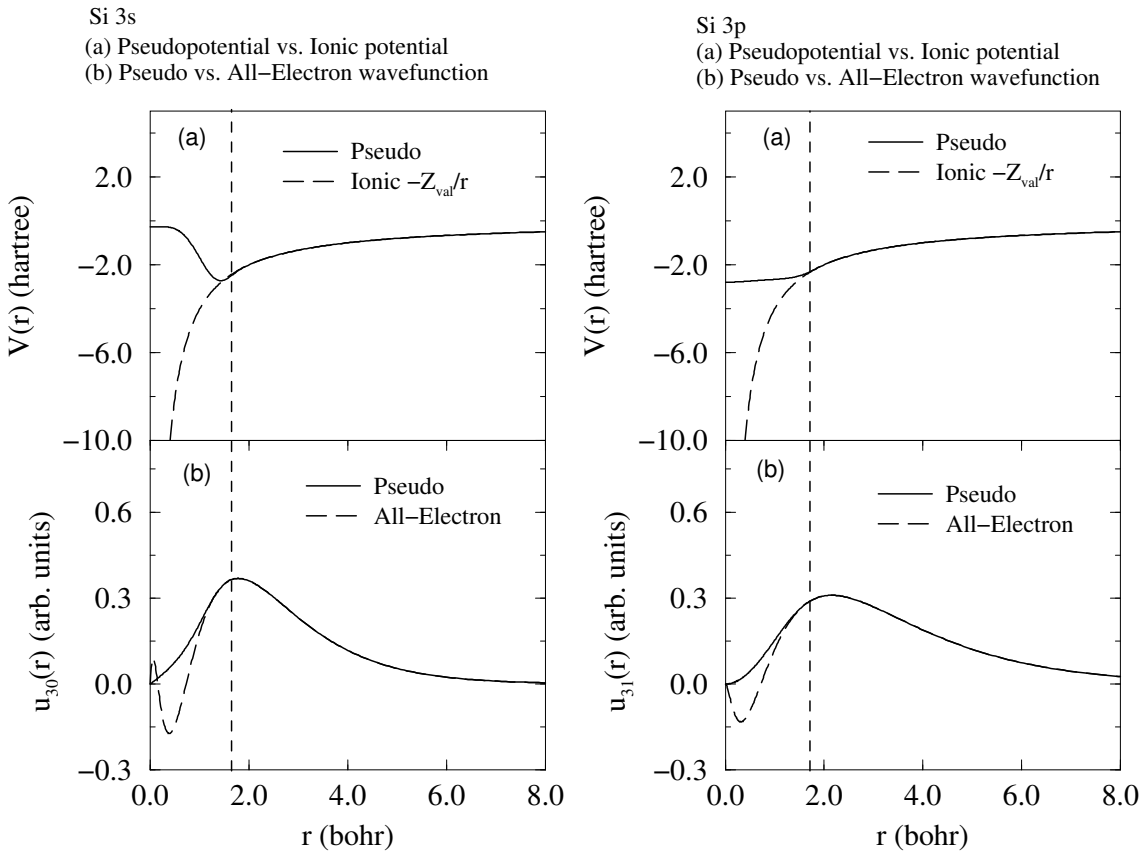

Figure 2.2: Plots for the pseudopotential versus the ionic potential $V(r) = -Z_{val}/r$ and the pseudo wave-function versus the all-electron wave-function for Silicon 3s and 3p respectively. The cut-off radius $r_c = 1.9$ bohr and it is denoted by the dotted vertical line. The data was generated using the `atom` program, a utility program in SIESTA.

Multiply each side with a core eigenbra $\langle\psi_{c'}|$ and use the fact that $\langle\psi_{c'}|\psi_c\rangle = \delta_{c,c'}$ and $\langle\psi_{c'}|\psi_v\rangle = 0$ to obtain $\lambda_c = -\langle\psi_c|\psi_{PS}\rangle$. Hence we get

$$|\psi_v\rangle = \left(\mathbf{1} - \sum_{c\in core} |\psi_c\rangle\langle\psi_c|\right)|\psi_{PS}\rangle, \tag{2.71}$$

where $\mathbf{1}$ is the unit operator. Now, since the left hand side of equation (2.71) satisfies Schrodinger's equation, that is, $\hat{\mathcal{H}}_e|\psi_v\rangle = E_v|\psi_v\rangle$ so must the right hand side,

$$\hat{\mathcal{H}}_e\left(\mathbf{1} - \sum_{c\in core} |\psi_c\rangle\langle\psi_c|\right)|\psi_{PS}\rangle = E_v\left(\mathbf{1} - \sum_{c\in core} |\psi_c\rangle\langle\psi_c|\right)|\psi_{PS}\rangle \tag{2.72}$$

This simplifies to

$$(\hat{\mathcal{H}}_e + \hat{V}_{PS})|\psi_{PS}\rangle = E_v|\psi_{PS}\rangle, \tag{2.73}$$

where

$$\hat{V}_{PS} = \sum_{c\in core} (E - E_c)|\psi_c\rangle\langle\psi_c|. \tag{2.74}$$

Hence we have constructed a pseudopotential such that eigenvalue of the pseudo valence wave-function equals that of the true valence wave-function. It must be stated that the construction of the pseudopotential and pseudo wave-function is not uniques. We can clearly see that the pseudo Hamiltonian is the original Hamiltonian plus $\hat{V}_{PS}$. $\hat{V}_{PS}$ cancels the effect of the strong Coulomb interaction so the resultant potential is a weak. This form of the pseudopotential is non-local (explicit dependence on the angular momentum for the core electrons) and energy dependent. Energy dependent pseudopotentials are not very helpful when the valence orbitals and core orbitals exhibit different symmetries [38]. It is convenient therefore to construct pseudopotentials that will act differently on electrons with different angular momenta.

2.9.3 Constructing Non-Local Ab Initio Pseudopotentials

We now briefly look into the construction of pseudopotentials for *ab initio* electronic structure calculations. Pseudopotentials have been studied extensively. Most Pseudopotentials are *norm conserving*. Then there are also *ultra-soft* pseudopotentials. Later, will explain the meaning of "norm conserving" and "ultra-soft pseudopotentials."

We now state the requirements that a pseudopotential should satisfy [63].

1. The pseudo valence wave-function $\psi_{lm}^{AE}(r)$ and the all-electron wave function $\psi_{nlm}(r)$ should have the same set of eigenvalues.

2. All-electron wave-function and pseudo wave-function should match outside the core.

$$u_l^{PS}(r) = u_{nl}^{AE}(r) \quad r \geq r_c$$

3. The logarithmic derivatives of all-electron and pseudo wave-function should match at $r = r_c$:

$$\frac{d}{dr} ln\left(\psi_{lm}^{PS}(r)\right)\Big|_{r=r_c} = \frac{d}{dr} ln\left(\psi_{nlm}^{AE}(r)\right)\Big|_{r=r_c}$$

.

4. The integrated charge inside the core for each wave-function should agree:

$$4\pi \int_0^{r_c} r^2 |\psi_{nlm}^{AE}(r)|^2 = 4\pi \int_0^{r_c} r^2 |\psi_{lm}^{PS}(r)|^2$$

Point (4) is the norm conservation requirement for the pseudopotential which simply means that the pseudo core charge must equal the true core charge. This guarantees that the pseudopotential scatters valence electrons and has the same phase-shift as the true potential. If the norm conservation condition is added to point (2), it leads to total charge conservation. Point (3) determines the *transferability* of the pseudopotentials. As we stated earlier, transferability refers to

the ability of the pseudopotential to work properly in atomic arrangements (e.g. a pseudopotential for silicon should perform equally well in SiO_2 and a-Si:H). Several schemes for generating norm conserving pseudopotentials exist [62, 63, 64, 65, 66, 67, 68, 69, 70]. Norm conserving pseudopotentials can be very accurate, but they are usually not very smooth.

The larger the cut-off radius r_c, the softer the pseudoptential and the smoother the pseudo wave-function but usually at the expense of norm conservation. Soft pseudopotentials can be described well with fewer basis functions. In the Vanderbilt ultra-soft pseudopotentials [71] method, the norm-conservation requirement is relaxed and pseudo wave-functions can be made as smooth as possible. This means that only few plane waves are needed to expand the pseudo wave-function. This surprisingly leads to pseudopotentials with very good transferability in plane wave calculations. For localized orbital methods (which we have used for all our first principles calculations), ultra-soft pseudopotentials are irrelevant.

Pseudopotentials are obtained from all-electron calculations. This is done by assuming a spherical screening approximation and solving the fully self consistent radial Kohn-Sham equations. Since the potential $V(r) = -\frac{Z}{r}$ acting on the electron has spherical symmetry we can write

$$\psi^{AE}_{nlm}(\mathbf{r}) = \frac{u^{AE}_{nl}(r)}{r} Y_{lm}. \tag{2.75}$$

Next, we solve the all-electron radial Kohn-Sham equations

$$\left[-\frac{1}{2}\frac{d^2}{dr^2} + \frac{l(l+1)}{2r^2} + V^{AE}_{KS}[\rho(\mathbf{r})] \right] u^{AE}_{nl}(r) = \varepsilon^{AE}_{nl} u^{AE}_{nl}(r) \tag{2.76}$$

We then parametrize the all-electron wave-function to obtain the pseudo wave-function $u^{PS}_l(r)$ using a well known scheme, (for example the Troullier-Martins scheme [70]) and invert Schrodinger's equation to obtain the screened pseudopotential

$$V^{PS,Screen}_l = \varepsilon^{PS}_l - \frac{l(l+1)}{2r^2} + \frac{1}{2u^{PS}_l(r)} \frac{d^2 u^{PS}_l(r)}{dr^2} \tag{2.77}$$

Finally subtract the Hartree and exchange-correlation energies to get the unscreened pseudopotential

$$V_l^{PS} = V_l^{PS,Screen} - V_H[\rho^{PS}] - V_{XC}[\rho^{PS}], \tag{2.78}$$

where $\rho^{PS}(\mathbf{r})$ is the pseudo-charge density obtained from pseudo valence wave-function $\psi_{lm}^{PS}(\mathbf{r})$.

For our work, we have used pseudopotentials generated using the Troullier-Martins parameterization [70]. The pseudopotentials are computed in an efficient manner using the separable form due to Kleinman and Bylander (KB) [72], in which matrix elements computational demand scales as $\mathcal{O}(N_{basis})$. One only has to make sure that the KB form of the pseudopotential does not lead to unphysical *ghost* states. A ghost state is an unphysical state above or below the valence band with radial nodes. A ghost state can be detected by a marked deviation of the logarithmic derivatives of the KB pseudopotential. Gonze *et al.* [73] have developed a method to track ghost states in pseudopotential calculations.

2.10 SIESTA: A Localized Basis Pseudoptential DFT Method

SIESTA is a DFT-based first principles pseudopotential code that utilizes a flexible linear combination of atomic orbitals (LCAO) as basis set [74, 75, 76]. It has the ability to accurately treat large systems in a relatively small time compared to plane wave codes. But the efficiency of a local basis method like SIESTA over plane waves-based codes does not necessarily make it better. Increasing the size of the basis set in a plane wave calculation guarantees the convergence of physical properties. Of course, increasing the number of localized basis functions improves the basis set but the level of basis needed can be system dependent and rarely diffuse functions may be required. However, once a well optimized LCAO basis set has been obtained, there is no reason why one can not perform highly accurate *ab initio* calculations with it [57].

SIESTA has been tested on hundreds of systems. Since we have used it exclusively for our *ab initio* electronic structure calculations, we find it necessary to discuss its main features. A detailed description of the method can be found in Ref. [76].

2.10.1 Basis Set

The basis orbitals centered on an atom I located at R_I is a product of numerical radial function and a spherical harmonic [76],

$$\phi_{Ilmn}(\mathbf{r}) = u_{Iln}(r_I)Y_{lm}(\hat{\mathbf{r}}_I), \tag{2.79}$$

where $\mathbf{r}_I = \mathbf{r} - \mathbf{R}_I$. In general, there will be several orbitals (labeled by n) with the same angular momentum l dependence but different radial dependence. In quantum chemistry, this is conventionally called a 'multiple ζ' basis. The individual basis elements are called first ζ, second ζ, third ζ, etc. and the basis set are referred to as single ζ (SZ)[7], double ζ (DZ), triple ζ, etc. We will briefly describe how the basis functions used in SIESTA are generated.

In the case of SZ basis set, the basis functions (that is the first ζ) are the eigenfunctions of an isolated atom in some confined within spherical potential wells of different shapes [77]. The first ζ is the solution to the radial Schrödinger equation

$$\left[-\frac{1}{2r}\frac{d^2}{dr^2}r + \frac{l(l+1)}{2r^2} + V_l^{PS}(r) + V_l^{\text{conf}}(r)\right] u_l(r) = \varepsilon_{nl}u_l(r), \tag{2.80}$$

where $V_l^{\text{conf}}(r)$ is a confining potential.

There several kinds of confining potentials [77, 78, 79]. Here, we will discuss two confining potentials that are used to strictly localize the basis functions. The first potential is due to Sankey and Niklewski [79] and the second is due Junquera *et al*. [77].

[7] By SZ basis, we mean a single radial function per angular momentum channel, and only for the valence orbitals with substantial electronic population

The confining potential in the Sankey-Niklewski approach is a hard-sphere potential of the form

$$V_l^{\text{conf}}(r) = \begin{cases} 0 & r < r_l^c \\ \infty & r \geq r_l^c, \end{cases} \tag{2.81}$$

where r_l^c is the orbital confinement radius. In general, one can specify r_l^c for each angular momentum channel. For strictly localized basis orbitals, the major problem is finding a well balanced and systematic way of defining the different cut-off radii r_l^c, since the accuracy and computational efficiency of the calculations depend on them. A scheme was proposed in which the confinement radii of different orbitals were defined by a single parameter, the orbital energy shift $\delta\varepsilon$, i.e., the energy increase suffered by the orbital due to the strict confinement [80]. It defines the radii in a well balanced way and allows the systematic convergence of physical quantities to the required precision.

$$\left[-\frac{1}{2r}\frac{d^2}{dr^2}r + \frac{l(l+1)}{2r^2} + V_l^{PS}(r) + V_l^{\text{conf}}(r)\right] u_l(r) = (\varepsilon_l + \delta\varepsilon)u_l(r) \tag{2.82}$$

Note that in the limit as $\delta\varepsilon \to 0$, the true pseudo-atomic orbitals, which are not strictly localized but decay to zero at infinity are recovered [80].

The hard nature of confinement potential the Sankey-Niklewski method [79] results in strictly localized basis but it suffers from a discontinuity in the first derivative of the basis function at r_l^c. The discontinuity is more pronounced for smaller values of r_l^c and tend to disappear for larger values [77]. Junquera *et al*. [77] proposed the generation of pseudo-atomic orbitals using a relatively *soft* confining potential

$$V_l^{\text{conf}} = V_l^0 \frac{exp\left(-\frac{r_l^c - r_l^i}{r - r_l^i}\right)}{r_l^c - r} \tag{2.83}$$

with the following properties: it vanishes in the core region, starts off at some radius r_l^i, and diverges smoothly at r_l^c. The pseudo-atomic functions resulting from this relatively soft potential are well localized and do not suffer from the discontinuity in their first derivatives at r_l^c. In both methods, the shape of the orbital can be changed by the ionic character of the atom. Orbitals tend to shrink in cations, and swell in anions [77].

We have briefly described above, how one can generate one radial function per angular momentum, the first ζ. The set of first ζ functions, the single ζ basis set, is sometimes not sufficient to accurately predict material properties. A second radial function can be added to each angular momentum channel in the SZ basis to obtain the DZ basis with two radial functions per channel. Several Schemes have been proposed to generate the second ζ. To generate the second ζ, SIESTA uses the idea of the *split valence* [28, 81] scheme in quantum chemistry.

In the split valence scheme adaptation by SIESTA, the second ζ function $u_l^{2\zeta}(r)$ must reproduce the tail of the first ζ function $u_l^{1\zeta}(r)$ within a matching radius r_m called the *split radius*, and continue smoothly toward the origin as $r^l(\alpha - \beta r^2)$, that is,

$$u_l^{2\zeta}(r) = \begin{cases} u_l^{1\zeta}(r) & r \geq r_m \\ r^l(\alpha - \beta r^2) & r < r_m \end{cases} \tag{2.84}$$

where the constants α, β are obtained by ensuring continuity and differentiability of $u_l^{2\zeta}(r)$ at r_m. r_m is determined by fixing the norm of $u_l^{1\zeta}(r)$ for $r \geq r_m$, the so called *split norm*. Typically, the split norm is set to 0.15 with exception of hydrogen whose split norm is set to 0.5. After $u_l^{2\zeta}(r)$ is generated, SIESTA actually defines the second ζ as

$$u_l^{1\zeta}(r) - u_l^{2\zeta}(r) = \begin{cases} 0 & r \geq r_m \\ u_l^{1\zeta}(r) - r^l(\alpha - \beta r^2) & r < r_m \end{cases} \tag{2.85}$$

instead of $u_l^{2\zeta}(r)$. As we can see, this gives rise to a strictly localized orbital having confinement radius r_m smaller than that of the single ζ orbital and hence reduces the number of non-zero matrix elements, without any loss of variational freedom. This process can be repeated easily to generate multiple ζ basis set.

To achieve well converged results, in addition to the pseudo-atomic valence orbitals, it is generally necessary to also include polarization orbitals, to account for the deformation induced by bond formations. Also, using pseudo-atomic orbitals for higher angular momenta is quite unsatisfactory because they tend to be too extended. Instead polarization orbitals can be included by applying a small electric field $\mathcal{E}$ in the z-direction and using first order perturbation theory [76]. We require that valence pseudo-atomic orbital $\phi_{lm}(\mathbf{r}) = u_l(r)Y_{lm}(\hat{\mathbf{r}})$ should not have any valence orbitals with angular momentum $l+1$. The polarized functions $u_{l+1}(r)$ are obtained from the lower angular momentum ζ functions $u_l(r)$ by solving the equation [76]

$$\left[-\frac{1}{2r}\frac{d^2}{dr^2}r + \frac{(l+1)(l+2)}{2r^2} + V_l^{PS}(r) - \varepsilon_l + \delta\varepsilon\right] u_{l+1}(r) = -ru_l(r). \tag{2.86}$$

It has been demonstrated elsewhere that addition of polarization orbitals to the multiple ζ sets results in well-converged calculations [76].

2.11 Electronic Structure Analysis

2.11.1 Mulliken Charge Analysis

As we stressed earlier, a localized basis set offers a natural way of quantifying electron properties such as atomic charge, orbital population, bond charge, charge transfer, whereas plane wave basis do not possess this advantage. Such a quantification could be very helpful if we want to know how electron and spin densities populate basis orbital centered on individual atomic sites

in solids and molecules. However, there is not a unique approach to analyzing orbital electron and orbital spin populations in atoms. Mulliken [58] and Löwdin [82] have introduced methods to perform such analysis. These methods are quite arbitrary. The Mulliken analysis is much simple and convenient for practical situations than the Löwdin analysis.

The Mulliken method assumes that electron population due to overlap is equally divided between the overlapping orbital. The arbitrariness in the method stems from this assumption. We will now demonstrate how the Mulliken charges are computed.

Consider the single particle Kohn-Sham orbitals $\{\psi_i(\mathbf{r})\}$. Suppose that each $\psi_i(\mathbf{r})$ is expanded in terms of basis orbitals $\{\phi_\mu(\mathbf{r})\}$. From the electron number conservation in equation (2.22), we have

$$\begin{aligned} N &= \int_{\mathbb{R}^3} \rho(\mathbf{r}) d^3\mathbf{r} = \sum_i \int_{\mathbb{R}^3} \psi_i^*(\mathbf{r})\psi_i(\mathbf{r}) d^3\mathbf{r} \\ &= \sum_i \sum_{\nu=1}^{N_{basis}} \sum_{\mu=1}^{N_{basis}} \int \phi_\mu^* \phi_\nu d^3\mathbf{r} = \sum_i \sum_{\nu=1}^{N_{basis}} \sum_{\mu=1}^{N_{basis}} a_{\mu i}^* a_{\nu i} S_{\mu\nu} \\ &= \sum_{\nu=1}^{N_{basis}} \sum_{\mu=1}^{N_{basis}} \left[\sum_i a_{\mu i}^* a_{\nu i} \right] S_{\mu\nu} = \sum_{\nu=1}^{N_{basis}} \sum_{\mu=1}^{N_{basis}} P_{\nu\mu} S_{\mu\nu} \\ &= \sum_{\nu=1}^{N_{basis}} (PS)_{\nu\nu} = \sum_{\nu=1}^{N_{basis}} q_{\nu\nu}, \end{aligned} \tag{2.87}$$

where

$$P_{\nu\mu} = \sum_i a_{\mu i}^* a_{\nu i}. \tag{2.88}$$

$q_{\nu\nu}$ is interpreted as the electron population in the basis orbital ϕ_ν. For an atomic site $\mathcal{A}$, we define the electron charge $Q_\mathcal{A}$ as the sum of the electron contributions from all basis orbitals centered on $\mathcal{A}$,

$$Q_\mathcal{A} = \sum_{\nu \in \mathcal{A}} q_{\nu\nu}. \tag{2.89}$$

We refer to $Q_{\mathcal{A}}$ as the Mulliken charge of the atom $\mathcal{A}$. We can clearly see that the Mulliken charge is meaningful if the basis functions are atom-centered, that is if they are localized.

In a similar fashion, we can construct the Mulliken charges at each atomic site for a single Kohn-Sham $\psi_i(\mathbf{r})$ by summing the electronic charge contributions over all basis orbital centered on that atom. In this case, the sum of the Mulliken charges over all the atomic sites equals unity.

Within the Mulliken analysis, the electron spin density is the difference between the sum of the up spin electron charges and the sum of the down spin electron charges for all occupied Kohn-Sham orbitals.

2.11.2 Quantifying Wave-Function Localization

We now describe ways of quantifying the degree of localization of an eigenstate of an electronic Hamiltonian.

The first measure of localization is the *inverse participation ratio* (IPR). The concept of IPR was introduced by Bell and Dean [83] in their study of vibrational modes in SiO_2. For a given Kohn-Sham eigenstate $\psi_i(\mathbf{r})$, the IPR measures wave-function localization of an electron state via the Mulliken charge concentration at each atomic site. It tells us how the Mulliken population at individual atomic sites participate in localizing an electronic wave-function. We define the IPR for a given eigenstate $\psi_i(\mathbf{r})$ as

$$\mathcal{I}\{\psi_i(\mathbf{r})\} = \sum_{n=1}^{N_{atoms}} [q_{ni}]^2 \tag{2.90}$$

where q_{ni} is the amount of Mulliken charge at an atomic site with index n and $\sum_{n=1}^{N_{atoms}} q_{ni} = 1$. For a completely localized state, all the Mulliken charge will sit at one site so $\mathcal{I} = 1$. For a completely extended state, the Mulliken charge population is the same at each atomic site so $\mathcal{I} = 1/N_{atoms}$. IPR is an *ad hoc* measure because it does not tell us anything about the real space

picture of a localized or extended eigen wave-function. Nevertheless, it has proved to be a powerful tool [84, 85, 86] for localization analysis in disordered systems.

We define a second quantity to characterize the spatial extent of the wave-function using the information entropy of a quantum state [87, 88]. This method has been used by Lewis *et al*. [89, 90] to measure the localization of DNA electron wave-functions. In this case, we define the measure of wave-function localization corresponding to a Kohn-Sham orbital ψ_i as

$$W(\psi_i) = \mathrm{e}^{S(\psi_i)}, \tag{2.91}$$

where $S(\psi_i) = -\sum_{n=1}^{N_{atoms}} q_{ni}\, ln(q_{ni})$ is the entropy and q_{ni} is defined above. In information theory, the entropy is a measure of disorder or randomness in the $\{q_{ni}\}$. As such, it has a more *a priori* justification for use as a measure of localization. This is especially true if we interpret localization as a *departure from randomness* [8]. We interpret $W(\psi_i)$ as the number atomic sites the state ψ_i can reach. It follows that W ranges from the total number of atoms for a uniformly extended state (that is, all the atoms can be reached) to unity for an ideally localized state (that is, only one atom can be reached).

The next measure of localization that we have used is a real space measure. In this case, we compute the spread of the Kohn-Sham wave-function $\psi_i(\mathbf{r})$ in real space as defined in the context of localized wave-functions in molecules [91] and maximally localized wannier functions in solids [92, 93]. This definition assumes that the wave-function has a well-defined center, and that the expected value of the position operator can be defined. We compute the spread as

$$\sigma^2 = [\langle \hat{\mathbf{r}}^2 \rangle - \langle \hat{\mathbf{r}} \rangle^2], \tag{2.92}$$

[8] The information entropy functional is the only form which consistently represents our state of knowledge in a probability theoretic sense. For a discussion of the $plog(p)$ functional versus p^2 (as for the IPR), see Jaynes (Ref. [88]) and references therein.

where $\hat{\mathbf{r}}$ is the position operator and the integration is performed over the supercell volume centered on a defect site. σ^2 tells us how much the wave-function decays in real space. The smaller the spread, the more localized the wave-function. We also compute the amount of the norm of the wave function q_Ω that is around the defect center, by integrating over a sphere Ω centered on the defect site and which has the same diameter as the lattice constant of the periodic supercell (so it is the largest sphere that fits into the supercell volume)

$$q_\Omega = \int_\Omega |\psi_i(\mathbf{r})|^2 d^3\mathbf{r}. \tag{2.93}$$

Clearly, a localized state will not completely sit in the sphere and therefore $q_\Omega < 1$. $(1 - q_\Omega)$ measures the amount of the norm of the wave-function outside the volume of integration Ω. For an extended state, $q_\Omega \approx 0.5$ since Ω encloses approximately 50% of supercell volume. In this case, we characterized the wave-function as localized if $0.8 \leq q_\Omega < 1$.

Chapter 3

Review of Existing Computer Simulation Techniques for *a*-Si

Computer simulations is a primary tool for the theoretical studies of physical, chemical and biological systems. It is mainly used to test the validity of proposed theoretical models, to confirm experiments, and make predictions yet to be realized in experiment. The advent of fast computers has dramatically increased the systems sizes and accuracy of results pertaining to the physics of disordered systems. However, there will always be practical limits to what properties of the system can be simulated, taking into account the nature of the algorithm and computation time.

Broadly speaking, there are two families of computer techniques for modeling amorphous silicon: deterministic methods and stochastic methods. In deterministic methods, one starts with a set of initial conditions and advance the system by integrating the equations of motion with respect to time. Starting with some initial trajectory, the future trajectory is completely determined. In most deterministic methods, the forces acting on the atoms are determined using Newton's second law. Deterministic methods that use classical Newtonian dynamics are known as *Molecular Dynamics*

(MD) [94, 95, 96]. On the other hand, the stochastic approach to computer modeling mainly involves probabilistic methods that takes the configuration from one point in phase space to another such that some condition is satisfied (e.g. using a stochastic scheme to advance a system from one conformation to another until it reaches the minimum of the inter-atomic potential describing the system). The most common stochastic modeling method is the *Monte Carlo* (MC) method. There are different types of Monte Carlo methods [94, 95, 96]. Before discussing MD and MC simulation methods for *a*-Si, we describe below, different inter-atomic potentials used in modeling and discuss their pros and cons.

3.1 Inter-atomic Potentials for *a*-Si

The most important ingredient in modeling amorphous silicon is the potential that describes the inter-atomic interactions. Ideally, we want a potential that

1. can accurately reproduce key structural, electronic and vibrational properties as measured in experiment,
2. has the ease model systems with sizes closer to the bulk,
3. is transferable, i.e. it should work well in different chemical environments, and
4. is computationally fast.

Different prescriptions of potentials for modeling *a*-Si exists. They range from highly accurate and transferable first principles potentials to empirical potentials. Each potential has its strengths and practical limitations.

3.1.1 Empirical Potentials

Empirical potentials estimate the interactions between the ions in a system. They are simple and straightforward to implement in a computer code and computationally cheap. In comparison to *ab initio* and tight-binding schemes, they have the ability to model very large systems and are several orders of magnitude faster. Empirical potentials usually have a parameter-dependent analytic form, where the parameters are obtained by fitting the potential to experimental data for different phases of the system and also to highly accurate first principles calculations. By so doing, the potential inherits the ability to describe atomic bonding properties. However, they do not give any information about the electronic structure. Most empirical potentials suffer from transferability. Nevertheless, they provide a good starting point for high quality quantum mechanical calculations.

For a-Si, the most frequently used empirical potentials are the Keating potential and Stillinger-Weber potential. We will briefly describe below the features of each potential.

3.1.2 Empirical Potential I: Keating Potential

The Keating potential [97] is an old and very simple short-range potential whose minimum corresponds to the diamond crystal lattice for Si. The potential consists of two-body (bond stretching) and three-body (bond bending) terms given by

$$V = \frac{3\alpha}{16d^2}\sum_{\langle i,j\rangle}(|\vec{r}_{ij}|^2 - d^2)^2 + \frac{3\beta}{8d^2}\sum_{\langle j,i,k\rangle}(\vec{r}_{ij}\cdot\vec{r}_{ik} - \frac{1}{3}d^2)^2 \tag{3.1}$$

where α is the bond stretching constant, β the bond bending constant, and d is the zero temperature bond length of the crystalline silicon. For a-Si, $\alpha = 2.965$ eV/Å^2, $\beta = 0.285\alpha$, and $d = 2.35$ Å. $\langle i,j\rangle$ means that atom i is bonded to atom j and $\langle i,j,k\rangle$ means atom i is bonded to both atoms j and k. Since crystalline silicon is perfectly tetrahedral, the potential requires every atom to have four

nearest neighbors. In a configuration with N atoms, there is a total of $2N$ bonds and $6N$ bond angles and the energy evaluation scales as $\mathcal{O}(N)$. This is because the potential is short ranged. The $\mathcal{O}(N)$ scaling implies that the Keating potential can be used on very large systems. The shortcoming of the potential is that it requires every atom to have exactly four nearest neighbors and that two atoms can interact if and only if they are bonded. This can lead to unphysical situations in which atoms that are not bonded to each other can be as close as possible without paying a price in the energy. The strict bonding requirement also makes the potential unsuitable for molecular dynamics simulations of dynamic processes.

3.1.3 Empirical Potential II: Stillinger-Weber Potential

The Stillinger-Weber potential (SW) [98] is a more realistic empirical potential for silicon. Unlike the Keating potential, interaction between the ions goes beyond the first shell (up to the second shell) and do not require an atom to have a fixed number nearest neighbors making it more flexible. It is written as

$$V = \sum_{\langle i,j \rangle} v_2(\vec{r}_{ij}) + \sum_{\langle j,i,k \rangle} v_3(\vec{r}_{ij}, \vec{r}_{ik}) \tag{3.2}$$

where the two body term v_2 is given by

$$v_2(\vec{r}_{ij}) = \epsilon A \left[B \left(\frac{|\vec{r}_{ij}|}{\sigma} \right)^{-p} - 1 \right] \exp\left(\frac{1}{|\vec{r}_{ij}|/\sigma - a} \right) \Theta(a - |\vec{r}_{ij}|/\sigma) \tag{3.3}$$

and the three body term is given by

$$\begin{aligned} v_3(\vec{r}_{ij}, \vec{r}_{ik}) &= \epsilon\lambda \exp\left(\frac{\gamma}{|\vec{r}_{ij}|/\sigma - a} + \frac{\gamma}{|\vec{r}_{ik}|/\sigma - a} \right) \\ &\times \quad (\cos\theta_{ijk} + \frac{1}{3})^2 \; \Theta(a - |\vec{r}_{ij}|/\sigma)\, \Theta(a - |\vec{r}_{ik}|/\sigma) \end{aligned} \tag{3.4}$$

Here,

$$\Theta(x - x_0) = \begin{cases} 0 & \text{if } x < x_0 \\ 1 & \text{if } x > x_0 \end{cases} \tag{3.5}$$

is the unit step function, θ_{ijk} is the angle between the bonds $\vec{r}_{ij}$ and $\vec{r}_{ik}$, and A, B, a, p, ϵ, σ, λ, and γ are the fitting parameters. In the original SW potential, the parameters were obtained by fitting to the liquid and crystalline phases of silicon so it does not guarantee the right description of the amorphous phase. Recently, Vink *et al.* [99] fitted the potential directly to the amorphous phase. We will refer to this as the modified SW potential (MSW). The ratio of the three-body coupling parameter $\epsilon\lambda$ to the two-body coupling parameters ϵA in the MSW potential increases by 50% when compared to the same ratio in the original SW potential. The modified potential yields amorphous silicon configurations that agrees with experimental data compared to the original potential. MSW potential also produces the correct experimental transverse acoustic peak and transverse optical peak in the phonon density of states. The values of parameters for each potential is given in table 3.1.

In all empirical methods, the forces are purely ionic. The force experienced by an ion i at position $\vec{r}_i$ is given by

$$\vec{F}_i = -\nabla_{\vec{r}_i} V = -\frac{\partial V}{\partial \vec{r}_i} \tag{3.6}$$

3.1.4 Semi-Empirical Tight-Binding

Semi-empirical tight-binding (SETB) methods lie between accurate *ab intio* and empirical methods. They are computationally cheap compared to *ab initio* methods and expensive compared to empirical methods. The phrase "semi-empirical" stems from the fact that the Hamiltonian is

Table 3.1. Parameters for the original Stillinger-Weber (SW) and modified Stillinger-Weber (MSW) potentials for amorphous silicon

Parameter	SW	MSW
A	7.049556277	7.049556277
B	0.6022245584	0.6022245584
ϵ (eV)	2.16826	1.64833
σ (Å)	2.0951	2.0951
λ	21.0	31.5
a	1.8	1.8
γ	1.2	1.2
p	4	4

empirical but at the same time has an ad hoc inbuilt quantum mechanical information so that both the ionic and electronic structure can be modeled. The quantum mechanical information is usually introduced in the Hamiltonian via parameters obtained from fitting to high quality theoretical and experimental quantum mechanical information. For systems that are large for *ab initio* methods to handle, the next obvious choice is SETB methods. Unlike *ab initio* methods, most SETB methods do not properly treat electron exchange and correlations. Also SETB methods can suffer from transferability in sense that SETB parameters obtained by fitting to one system may not work in another system.

The idea of the tight-binding scheme was first introduced by Slater and Koster [100]. We will briefly demonstrate how SETB methods work using a simple quantitative picture in which localized basis functions are used to expand the electronic eigenstates. Start with a single particle energy eigenstate ψ_i and expand it in terms of localized basis orbitals $\{\phi_{j\alpha}\}$:

$$|\psi_i(\vec{r})\rangle = \sum_{\alpha,j} a_{j\alpha} |\phi_{j\alpha}(\vec{r} - \vec{R}_\alpha)\rangle, \tag{3.7}$$

where α is the index of an atom at position $\vec{R}_\alpha$ and j is the index of an orbital centered on the atom. Let us use silicon for an example. Suppose we are working with a minimal atomic-like basis set (as described in section 2.10). Here, we have assumed that core electrons can be neglected and only valence electrons are of interest. Then there will be four orbitals (e.g. orbitals with s, p_x, p_y, and p_z symmetry) centered on each atom. So in a system with N atoms, a total of $4N$ atom-centered basis functions will be used to expand a given eigenstate. Next, solve the time independent Schrödinger equation

$$\hat{\mathcal{H}}|\psi_i\rangle = \varepsilon_i|\psi_i\rangle. \tag{3.8}$$

Here, $\hat{\mathcal{H}}$ is a *parameterized* single particle Hamiltonian with matrix elements

$$H_{j\alpha,j'\alpha'} = \langle \phi_{j\alpha}(\vec{r} - \vec{R}_\alpha) | \hat{\mathcal{H}} | \phi_{j'\alpha'}(\vec{r} - \vec{R}_{\alpha'}) \rangle \tag{3.9}$$

In most SETB schemes, the overlap between the basis functions is neglected (i.e. basis functions are orthogonal). There are several parameterization schemes for the tight-binding Hamiltonian for silicon. The first orthogonal Hamiltonian is due to Chadi [101] who used it compute the band structure and EDOS of diamond and zincblende Si and Si surfaces. Following Chadi, several orthogonal parameterizations schemes [102, 103, 104, 105, 106] as well as non-orthogonal schemes [107, 108, 109] have been developed.

Within semi-empirical tight-binding, the total energy is the sum of the single electron energies plus sum of pairwise potential energies

$$E = \sum_i \varepsilon_i + \frac{1}{2} \sum_i \sum_j V_{ij}(|\vec{R}_i - \vec{R}_j|). \tag{3.10}$$

The force experienced by an atom at position $\vec{R}_I$ is obtained using the Hellmann-Feynman theorem [110]

$$\vec{F}_I = -\frac{\partial E}{\partial \vec{R}_I}. \tag{3.11}$$

For a detailed explanation of tight-binding schemes, the reader may refer to an extensive review article by Goringe *et al*. [111].

3.1.5 Ab Initio Potentials

The *ab initio* method has been described in chapter 2. Forces computed using *ab initio* potentials are more accurate and reliable than empirical and semi-empirical forces. Also, unlike empirical and semi-empirical schemes, *ab initio* potentials are generally transferable. There

are two major approaches to *ab initio* modeling: Born-Oppenheimer (BO) methods and extended Lagrangian (EL) method. We briefly discuss below, the two approaches within the Kohn-Sham formulation of density functional theory discussed in section 2.3.

Within the BO approach, one has to fully compute the electronic structure at each time step to obtain information about the potential energy surface for a given ionic configuration. Let us briefly explain how this works within the Kohn-Sham DFT discussed in chapter 2. Suppose that $\Phi(\{\vec{R}_I\})$ is the potential energy surface corresponding to a configuration of ions, where $\{\vec{R}_I\}$ are the the ionic coordinates. Within the Kohn-Sham DFT, we obtain $\Phi(\{\vec{R}_I\})$ by minimizing the total energy functional $E(\{\vec{R}_I\},\{\psi_i\})$ with respect to the set of singly-occupied orthonormal Kohn-Sham electron wave-functions $\{\psi_i\}$, i.e.

$$\Phi(\{\vec{R}_I\}) = \mathrm{Min}_{\{\psi_i\}} E(\{\vec{R}_I\},\{\psi_i\}) \tag{3.12}$$

We refer to $\Phi(\{\vec{R}_I\})$ as the Born-Oppenheimer potential energy surface. Next, assume that the ions obey the laws of classical mechanics (Newton's law to be precise). Then force $\vec{F}_I$ on ion at position $\vec{R}_I$ is given by

$$\vec{F}_I = -\frac{\partial \Phi}{\partial \vec{R}_I}. \tag{3.13}$$

In short, to compute the time evolution of the system of ions using the BO approach, one has to first compute the potential energy surface from the total energy functional (by minimizing the total energy with respect to the electron density), followed by the computation of the ionic forces from the potential energy at each time step, which are then used to compute the ionic trajectories.

The EL approach, which was introduced by Car and Parrinello [112], exploits the continuity of the dynamical trajectories i.e. the fact that adjacent configurations are very close. The Car-Parrinello EL approach uses a fictitious classical dynamical system in which the ionic coordi-

nates $\{\vec{R}_I\}$ and the single particle Kohn-Shame orbitals $\{\psi_i\}$ both act as degrees of freedom [1]. The Lagrangian for the system is constructed in such a way that the two resulting equations of motion corresponds to Newton's equation for the ions and for the electrons. This results in the simultaneous integration of the ionic and electronic motion. However, the electron dynamics do not correspond to the real quantum mechanical time evolution. All that the Car-Parrinello method does is to avoid solving the Kohn-Sham equations explicitly and provide a new way to carry out the energy minimization. It therefore finds a way, albeit artificially, to include the electronic motion to the time evolution of the ions. By varying ψ_i for a fixed ionic position a single minimum on the energy surface is found, which should in principle equal the minimum BO potential energy surface Φ. We will not bother readers with the details of the method since we have not used it in this dissertation. For more on the basics of the Car-Parrinello approach and how it works in practice, the reader may refer to seminal papers by Car *et al.* in Refs. [112, 113].

The major problem with *ab initio* modeling is that it is computationally expensive. Therefore the system sizes accessible to such methods can be many orders of magnitude smaller that empirical methods. To realistically simulate device-sized systems, we need systems containing tens of thousands of atoms. Current *ab initio* techniques, although accurate, cannot treat such systems.

As a precursor to our modeling methodology for *a*-Si, which would be presented in the next chapter, we will review existing computer modeling schemes for amorphous silicon in the remainder of this chapter. We will briefly present each method and state their advantages and disadvantages. However, we will place special emphasis on the WWW technique that yields high quality and large atomistic models.

[1] The dynamical system is fictitious due to the fact that an electron is assigned a fictitious mass, resulting in a fictitious kinetic energy in the Lagrangian

3.2 Molecular Dynamics Simulations

3.2.1 Statistical Eechanics Ensembles

As we explained in the introduction to this chapter, an MD simulation is simply the time evolution of a system of ions in which the atomic positions are advanced using Newton's second Law. Generally, one is interested in the macroscopic properties of a system of interacting particles. Since MD simulates microscopic processes, we should find a way to extract macroscopic information from the microscopic information.

Statistical mechanics relates the thermodynamical properties of a microscopic system such as atomic positions and momenta with the macroscopic informations in the system such as the free energy, heat capacities, pressure etc. [114]. The explicit description of macroscopic properties of a system is non-trivial. However, it is easier to describe the macroscopic properties of a system in equilibrium. To do this, we fix some parameters to describe a statistical mechanical ensemble [2]. The macroscopic parameters that define an ensemble are classified into two categories, intensive quantities (i.e. system-size independent parameters, namely pressure (P), temperature (T), and chemical potential (μ)) and extensive quantities (i.e. system-size dependent parameters, namely total energy of the system(E), total volume of the system (V), and total number of particles (N)).

The most basic ensemble is the *micro canonical* or (N,V,E) ensemble in which N, V, and E are always constant. Physically, the (N,V,E) ensemble corresponds to an isolated system surrounded by heat insulating walls so that energy does not leave or enter the system. Another ensemble is *canonical* or (N,V,T) ensemble in which N, V, and T are always constant. Physically, the (N,V,T) ensemble corresponds to a system in thermal contact with a comparably large heat

[2] A statistical mechanics ensemble is a defined as a set configurations in phase-space characterized by the set of macroscopic quantities

reservoir. There is no exchange of particles between the system and reservoir but energy transfer is allowed to keep the temperature of the system constant [114]. This results in large fluctuations in the energy of the system with respect to time. Finally there is the (N,P,T) ensemble which corresponds to a system coupled to a heat and pressure reservoir. There are volume fluctuations of the system to ensure a balance between the external reservoir pressure and the internal system pressure [114]. In the thermodynamic limit, the equilibrium properties of all ensembles will be the same [3]. Detailed descriptions of the statistical mechanics ensembles and related phenomena can be found in standard statistical mechanics texts [115, 116] and extensive review articles [114].

3.2.2 MD Simulations in Practice

Now that we have briefly described the ensembles that one can use in MD simulations, we will describe how a typical MD simulation of an ensemble proceeds. Generally, an MD simulation consists of three major steps [95]:

Step 1: **Initialization**

Begin with a set of initial atomic positions $\{\vec{r}_i\}$ that is representative of the system we want to simulate. Also specify the initial particle velocities [4] $\{\vec{v}_i\}$ and shift the velocities by the center of mass velocity such that the net momentum vanishes. Depending on the ensemble, one may have to specify the initial temperature T and other parameters when necessary.

[3] The thermodynamic limit is the limit of an infinite system such that the density N/V is constant

[4] Usually, one draws the initial velocities from the Maxwell-Boltzmann distribution or from the uniform distribution. The initial choice does not really matter since the velocities will follow the distribution that characterizes the ensemble in the equilibrium state.

Step 2: **Computations and Equilibration**

Compute the forces acting on each atom and integrate Newton's equations of motion over a small time interval δt to obtain the new atomic positions and velocities. We compute the averages of dynamical and static macroscopic observables of interest from the positions and velocities. Keep repeating this step for several time intervals until the system reaches equilibration. Equilibration simply means that the system is in a state whereby certain properties are no longer varying with time.

Step 3: **Production**

After the system reaches an equilibrium state, we proceed to analyze the trajectories to compute observables of interest such as the radial distribution function, diffusive properties etc.

There are several flavors of MD algorithms for simulations in different ensembles. The reader may refer well-written texts on computer simulations for detailed algorithms for different MD schemes [94, 117, 96, 118, 119, 120].

We would like to state that choosing the time interval δt is very crucial to the success of the simulation. If δt is too large, the particles will move very far along the trajectory and hence the atomic motion would be poorly simulated. If δt is too small, the simulation becomes time-consuming as several time steps will be required to achieve equilibration [95]. Generally, δt is chosen to be a tiny fraction of timescales for atomic processes such as vibrations and collisions. For most solids and and liquids, this translates to time intervals of the order of a femtosecond. Hence it is not feasible to use MD to properly simulate atomic processes that occur on timescales greater than a microsecond.

3.2.3 Molecular Dynamics Modeling of a-Si

Now that we have briefly described what MD is all about, we would like to describe how amorphous silicon models were generated using MD and the feasibility of the method. The most common MD scheme for modeling amorphous silicon is known as the "cook and quench" (CQ) method. The starting point of most CQ schemes is crystalline silicon. Thermal disorder s introduced at a high temperature so that the system completes melts. In the liquid state, the energy of the system is quite high. The liquid is allowed to cook until it reaches equilibration. The liquid is then quenched rapidly from the high temperature to a desired low temperature to form a solid. The solid can either be a crystal (or polycrystal) or amorphous. According to Elliot [4], the distinguishing feature of producing amorphous solids is that the amorphous solid is formed by the continuous hardening (increase in viscosity) of the melt. Hence, a criterion for generating an amorphous solid from the liquid is that the cooling rate should be fast enough to introduce topological irregularities and avoid crystal formation. In computer simulations, quenching is achieved by decreasing the velocity at each time step. The simplest way to drive the system at an instantaneous temperature T_i to some target temperature T is reduce the velocity by a factor ν [117]

$$\nu = \sqrt{1 - \frac{\delta t}{\tau}\left(1 - \frac{T}{T_i}\right)}, \tag{3.14}$$

where δt is the length of the time step and τ is the *relaxation time*, which determines how rapid the system is to be driven to reach the target temperature T. Choosing the relaxation time is very important in generation of amorphous solids and it currently a subject of intense research [122]. Once the cooling process is completed at some desired temperature (usually room temperature), the configuration is finally quenched at $T = 0$, which simply is the potential energy surface minimization.

The CQ method simulates the process called *annealing* which is used in experimental laboratories and industries. Samples are annealed through heating to get rid of any artifacts that are frozen in the material and rapidly cooled. The annealing process usually takes minutes and sometimes hours. This translates to cooling rates of $10^{-2} - 10^{3}$ Ks^{-1}. One problem with the CQ method is immediately apparent; the cooling rate, which is typically around $10^{12} - 10^{15}$ Ks^{-1} is too high compared to experiments.

The CQ method is not be the best scheme to model of amorphous silicon but it works well for many systems e.g. SiO_2. Several factors are responsible for this. First, the extremely fast cooling rates does not allow the liquid to fully relax into an amorphous phase. The second problem is the starting liquid phase is more dense than the amorphous solid phase since the average coordination number of the liquid is 6 compared to 4 in the amorphous phase. Most MD routines do not adequately treat the density differences (this implies volume expansion and contraction since the particle number is fixed or equivalently, pressure is controlled) [123]. Together, these problems in the CQ method lead to models of amorphous silicon whose structural and optical properties are generally unsatisfactory compared to experiment

The structural defects in CQ models of *a*-Si lead to localized defect states in the electronic bang gap. CQ methods do not have any feature to enforce bond coordination. Another feature of CQ models is the strain in the network. Primary measures of network strain are the standard deviation of the bond angles and the total energy. The average bond angle of CQ model of *a*-Si lies in the range 13.6°–16.3° compared to experiments which is in the interval 9.4°–12.0°. The bond angle standard deviation leads to significant broadening of the second peak in the RDF and introduce localized gap and band tail states. Generally in MD simulations, the final configuration sits in a local minimum of the energy landscape which is far away from the global minimum. This leads to

Table 3.2. Structural properties of quench-melt CRN models of *a*-Si. The models are labeled as follows: SCP-Štich, Carr, and Parrinello, CGM-Cooper, Goringe, and McKenzie, KL-Kim and Lee, UK-Urbassek and Klein, KRR-Kluge, Ray , and Rahman, LL-Luedtke and Landman, and JBK-Justo *et al*. N is the system size, n_c is the average coordination, and $\Delta\theta$ is the root mean square deviation of the bond angles. The table is taken from Ref. [123].

Model	SCP	CGM	KL	UK	KRR	LL	JBK
Ref.	[124]	[125]	[126]	[127]	[128]	[129]	[130]
Potential	DFT	DFT	SETB	SETB	SW	SW	EDIP
N	64	64	64	128	216	588	1728
$\Delta\theta$ (deg)	15.5	15.2	16.3	–	13.6	14.7	14.0
n_c	4.03	3.96	4.28	4.001	4.12	4.12	4.054

high energy configurations. In short, MD modeling methods are not the best as far as amorphous silicon is concerned. Serge Nakhmanson [123] has carefully documented the structural and optical properties of CQ models of amorphous silicon in his thesis so the reader may refer to his work for thorough explanations of the individual approaches. Here, we summarize the structural properties of CQ models in table 3.2. We generally observe the networks have a high bond angle deviation and significant coordination defect concentration.

3.3 Monte Carlo Simulations

Monte Carlo (MC) methods are stochastic-based simulation methods. Unlike MD which uses the atomic forces to advance particles in a system over a period of time, MC is not a time evolution scheme. In MC simulations, the system "hops" from conformation to conformation. The "hopping' probability usually depends on the energy difference between the initial and final configurations as well as the simulation temperature. One advantage MC has over MD is that it can search for different trajectories for system and even jump over energy barriers that would be otherwise difficult for MD to overcome. For example, if we want to find the global minimum potential energy of a system using MD, the system may get trapped in a local minimum on the energy surface and it might take a considerable amount of time for the configuration to jump out the local minimum and settle in the global minimum. But MC can easily push the system out of the local minimum and eventually find the global minimum. Also, MC methods are useful for computing multidimensional integrals that is otherwise difficult to compute analytically.

For our work, we are interested in Monte Carlo methods based on *Markov chains*. By definition, a sequence of successive events $\Gamma_0, \Gamma_1, \cdots, \Gamma_n$ is a Markov chain for any $n \geq 1$ if the outcome of any event depends on *only* the preceding event, i.e., an immediate future event is uniquely determined by the present event, regardless of the past. Mathematically, this translates to

$$\mathcal{P}(\Gamma_n|\Gamma_{n-1}, \Gamma_{n-2}, \cdots, \Gamma_0) = \mathcal{P}(\Gamma_n|\Gamma_{n-1}), \tag{3.15}$$

where $\mathcal{P}(A|B)$ is the conditional probability that an event A will occur given that the event B has occurred [94]. In this case, we refer to $\mathcal{P}(A|B)$ as the *transition probability* from state A to state B and would be henceforth denoted by $\mathcal{P}(A \rightarrow B)$ for clarity. For a Markovian chain to reach equilibrium, it must satisfy *ergodicity* and *detailed balance*. Ergodicity means that a state Γ_j can be

reached from a state Γ_i in a finite number of steps. Detailed balance means that the average number times that a state Γ_j can be reached from a state Γ_i is equal to the average number of times Γ_i can be reached from Γ_j [94].

We will now give a general outline of an MC algorithm.

Step 1: Begin with an initial state Γ_{old}

Step 2: Randomly change the initial state to obtain a new state Γ_{new}. For example, we may randomly displace atomic positions, deform the supercell to change its volume or use particle annihilation and creation operators to change the total number of atoms

Step 3: Compute the probability of transition from Γ_{old} to Γ_{new}, $\mathcal{P}(\Gamma_{\text{old}} \rightarrow \Gamma_{\text{new}})$. If $\mathcal{P}(\Gamma_{\text{old}} \rightarrow \Gamma_{\text{new}}) = 1$ then replace Γ_{old} with Γ_{new} and go to step 2. Otherwise, go to step 4 below.

Step 4: Generate a random number $r \in [0, 1]$. If $\mathcal{P}(\Gamma_{\text{old}} \rightarrow \Gamma_{\text{new}}) < r$ then replace Γ_{old} with Γ_{new} and go to step 2. Otherwise discard Γ_{new} and begin again from step 2.

We keep repeating steps 2 through 4 until convergence is achieved.

One of several Monte Carlo schemes is the Metropolis Monte Carlo method [121]. In this case, probability of transition follows the Boltzmann distribution [5] given by

$$\mathcal{P}(\Gamma_{\text{old}} \rightarrow \Gamma_{\text{new}}) = \begin{cases} 1 & \text{if } \Delta E < 0 \\ exp(-\beta \Delta E) & \text{if } \Delta E \geq 0 \end{cases} \tag{3.16}$$

where the energy difference $\Delta E = E(\Gamma_{\text{new}}) - E(\Gamma_{\text{old}})$, $\beta = \frac{1}{k_B T}$, T is the temperature of the system, and k_{B} is the Boltzmann constant. Because Monte Carlo methods are not time dependent, it can not be used to simulate dynamical (time-dependent) quantities. Nevertheless, it is quite useful and flexible as a modeling scheme compared to MD.

3.4 Metropolis Monte Carlo Modeling of a-Si

As we explained earlier, MC simulations afford us the opportunity to sample the energy landscape of a given system. It is even more flexible if one simulates a Markov process using the Metropolis algorithm. All that is needed is the energy before and after an event. A well known metropolis MC method for modeling amorphous silicon, silicon based compounds, amorphous germanium, and amorphous carbon is the Wooten, Winer, and Weaire (WWW) bond switching algorithm [131] and the modifications to the algorithms due to Barkema and Mousseau [132]. The modified algorithm is extremely fast compared to the original and it can be used to generate extremely large and high quality models of amorphous silicon of the order of 10^5 atoms. We will first describe the WWW bond switching algorithm in its original form and then focus on the improved version.

[5] Boltzmann's distribution describes the energy distribution of N possible states a system in thermal equilibrium at temperature T. The Boltzmann probability p_i that a system is in a state i with energy E_i is given by $p_i = \frac{exp(-\beta E_i)}{\sum_j^N exp(-\beta E_j)}$, where $\beta = \frac{1}{k_B T}$.

3.4.1 The Original WWW Algorithm

The starting point of the Original WWW method is crystal diamond structure of silicon. With periodic boundary conditions imposed, every atom in the supercell has exactly 4 neighbors and only six-fold rings exist. As we stated in chapter 1, the structure of amorphous silicon is believed to be a CRN. So the aim of the method is to generate a CRN starting with the diamond crystal. Since every atom in the crystal has exactly 4 neighbors, a crystal of N atoms will consist of exactly $2N$ bonds. The interaction between the atoms is described by the classical Keating potential given by equation 3.1. In this case, a Monte Carlo move is a *bond switch*.

Before explicitly describing the modeling process, let us explain the bond switching process using the cartoon figure in Fig. 3.1. Begin with a chain of 3 bonds A-B, B-C, and C-D (consisting of four atoms A, B, C, and D) as shown on the left side of Fig. 3.1. Next, delete the bonds A-B and C-D and respectively replace them with bonds A-C and B-D as shown on the right side of Fig. 3.1. In the computer code, all we do when a bond switching event is successful is to delete the bonds A-B and C-D from the $2N$ of bonds and replace it with the new bonds A-C and B-D. This is an artificial process since for example, two atoms close enough will not be bonded but two atoms far away can form a bond. All we follow here is what the bond list says. The bond list addition and subtraction always ensures that every atom has exactly four neighbors so that the Keating potential is satisfied. The bond switching events leads to the creation of five-fold and seven-fold rings which is a primary feature of CRN models of *a*-Si. Now that we know how to switch a bond, we explain how to generate a CRN model of amorphous silicon starting with crystalline silicon and using the WWW method. The modeling process consists of two main stages. In the first stage, the starting crystal is heavily randomized by a large number of bond switches at a temperature T which is of the order of the melting point of the material. A bond switch event is accepted with a Metropo-

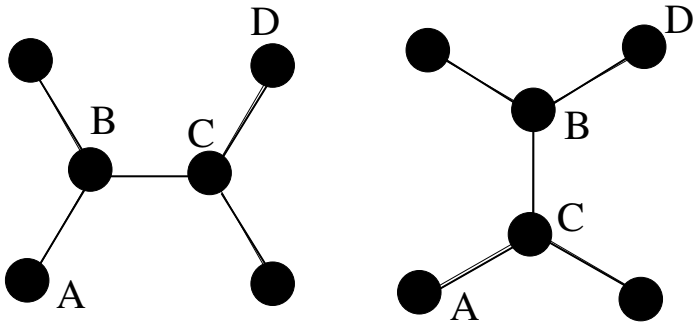

Figure 3.1: WWW bond switching event. Figure on the left denotes the configuration before bonds are switched and the one on the right is the configuration after bonds are switched.

lis probability $\mathcal{P}$ given by equation 3.16 i.e. $\mathcal{P} = \min(1, exp(-\beta\Delta E))$, where $\Delta E = E_f - E_i$, E_i is the total energy before the bond switch, E_f is the total energy after the bond switch, followed by a *total energy minimization* at zero temperature and, $\beta = (K_B T)^{-1}$. The initial stage is very crucial in the sense that a significant number of bond switches should be performed so that the randomized model has no traces of crystallinity.

In the second stage, we perform some more bond switches whiles annealing the system, and making sure that thermal equilibrium is achieved at each new temperature. Because of the nature of the Keating potential, unphysical short and long bonds will eventually become normal as the simulation proceeds. The final structure is a low strain, defect free CRN model of amorphous silicon. We will discuss the properties of these models in section 3.4.3.

3.4.2 The Modified WWW Algorithm

We will begin by briefly explaining the need to optimize the original WWW algorithm. First, the starting configuration is a crystal and so even after randomizing the system using bond switches, the system may retain traces of crystallinity. Secondly, after every bond switching event in the original method, we have to fully relax the structure i.e. minimize the total energy. Then, we

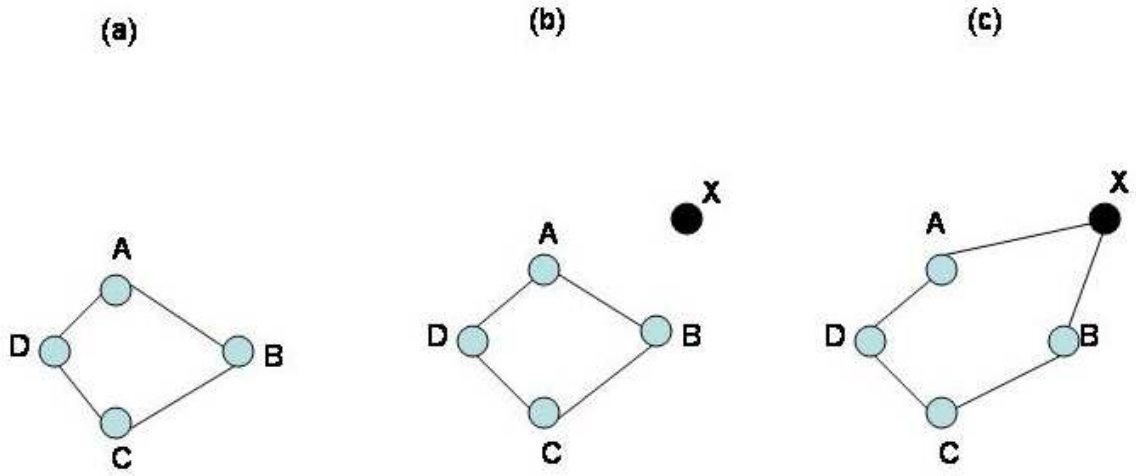

Figure 3.2: Picture depicting a single step in the artificial tetravalent network construction.

have to decide, using the Metropolis probability, whether or not the event will be accepted. For large systems, structural relaxations can be time consuming. If a move is eventually not accepted, time would have wasted on relaxing the structure. Thirdly, there is no scheme that tracks an attempted bond switch event so we do not have to select that event again. This also makes the modeling process time consuming.

Recently, Barkema and Mousseau (BM) [132] have optimized the algorithm by eliminating the problems outlined above. We will now explain how the BM method improves over the original WWW method.

1. The starting configuration is completely random. This, unlike the original method guarantees no trace of crystallinity. To generate the starting structure, N atoms are randomly place in a cubic box at crystalline density using periodic boundary conditions. A constraint that no two atoms a separated by less 2.3 Å is imposed. Next, we have to artificially make every atom four-fold by constructing a list of $2N$ bonds. Artificially, making the network perfectly 4-fold

is tricky and can be confusing. Let us briefly examine how this works. Ideally, two atoms separated by a cut-off distance r_c or less should be bonded. Begin by randomly constructing a loop consisting of four atoms $\{A, B, C, D\}$ and four bonds $\{AB, BC, CD, AD\}$ such that each bond length is less than r_c as shown in Fig. 3.2(a). At this point, only these four bonds exist in the network. No other bonds exist. From this point, some old bonds are broken from which new bonds are generated. This is achieved by gradually expanding the loop until each atom is visited twice. The trick here is to know how the loop is expanded. Let us explain this using a single step in the loop expansion process. Randomly select three atoms X, A, and B satisfying the following conditions:

(a) atom X has at most two neighbors and it is not bonded to neither A or B (according to the bond list) but it is within the cut-off distance r_c from both A and B

(b) atoms A and B are bonded i.e., the bond AB exist.

In Fig. 3.2(b) we have depicted how the atoms are selected. We expand the loop by deleting the bond AB from the list and replacing it with the bonds AX and BX. This is shown in Fig. 3.2(c). Each time the loop is expanded the artificial bond list is updated. After every atom is visited twice, the network becomes perfectly fourfold. This final structure is an extremely strained random network with unphysical bonds and a bond angle standard deviation of about 30°.

2. As we stated earlier, in the original approach, the network is fully relaxed after a bond switching event and then the Metropolis probability to either accept or fail to accept the new configuration. Barkema and Mousseau [132] have devised a smart way to reject a bond switching event without fully minimizing the total energy. Right after bonds are switched, a threshold

energy is determined before performing structural relaxation. Vaguely think of this as a reverse process. To determine the threshold energy, generate a random number $s \in (0,1)$ and equate it to the Metropolis probability $exp(-\beta\Delta E)$, where $\Delta E = E_t - E_i$, E_i is the total energy before the bonds are switched and E_t is the threshold energy. From the equation, we obtain

$$E_t = E_i - \frac{1}{\beta}\ln(s). \tag{3.17}$$

Once E_t is determined, we begin the relaxation process. The relaxation is mostly local as opposed to global in the original case. Local relaxations are permissible since after bonds are switched, only a small cluster of atoms around the bond switch site experience significant changes in forces. Local relaxations are performed up to within the fourth shell of the atoms involved in the bond switch event. The local moves significantly reduces the computation of the forces and energy from $\mathcal{O}(N)$ to $\mathcal{O}(1)$ [133]. At each step of the local relaxation, the total energy E is given by $E = E' + E_c$, where E' is the energy of the atoms outside the cluster which is fixed and E_c is the energy of the atoms forming the cluster [133].

For a well relaxed configuration, the potential energy is harmonic about the minimum. We can use this fact to determine an estimate of the total final energy E_f using the current E and the $\{\vec{f_i}\}$ on the atoms at each step of the relaxation process,

$$E_f \approx E - c_f \sum_i |\vec{f_i}|^2, \tag{3.18}$$

where $c_f < 1$. A bond transposition is accepted only if $E_f \leq E_t$. If at any stage in the relaxation process $E_f > E_t$, the bond switch is immediately discarded and a new event is attempted. By so doing, time is not wasted to fully relax the system before the decision to accept or fail to accept is reached. After a bond transposition is accepted, we globally relax

the system to relieve the system of the strain at the boundary separating the the cluster and the rest of the atoms.

3. Every attempted bond switch is marked so that an event is not duplicated. For a system of N atoms, there are ${}^4P_2 \times N = 18N$ possible bond transposition attempts. After all the attempts are made, the process is complete. The final model is relaxed at zero pressure.

With the optimization procedures mentioned above, the modified WWW algorithm produces high quality models of *a*-Si. The final models provide a good starting point for MD simulations using reliable classical potentials and accurate *ab initio* potentials.

3.4.3 Properties of WWW Models of a-Si

We now discuss the properties of the modified WWW models of *a*-Si. We will also compare them to models generated from the original algorithms to see the marked improvements that the modified algorithm provides.

We begin with the structural properties of the models. As we stated in chapter one, the structure of an amorphous matrix is characterized by its short range properties, in particular the distributions of the pair inter-atomic distances (radial distribution function) and the distribution of the bond angles. The short-range parameters that can be readily computed from the distributions are the mean bond length $\langle r \rangle$, bond length standard deviation Δr, mean bond angle $\langle \theta \rangle$, bond angle standard deviation. $\Delta\theta$. We also compute the dihedral angle distribution and ring statistics which we have already defined in chapter 1.

In Fig. 3.3, we show the plots for the bond angle distribution, dihedral angle, and radial distribution function, of a 1000-atom CRN model of amorphous silicon generated using the BM algorithm. In Fig. 3.3(a) , we see a nice match of the computed RDF with recent X-ray diffraction

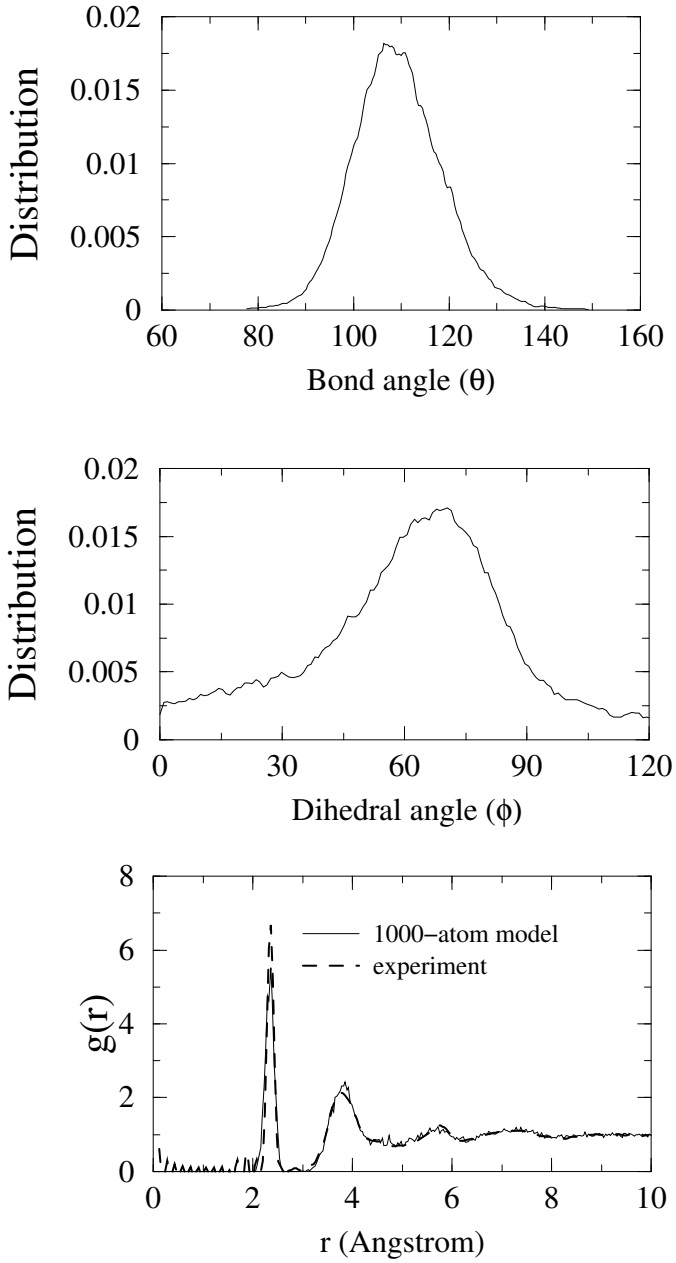

Figure 3.3: Structural properties of 1000-atom model generated using the modified WWW algorithm. Upper panel is the bond angle distribution, middle panel is the dihedral angle distribution, and lower panel is the radial distribution function (RDF). Experimental results for the RDF is taken from Ref. [11].

experiments by Laaziri *et al*. [11]. The experiments results were measured from annealed samples of *a*-Si prepared by ion bombardment. All the peaks corresponding to the coordination shells are correctly reproduced. In Fig. 3.3(b), we show a plot for the bond angle distribution. The bond angle standard deviation can be extracted from the second peak of radial distribution function, using the relation proposed Beeman *et al*. [12]. Using the experimental RDF results by Laaziri *et al*. [11], the relation yielded bond angle widths of 10.45° and 9.63° respectively for as-implanted and annealed samples. The configuration presented here has a bond width of 9.0° which compares favorably with annealed experimental samples. This configuration is expected to be the right structure since it was well relaxed at zero temperature. In addition to the radial and bond angle distributions, the model has no coordination defects. In Fig. 3.3(c), we show the dihedral angle distribution for the model. For the diamond crystal, the dihedral angle is a Dirac delta function centered on $\phi = 60°$. For the configuration presented here, we see a thermal broadening about $\phi = 60°$. This is expected for the disordered system and depicts the deviations of bond lengths and bond angles from the ideal crystalline system.

Even though structural analysis gives us some key information about a system, it does not tell us all about the local bonding environments. The other major quantities that are used to confirm how realistic a model is are the electronic and vibrational density of states. We have discussed the relevance of the electronic and vibrational density of states to amorphous networks in chapter 1. Strained regions of the network that are not immediately apparent from structural analysis can be picked up by the EDOS in the form of localized gap and tail states. We have therefore computed the EDOS of the 1000-atom CRN model from first principles. To do this, we used the *ab initio* code SIESTA [74, 75, 76] to relax the network. The exchange-correlation functional was approximated using LDA, where the correlation functional was obtained from parametrization of

Ceperly and Alder [47]. Valence electron states were expanded using single ζ basis set described in section 2.10. Norm conserving pseudopotentials, expressed in Kleinman-Bylander [72] form. Troullier-Martins [70] scheme was used to parametrize the pseudopotential. The network was relaxed using a conjugate gradient optimization scheme such that the maximum force on each atom is less than 0.04 eV/Å. Even though a 1000-atom model is somewhat large for a first principles calculation, the energy relaxation converges much faster because of the high quality nature of the CRN. Structural relaxation is also a way to check the stability of the model. The single-particle Kohn-Sham eigenvalues were obtained from the direct diagonalization of the Kohn-Sham Hamiltonian.

In Fig. 3.4, we show a plot of the EDOS of a fully relaxed network. The distribution has been smoothed with a Gaussian approximation for the delta function. As can clearly be seen, the EDOS is devoid of band gap states. A band gap state tells us about the nature of defects in the network. The clean gap possessed by this model indicates the system is well-relaxed. Defect-free models are important if one wants to study the role of defects in amorphous semiconductors. The graph in the inset shows that the decay of the valence states, known as the *Urbach* tails, is exponential i.e. it proportional to $\exp(-E/E_o)$, where the decay parameter $E_o \approx 0.25$ eV. This value for the decay parameter is in good agreement with a previous calculation on a large model by Dong and Drabold [134].

We explained in section 1.4 that computing the dynamical matrix scales as $\mathcal{O}(N^4)$ if the force are computed by directly diagonalizing the Hamiltonian. For a 1000-atom model this quite expensive so we have computed the VDOS for a 216-atom CRN model. The match of the Gaussian broadened computed VDOS with the inelastic neutron scattering experimental data [16] in Fig. 3.5 is not too bad. Apart from the small shift in the transverse optical peak to the right of the experimental

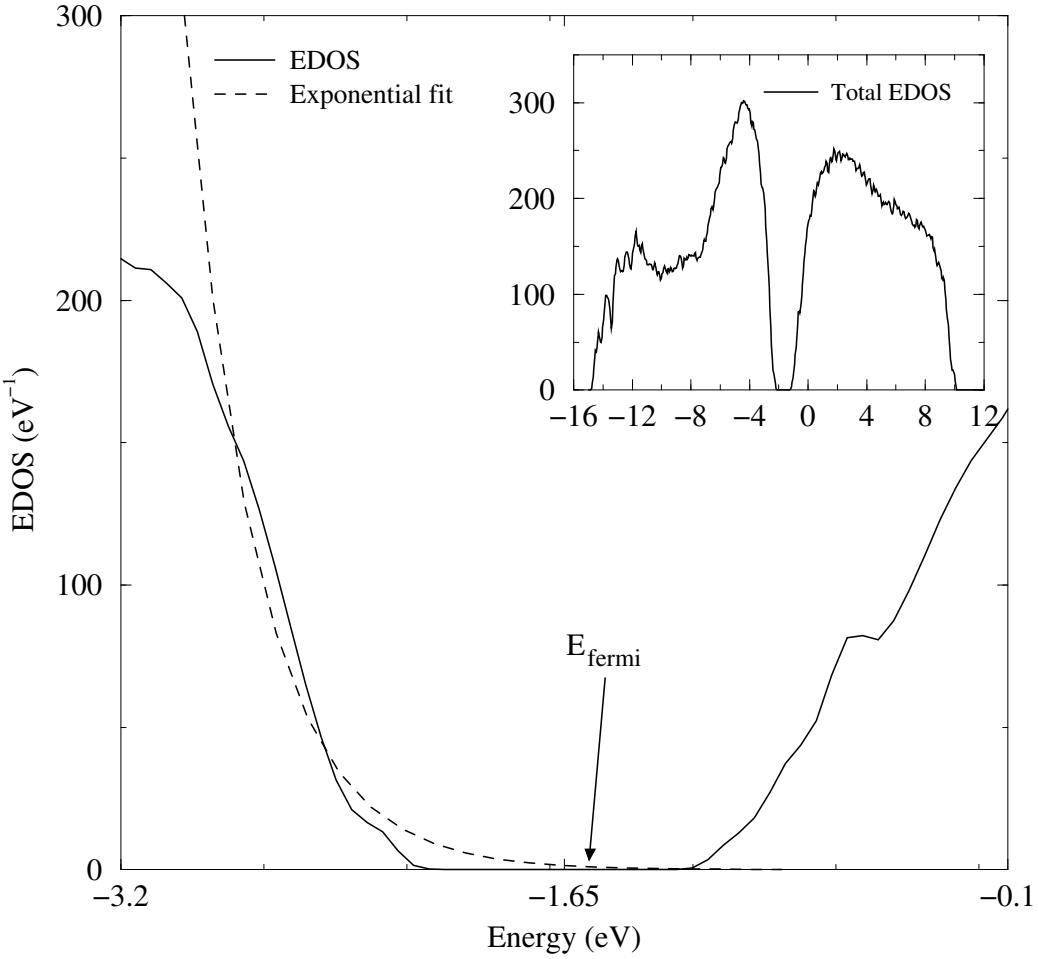

Figure 3.4: Electronic density of states (EDOS) for 1000-atom model generated using the modified WWW algorithm calculated from first principles. The main graph details the density of states near the gap and the graph in the inset shows the total density of states. The dashed curve is an exponential of the form $\rho(E) \propto \exp(-E/E_o)$, where $\rho(E)$ is the EDOS and the decay parameter $E_o = 0.25$ eV.

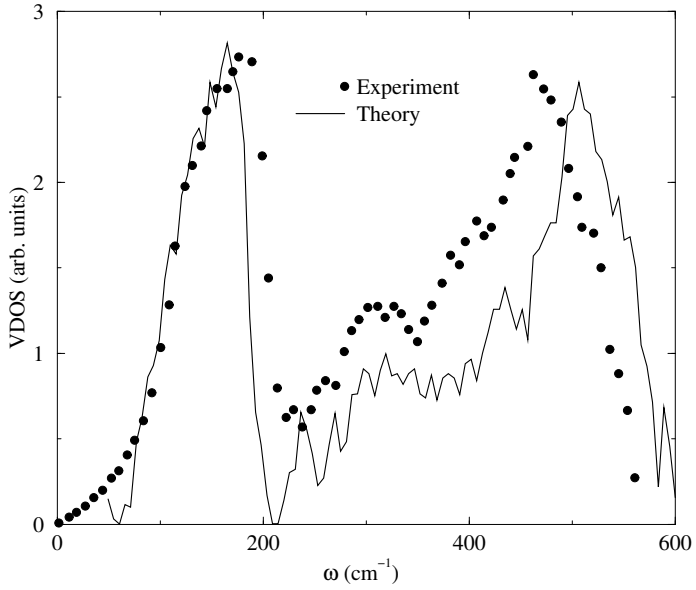

Figure 3.5: Vibrational density of states (VDOS) for 216-atom model generated using the modified WWW algorithm calculated from first from. Experimental data was taken from Ref. [16]. The shift in the computed transverse optical (TO) peak to the right of the experimental TO peak has been discussed in the text.

peak, the general features of the distribution are well reproduced. The shift in the peaks can be attributed to the procedure for preparing the sample in experiments and the difficulty in simulating the high frequency region of the vibrational spectrum due to the flattening of the phonon dispersion curve. The gap around $\omega = 200\mathrm{cm}^{-1}$ is due to the finite size effects in the sense that small values of ω correspond to wavelengths that are larger than the length of the simulation cell [123].

Finally, we compare and contrast the energetic and structural properties of modified WWW models to models obtained from the original WWW algorithm. The models generated from the original method is are two 4096-atom configurations due to Djordjević, Thorpe, and Wooten (DTW) [135]. We will refer these models as DTW4096(a) and DTW4096(b). For the models generated from the modified algorithm, we use 1000-atom (BM1000), 4096-atom (BM4096), and 10,000-atom (BM10K) models.

In table 3.3, we see that the total Keating energy per atom for the models obtained from the improved algorithm is smaller than all the DTW models, indicating that besides its efficiency, the modified algorithm yields models that have a significantly lower strain. Another important structural property is the bond angle standard deviation $\Delta\theta$. It is clear from table 3.3 that $\Delta\theta$ for the models generated from the improved algorithm agree closely with experimental value of 9.63° while the original models still shown a slightly larger deviation. We also show the ratio of the mean bond length $\langle r \rangle$ to the zero temperature crystalline bond length $d = 2.35$ Å. The results clearly show the deviation of the bond lengths from the d is not too large.

Table 3.3. Structural and energetic properties of models relaxed with the Keating springs. DTW4096(a) and DTW4096(b) are 4096-atom models are obtained from the original WWW algorithm and described in detail in Ref. [135]. All other models are due to the improved WWW algorithm due to Barkema and Mousseau (BM) [132]. BM1000: 1000-atom model; BM4096: 4096-atom model; BM10K: 10,000-atom model.

	DTW4096(a)	DTW4096(b)	BM1000	BM4096	BM10K
E (eV/atom)	0.336	0.367	0.250	0.304	0.301
$\langle r \rangle / d$	0.996	0.997	0.989	0.98	0.98
$\langle \theta \rangle$	109.2	109.3	109.3	109.3	109.3
$\Delta\theta$	10.51	11.02	9.00	9.89	9.88
rings/atom					
4	0.015	0.000	0.000	0.000	0.000
5	0.491	0.523	0.472	0.490	0.48
6	0.698	0.676	0.761	0.739	0.742
7	0.484	0.462	0.507	0.467	0.512
8	0.156	0.164	0.125	0.148	0.142

Chapter 4

Reverse Monte Carlo Modeling of Amorphous Silicon

In chapter 3, we reviewed the existing modeling schemes for generating CRN models of *a*-Si and the WWW [131, 132, 135] modeling method was by far, the most efficient and reliable scheme for generating high-quality CRN configurations of *a*-Si. All the modeling schemes discussed in chapter 3 begin with some initial configuration and a potential describing inter-atomic interactions and the system is then modeled using molecular dynamics or Monte Carlo. The key properties of the final model are compared to experiments to determine whether or not they are realistic. The resulting models can also be used to make predictions. We will refer to all these modeling schemes as *forward modeling*. The name forward modeling is so chosen because one begins with some initial configuration and uses some modeling scheme to generate a final configuration, whose properties may be consistent or inconsistent with experiments.

In this chapter we develop a *reverse modeling* approach to model amorphous semiconductors. In general, the starting point for reverse modeling is experimental data sets and some

initial configuration. The key idea of reverse modeling is to generate output configurations that satisfy the experimental data used in the modeling process. Generally, reverse modeling methods have the advantage that no inter-atomic potentials are needed and we are guaranteed that the model is consistent with experiments whereas forward modeling methods can be expensive and they do not guarantee models satisfying experiments. Reverse modeling methods also provide a platform for the interpretation of experimental information via atomistic modeling [136]. However, reverse modeling methods suffer from the problem of non-uniqueness in the sense that different models will satisfy the same set of experimental data used in the modeling scheme but may entirely fail to match experimental data not used in the scheme. *So if one uses a reverse model model scheme, one has to make sure that the basic structural and chemical properties are satisfied*.

Here, we focus on a Metropolis Monte Carlo method with reverse modeling known as *reverse Monte Carlo* (RMC) [136, 137, 138, 139, 141, 142, 143, 140], for modeling *a*-Si. Our primary objective is to produce structural configurations that are consistent with experimental data but at the same time we go one step further to generate realistic configurations for comparison with models obtained via other routes. We emphasize that producing realistic models (meaning models which agree with all experiments) requires more than spatial pair correlations, and identify additional constraints which lead to realistic models.

The existing RMC models of amorphous semiconductors are found to be inadequate and fail to produce some of the basic experimental features of amorphous tetrahedral semiconductors. Gereben & Pusztai [137, 138] have carried out RMC simulation of tetrahedral semiconductors using a number of models ranging from completely disordered configuration to randomized diamond structure. Although a certain degree of tetrahedral character in the bond angle distribution was reflected in their work, most of their models show an unphysical peak in bond angle distribution

around 60°. The work of Walters and Newport [139] on amorphous germanium made some progress toward getting the correct bond angle distribution, but the number of 3-fold coordinated atoms were quite high in their model and in absence of any discussion on local strain and electronic properties it is difficult to say how reliable their models are when it comes looking at the electronic properties.

The inadequacy of earlier RMC schemes to generate realistic CRN models of *a*-Si and the need for a reliable algorithm that will result in realistic CRN models was the motivation for carrying out this project. In section 4.1, we briefly mention the basic philosophy of reverse Monte Carlo modeling and some of its salient features. This is followed by role of constraints in RMC modeling in section 4.2 where we illustrate how a set of judiciously chosen constraints can be used to construct a reliable model of amorphous silicon. Finally we compare our results with those obtained from earlier RMC models and a model obtained via the WWW algorithm.

4.1 Basics of RMC

The RMC method has been described in detail elsewhere [140]. Here we briefly outline the basic philosophy of RMC. At the very basic level, RMC is a technique for generating structural configurations based on experimental data. The logic is very appealing: any model of a complex material worthy of further study should, at a minimum, agree with what is known (that is, the experiments). By construction, the RMC scheme enforces this (and for contrast, a molecular dynamics simulation may not). In an ideal implementation, one should find a model agreeing with *all* known information, but this is not easy to accomplish. As far as we are concerned, the approach was originally developed by McGreevy & Pusztai [140, 137] for glassy materials and liquids us-

ing different experimental data. In recent years however, progress has been made toward modeling polycrystalline systems as well [144].

We now briefly explain how the original RMC method works.

1. Begin by randomly placing atoms in a periodic supercell at the right number density, and in addition, impose the condition that no two atoms are closer within a cut-off distance. The cut-off is of the order of the equilibrium bond length of the material.

2. Compute the non-negative initial cost function ξ_i using an experimental data of choice. The most commonly used experimental data in RMC algorithms is the static structure factor $S_E(Q)$ obtained from diffraction experiments. For N experimental data points, we have

$$\xi_i = \sum_{j=1}^{N} \sigma_j [S_E(Q_j) - S_i(Q_j)]^2, \tag{4.1}$$

where $S_i(Q)$ is the simulated static structure factor corresponding to the initial configuration $\mathcal{C}$ and σ_j is the uncertainty associated with each data point.

3. Make a move by randomly displacing a single atom or group of atoms and compute a new cost function:

$$\xi_f = \sum_{j=1}^{N} \sigma_j [S_E(Q_j) - S_f(Q_j)]^2, \tag{4.2}$$

where $S_f(Q)$ is the simulated static structure factor corresponding to the the new configuration $\mathcal{C}'$.

4. Compute the cost difference $\Delta\xi = \xi_f - \xi_i$. The Metropolis Monte Carlo begins at this point. The probability of accepting the new configuration is given by

$$\mathcal{P} = \begin{cases} 1 & \text{if } \Delta\xi < 0 \\ exp(-\beta\Delta\xi) & \text{Otherwise} \end{cases} \tag{4.3}$$

If the new configuration is accepted, replace $\mathcal{C}$ with $\mathcal{C}'$ and ξ_i with ξ_f and go to step 3. Otherwise, keep $\mathcal{C}$ and ξ_i, discard $\mathcal{C}'$ and ξ_f, go to step 3 and make a new move.

Keep repeating steps 3 and 4 until convergence is achieved. We must also mention that the cost function is not restricted to only a single experimental data. Multiple experimental data sets can be used.

Besides glasses and liquids, the RMC method has been used to model magnetic materials and polymers. However, the method has never been accepted without some degree of controversy and the most popular criticism is the lack of unique solution from RMC. RMC can produce multiple configurations having the same pair correlation function. This lack of uniqueness, however, is not surprising, since usually only the pair correlation function or structure factor is used in modeling the structure, while there exists an infinite hierarchy of higher order correlation functions carrying independent structural information are neglected. In other words, RMC samples from the space of all models consistent with some limited body of data – in its simplest form (analyzing a single experiment) RMC is an ideal gauge of how *non-specific* the data is with respect to identification of an atomistic model. If the modeler possesses *a priori* information independent of that implicit in the experiment being fit to, it is necessary to add this information to the modeling in some fashion to receive a model in joint agreement with the experiment and the additional information.

4.2 A New RMC model

We begin by including the minimal information that is necessary to model a configuration of a-Si. In so doing, we use the radial distribution function obtained from a high quality model of amorphous silicon. This latter model was generated by Barkema and Mousseau [132] using a

modified form of WWW algorithm [131] having bond angle distribution close to 10° with 100% 4-fold coordination which we have described in chapter 3. In addition to this RDF, we also impose the conditions that the average bond angle of all the triplets Si-Si-Si should be near 109.5° and the corresponding root mean square deviation should be no less than 10°. The number of 4-fold coordinated atoms is driven to a specified value during the simulation by including a constraint on the average coordination number. It is to be noted that while there is no limit to the number of constraints that can be included in the system, there is no guarantee that mere inclusion of more constraints will necessarily give better results. Forcing a completely random configuration with too many competing constraints may cause the configuration to be trapped in the local minimum of the function ξ and may prevent the system from exploring a large part of the search space. By adding only the essential constraints that describe the chemical and geometrical nature of the bonding correctly, Eq. 4.1 can be written as :

$$\begin{aligned} \xi &= \sum_{i=1}^{M} \lambda_1 \{S_E(x_i) - S(x_i)\}^2 + \lambda_2 (\theta_0 - \theta)^2 \\ &+ \lambda_3 (\delta\theta_0 - \delta\theta)^2 + \lambda_4 \{1 - \Theta(x - x_c)\} \\ &+ \lambda_5 \{\phi_0 - \phi\}^2 \end{aligned} \tag{4.4}$$

where,

$$\begin{aligned} \theta &= \frac{1}{N_\theta} \sum_{i\{j,k\}} \theta_{ijk} \\ \delta\theta &= \sqrt{\langle (\theta - \theta_0)^2 \rangle} \end{aligned}$$

In Eq. 4.4, θ and $\delta\theta$ are the average angle and the rms deviation while ϕ and ϕ_0 are the current and proposed concentration of the 4-fold coordinated atoms. It is important to note that

each of the terms in Eq. 4.4 is non-negative, and should decrease ideally to zero during the course of minimization. Since the cost of energy associated with the bond length relaxation is more than the bond angle relaxation, atomic arrangements with large bond angle distribution but having correct RDF frequently result. The coefficients, λ_1 to λ_3, for the different terms in Eq. 4.4 can be chosen appropriately to minimize this effect. In general the coefficients λ are constant during the course of simulation but the minimization procedure can be slightly accelerated by making them vary in such a way that the contribution from each of the term are of the same order during the course of simulation. The coefficient λ_4 is usually assigned a large value in order to include a hard sphere cut-off as mentioned earlier so that no two particles can come closer to x_c while the coefficient λ_5 maintains the number of 4-fold coordinated atoms to a specified value. In RMC simulation of amorphous tetrahedral semiconductors one usually encounters the problem of having a pronounced peak at 60°. This peak is a characteristic feature of unconstrained RMC simulation and is due to the formation of equilateral triangles by three atoms. In the work of Gereben and Pusztai [137, 138], attempts were made to overcome this difficulty by constraining the bond angle distribution as well as by making an initial configuration which is 100% 4-fold obtained from a diamond lattice. The resulting structure is, however, found to be unstable and on relaxation using a suitable potential, the configuration tends to get back toward the starting structure, i.e., randomized diamond in this case [1].

In the approach of Walter and Newport [139], the initial random configuration was examined and any "triples", i.e. , three atoms forming an equilateral triangle was removed before the beginning of RMC fit. By selective removal of such unwanted triplets, they have been able to generate configuration of a-Ge without having a peak at 60°. The approach that we have taken in our work is more general and starts with a completely random configuration. This eliminates, in

[1] This observation is equally true in the context of original WWW model as pointed out by Winer and Wooten [145] and was remedied in their work by making the number of bond switching per atom above a certain minimal value.

the first place, any possible local ordering that may exist in the starting structure (e.g., randomized diamond structure retains the memory of tetrahedral ordering). Furthermore, we have not included or excluded any special configuration in our starting structure, e.g., three atoms forming an equilateral triangle. Based on experimental consideration, we have included only the key features of amorphous tetrahedral semiconductors – an average bond angle of 109.5° having rms deviation of 10° which is consistent with the RDF obtained from a WWW relaxed model used in our calculation. For the 500-atom model reported in this work, we have chosen a cubic box of length 21.18Å which corresponds to number density 0.0526 atom/Å^3. The initial configuration is generated randomly so that no two atoms can come closer to 2.0Å. The configuration is then relaxed by moving the atoms to minimize the cost function ξ. In addition to applying standard Monte Carlo moves in which a single or a group of atoms is randomly displaced, a variety of Monte Carlo moves have been implemented in our work. For example, in one of such moves, a 3-fold or 5-fold atom is selected and the nearest neighbor distance is examined. If the distance is greater than 2.7Å, the neighboring atom is displaced in order to bring the distance within a radius of 2.7Å. The maximum displacement of a Monte Carlo move is limited to 0.2-0.4Å throughout the simulation. Since we are interested in the electronic structure as well, we confine ourselves within a reasonable system size for studying the generated structure using a first principles density functional Hamiltonian. The density functional calculations were performed within the local density approximation (LDA) using the local basis first principles code SIESTA [74, 75, 76] described in chapter 2. We have used a non self-consistent version of density functional theory based on the linearization of the Kohn-Sham equation, the Harris functional approximation [53] described in chapter 2 along with the parameterization of Perdew and Zunger [49] for the exchange-correlation functional. We have used the Troullier-Martins [70]

norm conserving pseudopotentials factorized in Kleinman-Bylander form [72] were used to remove core electrons. Valence electrons were expanded using the minimal single ζ basis set.

4.3 Results

The results for the model including all the constraints are presented in Figs. 4.1-4.4. Since the structure factor is generally considered to be more sensitive to an arbitrary small change in the atomic positions than the radial distribution function, we have plotted the structure factors for the constrained RMC and WWW model in Fig. 4.1. It is evident from Fig. 4.1 that the agreement between the RMC and WWW model is very good both for small and large values of Q. In order to further justify the credibility of our model, we have plotted in Fig. 4.2 the structure factor from the experiment of Laaziri *et al*. [11] along with the same obtained from our RMC model. Once again we find that the agreement between the structure factor from RMC and the experimental results is quite good except for the few points near the first peak. It is very tempting to think this deviation as a finite size effect coming from the finiteness of our model. We have therefore calculated the structure factor for WWW models containing 300 to 4096 atoms of Si but the deviation continues to remain. Holender and Morgan [147] also observed similar deviation near the first peak in their work with a much larger model containing 13824 atoms which was compared with the experimental data obtained by Fortner and Lannin [146].

In Fig. 4.3, we have plotted the bond angle distributions (BADF) for both the RMC and WWW model. As we have discussed in section 4.1, the radial distribution function or structure factor can not alone provide all the necessary information that are needed to characterize an atomic configuration obtained from a reverse Monte Carlo simulation. A further characterization beyond

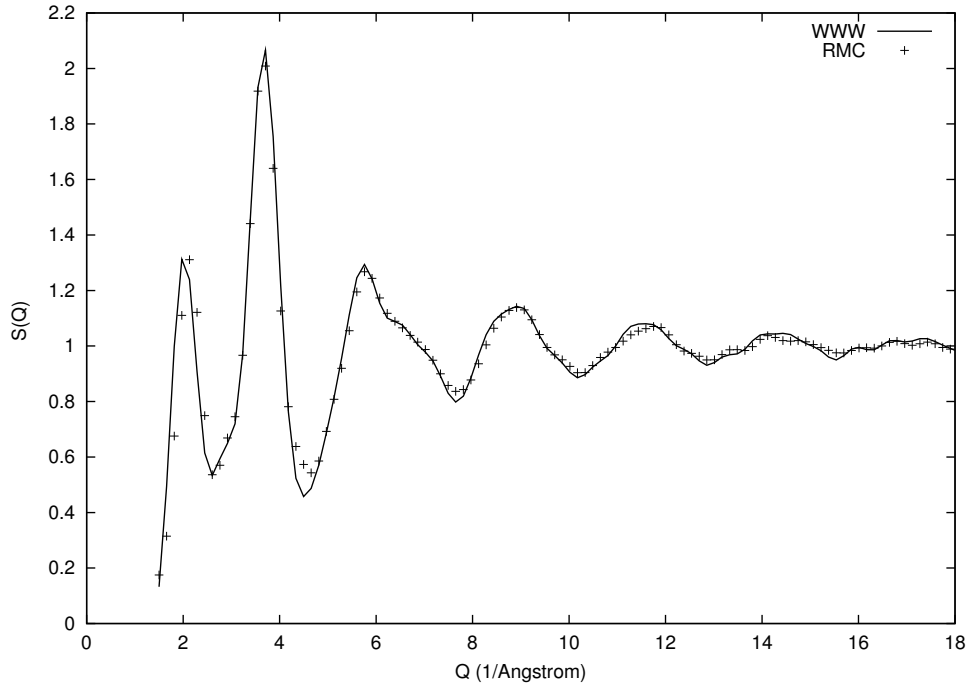

Figure 4.1: Structure factor obtained from a RMC (+) model containing 500 atoms of a-Si. The solid line is obtained from a WWW sample of identical size and number density of atoms.

pair correlation function is therefore vital and necessitates the need for getting some idea about the 3-body correlation function. It is clear from the Fig. 4.3 that the distribution obtained from the RMC model follows the tetrahedral character observed in amorphous semiconductors. The average bond angle in this case is found to be 109.01° with rms deviation of 12.5°. An important aspect of the bond angle distribution in Fig. 4.3 is that most of the angles are lying between 70°-150° compared to 80°-140° in WWW case. We emphasize at this point that the earlier works on modeling amorphous tetrahedral semiconductors using RMC predicted a much wider bond angle distributions. Gereben and Pusztai [137] have observed a pronounced, unphysical peak at 60° except for the model starting with diamond structure while Walters and Newport [139] have reported a bond angle distribution of a-Ge which is as wide as 60°-180°. It is an important development here that by adding three more constraints (λ_2, λ_3, and λ_4) we have achieved significantly improved results. Both the radial

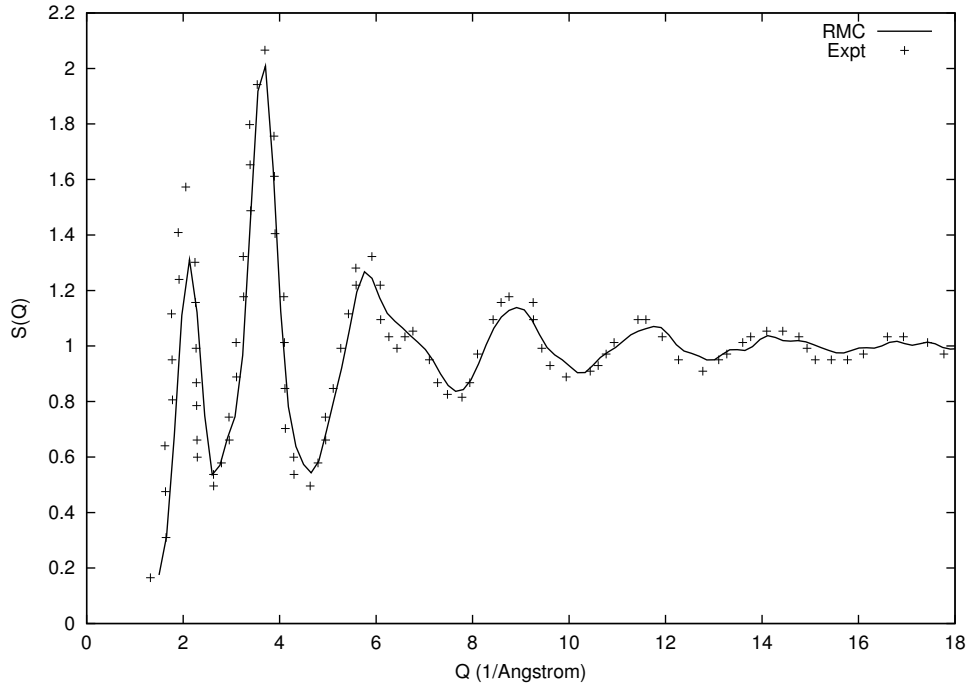

Figure 4.2: Structure factor obtained for a 500-atom model of a-Si from RMC (solid line) and the experiment of Laaziri *et al*. [11] as indicated in the figure.

and the bond angle distribution functions reported here are at par with the results obtained from molecular dynamics simulation and is comparable to those obtained from WWW model. The fact that the inclusion of these two constraints leads to a significant improvement is not surprising. For a large continuous random network (CRN) model of amorphous tetrahedral semiconductor, one can approximate the bond angle distribution as nearly Gaussian [2]. This approximated Gaussian distribution can defined by the first two moments of the distribution function. By specifying these two moments as constraints in Eq. 4.4, we correctly describe the tetrahedral bonding geometry of the atoms which along with the radial distribution function produces a configuration more realistic than those obtained from models based on RDF or structure factor only. This suggests that in addition to

[2]Although we are not aware of any mathematical proof, large continuous random network models of amorphous tetrahedral semiconductors obtained by WWW algorithm exhibit a bond angle distribution P(cosθ) very close to a Gaussian distribution. See for example, J. Dong and D.A. Drabold, Phys. Rev. B **54**, 10284 (1996).

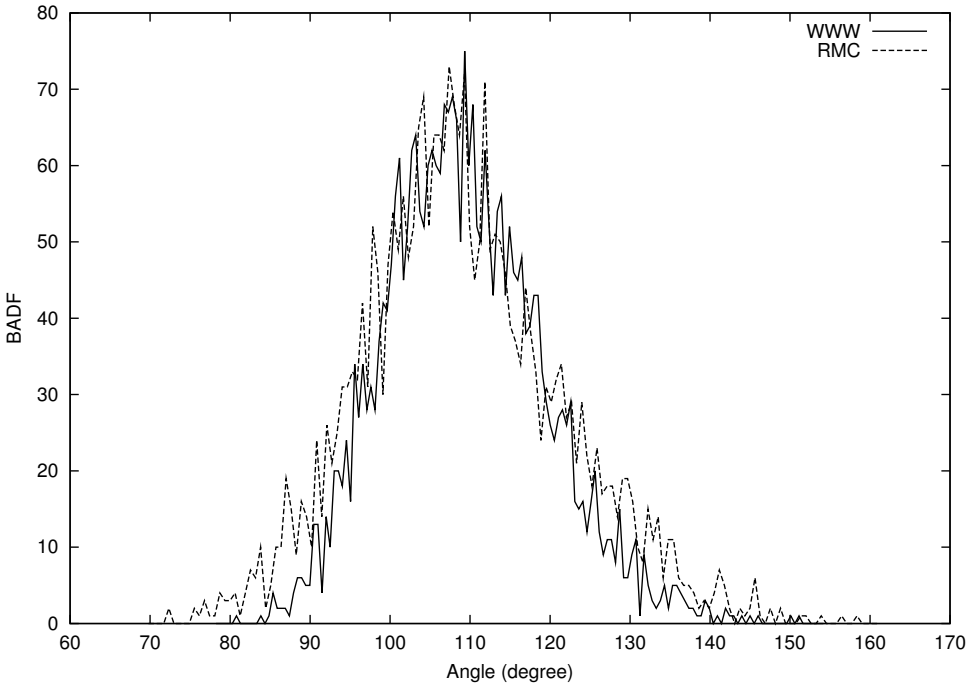

Figure 4.3: The bond angle distribution functions (BADF) for 500-atom a-Si from constrained RMC (dashed line) and WWW model (solid line). The rms deviation for the models are 12.5° and 9.9° respectively.

the radial distribution of the atoms, one needs to include some relevant information about the nature of 3-body correlation among the atoms to construct a realistic configuration.

Having studied the radial and bond angle distribution we now address the electronic density of states calculations. While the width of the bond angle distribution function (BADF) and the structure factor together indeed gives some idea about the quality of the model, some of the features e.g., the existence of spectral gaps and the position of defects states in the spectrum can be studied by looking at the electronic density of states only. The structure obtained from RMC simulation is first relaxed using the density functional code SIESTA and is found to be close to an energy minimum in the local density approximation (LDA). This is an important test for determining the stability of the structure obtained from RMC simulation and as far as we are concerned almost all earlier works

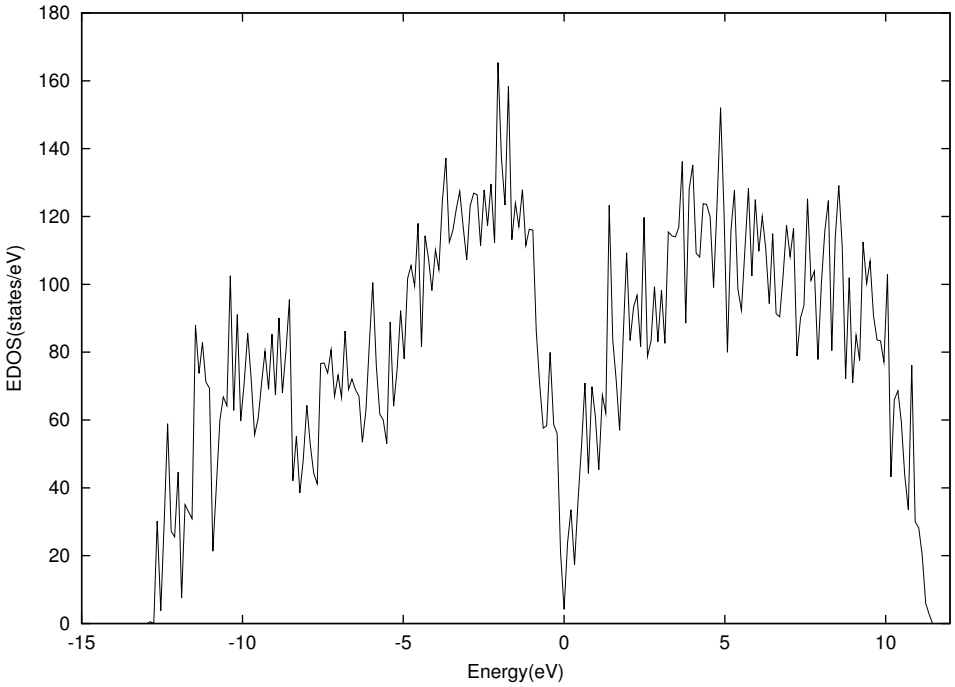

Figure 4.4: The electronic density of states (EDOS) of 500-atom model of a-Si obtained from RMC simulation described in the text. The Fermi level is at E=0.

on RMC have completely neglected this issue. In Fig. 4.4, we have plotted the electronic density of states (EDOS) for the constrained model. The EDOS appears with all the characteristic features of a-Si with the exception of a clean gap in the spectrum. This behavior is not unexpected in view of the fact that 88% of the total atoms are found to be 4-fold coordinated with an average coordination number 3.85. The presence of the defect states makes the gap noisy and at the same time the use of LDA underestimates the size of the gap. This EDOS is in significantly better agreement with optical measurements than conventional RMC models with much higher defect concentrations and spurious bond angles. It is interesting to observe that the average coordination number from our model is very close to the experimental value of 3.88 reported by Laaziri *et al*. [11].

4.4 Conclusion

We have presented a model of amorphous silicon based on reverse Monte Carlo simulation. One of the novel features of our model is to start with a completely random structure and then to relax toward a realistic configuration by adding a number of physically relevant constraints. The characteristic features of the tetrahedral bonding are taken into account by adding constraints on average bond angle and its deviation from the mean while the number of 4-fold coordinated atoms is maintained at a specified value by further use of a constraint on average coordination number. The radial and the bond angle distribution obtained from our model is found to be in excellent agreement with a high quality CRN model produced by WWW algorithm. We have also compared the structure factor with the experimental data obtained by Laaziri *et al*. [11] and observed a reasonably good agreement. By relaxing the model using the first principles density function code SIESTA, we find that the model is close to the energy minimum for LDA and is stable. The electronic density of states (EDOS) obtained from our model contains all the essential feature of amorphous silicon including a signature of the band gap. Although the model does not produce a clean gap in the spectrum, the quality of the EDOS is at par with models obtained from molecular dynamics simulation. Our RMC algorithm presents a significant improvement on previous RMC studies and makes it possible to compare for the first time, albeit qualitatively, the structural and electronic properties of RMC models with its WWW counterpart. We expect that further developments toward this direction will eventually make RMC as an useful modeling tool incorporating experimental information and can be used effectively without any criticisms in modeling complex materials.

Chapter 5

Experimentally Constrained Modeling of *a*-Si Using Fluctuation Electron Microscopy Experimental data

Medium range order (MRO) in amorphous silicon is generally considered to be any structure that lies between short range (2-5 Å) and long range ($>$ 300 Å) disorder [148]. Quantifying this order therefore requires information beyond the radial distribution function (RDF). Until recently, direct experimental data about atomic structure in disordered systems were limited to the structure factor. If one assumes that the material is isotropic, then the Fourier transform of the structure factor (with the angular information integrated out) yields the RDF [11]. The problem with the RDF is that it is insensitive beyond short range. Probes of the vibrational characteristics of the material, such as inelastic neutron scattering, Raman spectroscopy, and also, X-ray absorption spectroscopy and nuclear magnetic resonance have been employed with success, but yield structural information that

can be difficult to interpret [12, 149, 150, 151, 152]. In ionic and covalent glasses, hints about MRO come from the first sharp diffraction peak (FSDP) in the total factor structure [153]. This feature is believed to correspond to large spatial distances in the material, and understanding their origin is key to unraveling details of medium range order. Recently, the design of MRO in zinc-chloride networks using crystal engineering has been demonstrated and the degree of MRO measured from the FSDP [154]. The degree of MRO present in disordered systems is believed to affect the electronic, optical, and transport properties of these materials. The well known Staebler-Wronski effect, for example, which is the creation of metastable dangling bonds in hydrogenated amorphous silicon upon exposure to visible light [3], has been observed to occur in the material with diminishing medium range order [155].

Obtaining information about correlations other than the pair correlation in disordered systems remains a fundamental, yet presently unresolved, issue [153, 156, 157]. Recently, Treacy and Gibson have developed a low resolution electron microscopy technique known as fluctuation electron microscopy (FEM) [158]. In a nutshell, FEM is the statistical analysis of the fluctuations in the diffracted intensities due to scattering from mesoscopic volumes in amorphous thin films. It has been shown in detail elsewhere that the fluctuations in the diffracted intensities is, in a non-trivial way, related to three and four body correlation functions [159]. These correlation functions are much more sensitive to MRO than the RDF. Many experimental FEM studies have been done on *a*-Si and *a*-Ge. The conclusion reached in all cases was that MRO, as gauged by the fluctuations in the image intensity, is present in the as-deposited forms of these disordered systems [159, 160].

The CRN model for silicon, which we described in previous chapters has been successful in describing almost all known physical, electronic, and vibrational properties of the material. However, the CRN model for *a*-Si fails to produce FEM signal that matches experiment. An alter-

native model, the para-crystalline (PC) model of *a*-Si, which consists of nanometer sized crystalline grains (with a total volume fraction of $10 - 50\%$) embedded in an amorphous matrix was proposed [161, 162]. The PC model has all the known structural, electronic, and vibrational characteristics of *a*-Si and in addition, produces a computed FEM signal that matches experiment. Depending on the orientation of the crystalline grains to the incoming electron beam, regions of the material will be near or away from Bragg conditions. Crystals near Bragg conditions will diffract strongly, leading to large fluctuations in the diffracted intensity while those away from Bragg conditions will diffract weakly leading to small fluctuations in the diffracted intensity. Recently, it has been shown that the FEM signal in the PC model depends on the shape, size, and orientation of the nanometer-sized crystalline grains in the amorphous matrix [163].

If the CRN model for *a*-Si is an ideal atomistic model for the material, then it should be in reasonable agreement with experiments. As stated earlier, the shortcoming of the CRN model is its failure to satisfy FEM experiments. There have been previous claims that the PC model is the true representation of *a*-Si and that the CRN model is hypothetical [159, 163]. However, although the PC model satisfies all the known properties of *a*-Si and produces FEM signal comparable to experiment, we do not know whether or not it is unique. The aim of this work is investigate whether or not there exist a class of realistic models of *a*-Si, which are not PC models, and have FEM signature that matches experiment.

We propose two ways to build such models.

1. One can simply begin with a CRN model having no experimental FEM signal and build the experimental information into the model using the RMC algorithm described in chapter 4 and impose the necessary topological and energetic constraints. As we will describe later, this scheme is very easy to implement.

2. One can begin with an initial completely random network and use a reliable forward modeling method (modified WWW algorithm due Barkema and Mousseau [132] in this case) to model a CRN configuration and simultaneously enforce experimental FEM information using a reverse modeling method (RMC in this case). In chapter 3 we described the modified WWW algorithm and in chapter 4 we discussed the RMC method in detail.

We refer to the second modeling scheme as experimentally constrained modeling (ECM). One can think of ECM as part forward modeling and part reverse modeling. Hence, the underlying philosophy behind ECM is to build atomistic configurations from scratch using well-known forward modeling methods and at the same time use RMC to ensure that the system satisfies all known experimental, topological, and chemical properties. The beauty of ECM is that its applicability is flexible, in the sense that it can always be tailored to suit a disordered system of interest. Recently, Biswas *et al*. used ECM to successfully model glassy $GeSe_2$ using *ab initio* methods and RMC [164]. They obtained a model that matches well with experiment and is at the minimum of the inter-atomic potential describing the system. Details of the work is outlined in a recent dissertation by D. Tafen in Ref. [165].

5.1 Fluctuation Electron Microscopy

FEM theory and practice has been described in detail in a recent Ph.D. thesis by Paul Voyles [1] so we will briefly explain how experimental data is measured and explain the theoretical aspects needed for this work.

FEM is essentially a high energy electron diffraction technique, which gains its sensitivity to MRO by measuring fluctuations in the diffracted intensity from MRO-scale volumes of a sample.

The two key elements of FEM experiments are the radius of the objective aperture in reciprocal space Q and the scattering vector $\mathbf{k}$ accepted by the objective aperture. The corresponding real space resolution is $0.61/Q$. Here, we are interested in the fluctuation in the intensity for varying $\mathbf{k}$ at a fixed spatial resolution. The kinematic coherent image intensity $I(\mathbf{r}, \mathbf{k}, Q)$ due to scattering from a volume of size proportional to $1/Q$, centered on a pixel at position $\mathbf{r}$ is given by [158, 159, 166]

$$I(\mathbf{r}, \mathbf{k}, Q) = \lambda^2 (f(k))^2 \sum_i \sum_j a_Q(\mathbf{r} - \mathbf{r}_i) a_Q(\mathbf{r} - \mathbf{r}_j) F_{ij}(\mathbf{k}),$$

where the summations runs over all atoms within the sampling volume, $f(k)$ is the electron scattering factor, and λ is the electron wavelength. $a_Q(|\mathbf{r} - \mathbf{r_i}|)$ is the microscope point spread function of width $0.61/Q$ centered on atom i, which, if we neglect the effects of aberrations, is acceptable at low resolution, and is given by the Airy function [158]

$$a_Q(\mathbf{r} - \mathbf{r}_i) = Q \frac{J_1(2\pi Q|\mathbf{r} - \mathbf{r}_i|)}{|\mathbf{r} - \mathbf{r}_i|},$$

where J_n is the Bessel function of the first kind of order n. $F_{ij}(\mathbf{k})$ describes the incoming illumination, which within the hollow-cone approximation [158, 159] is given by

$$F_{ij}(\mathbf{k}) = J_0(2\pi k|\mathbf{r}_i - \mathbf{r}_j|)$$

The dimensionless normalized variance $V(k, Q)$ which is used to determine the degree of MRO in the material is defined as

$$V(k) = \frac{\langle I^2(\mathbf{r}, k, Q)\rangle - \langle I(\mathbf{r}, k, Q)\rangle^2}{\langle I(\mathbf{r}, k, Q)\rangle^2},$$

where $\langle\,\rangle$ denotes average over pixels. It has been shown elsewhere [159] that $I^2(\mathbf{r}, k, Q)$ is related to three and four body correlation functions, which in turn, gives information about MRO. A large $V(k)$ implies a high degree of MRO and a small $V(k)$ implies little or no MRO.

In Fig. 5.1, we show a simple picture of how diffraction measurements are taken in FEM experiments. Because the diffraction of amorphous silicon is isotropic, a hollow-cone illumination is used, which means that the incident electron beam is rotated at a constant inner angle so electrons that are scattered through a scattering vector with constant magnitude k, will travel down the optic axis of the microscope and form the image [1]. Scattered electrons will pass through the objective aperture of width Q and continue on to form the image. Unscattered electrons are blocked by the aperture, so images with no sample will appear dark [167]. A typical hollow-diffraction image is shown in the inset of figure 5.2. We also show a typical plot for the variance of the diffracted intensity measured in FEM experiments in fig 5.2. FEM experimentalists use the height of the peaks to measure the degree of MRO in the sample [1].

5.2 Experimentally Constrained Modeling

We begin this section by describing how a CRN model of a-Si that has no FEM signal to begin with, can be modified to have the signal using RMC and appropriate constraints. To achieve this, we use RMC to minimize the cost function

$$C = \lambda U_{MSW} + \lambda_1 C_1 + \lambda_2 C_2 + \lambda_3 C_3, \tag{5.1}$$

where

$$C_1 = \sum_{i=1}^{N_{exp}} (V(k_i) - V_{exp}(k_i))^2, \tag{5.2}$$

$$C_2 = \sum_{j=1}^{N_{atoms}} (n_j - 4)^2. \tag{5.3}$$

$$C_3 = 1 - \Theta(r - r_c), \tag{5.4}$$

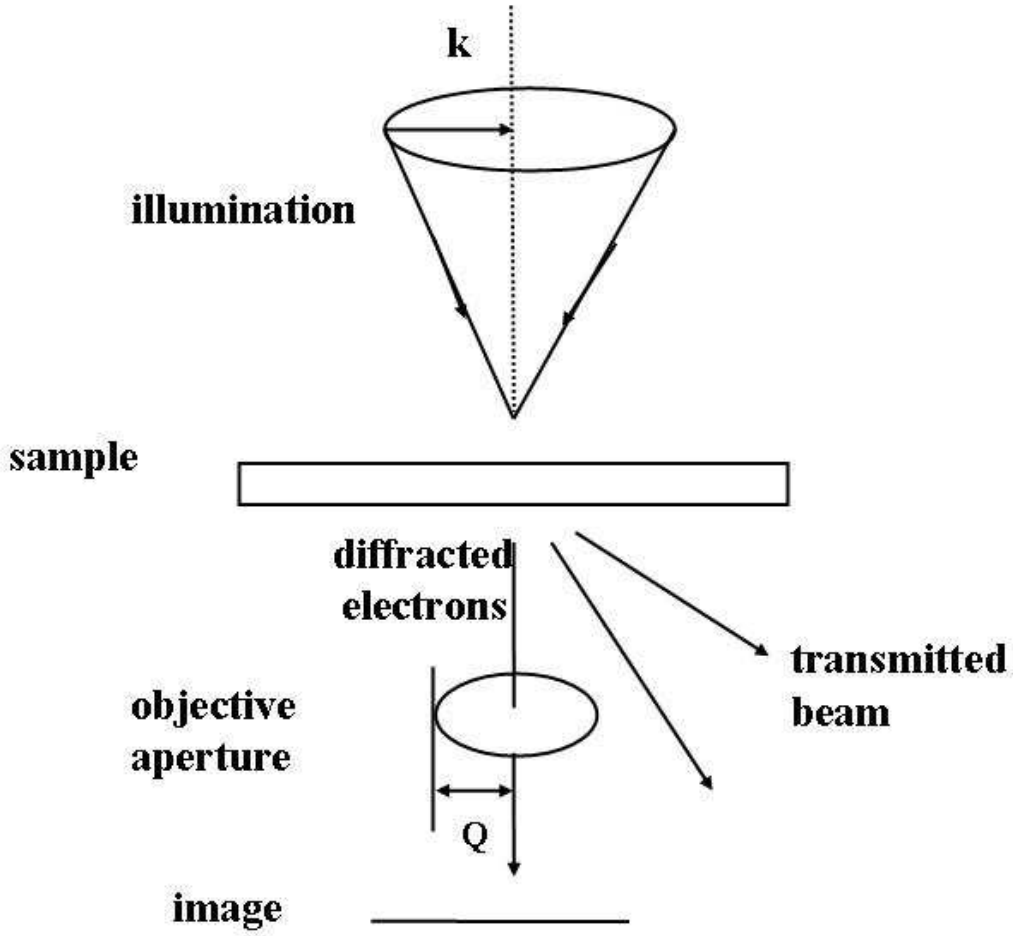

Figure 5.1: **k** and Q, the two parameters that control the images captured in FEM experiments. The picture was taken from Ref. [1].

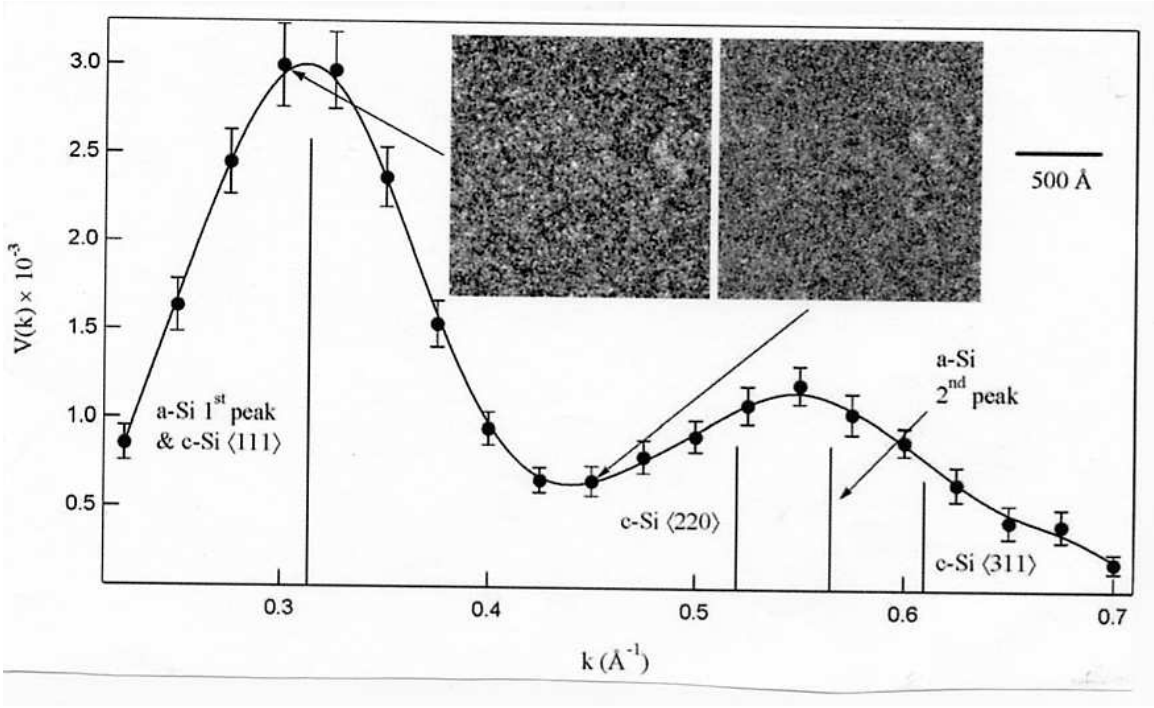

Figure 5.2: Typical $V(k)$ data for a-Si. Positions of diffraction peaks in c-Si and a-Si are indicated.Insets show dark field images. captured in FEM experiments. The picture was taken from Ref. [167].

U_{MSW} is the modified Stillinger-Weber (MSW) [99] potential, and λ, λ_i are coupling constants. Because the starting configuration is well relaxed with the Keating potential [97], we have used the MSW potential as an energetic constraint to regulate the energy landscape of the system. If the energetic constraint is not imposed, one ends up with a model having the experimental FEM signal but energetically unstable. Such a model contradicts experiments. C_1 is the part of the cost that controls the FEM signal, where N_{exp} is the number of experimental data points, $V(k_i)$ is the simulated FEM signal, and $V_{exp}(k_i)$ is the experimental FEM signal. C_2 maintains the 4-fold coordination of the network, where n_j is the coordination for atom j. C_3 ensures that no two atoms are separated by a distance less than r_c, where Θ is the Heaviside step function. Typically, the minimum value of r_c is 2.0 Å. One difficulty in the simulation is appropriately choosing the λ and λ_i so that there is no bias toward a particular component of the cost function and the system does not get trapped in some local minimum in phase-space. However, preference is given to C_1 since it controls the amount of FEM signal we put in the system. We have used 1000-atom and 4096-atom configurations due to Barkema and Mousseau [132] as starting models. The 1000-atom model is perfectly fourfold with a bond angle standard deviation of 9.2° The 4096-atom model has an 99.9% fourfold and bond angle standard deviation of 10.02°. Throughout this paper, we will refer these models as modified Barkema-Mousseau (MBM) models.

We now describe how to generate ECM configurations. We begin with a truly random configuration as the starting structure rather than a disordered crystal, thus guaranteeing that the configuration has no trace of crystalline regions. We used the modified version of the WWW bond switching algorithm and constrained RMC for the ECM process. To achieve this, we minimize cost

$$C = \lambda_1 C_1 + \lambda U_K \tag{5.5}$$

where U_K is the Keating potential and λ, λ_1 is are coupling parameters. C_1 is minimized using RMC and U_K is minimized using Newton's method. In this case, we modeled only a 1000-atom system. Throughout this chapter, we will refer this configuration as ECM configuration.

The experimental data used for this work is due to Voyles and Muller [168]. For all 1000-atom models, we use a real space resolution of $0.61/Q = 6$ Å and for the 4096-atom model we use $0.61/Q = 14$ Å.

5.3 Discussion of Results

We have used the RMC scheme described above to obtain the 1000-atom and 4096-atom MBM models. Each model has FEM signal that matches experiment. In Fig. 5.3, we show the $V(k)$ plots for the starting CRN models, the MBM models, and experiment. We can clearly see a nice match between the experiment and the MBM model. Because of the coordination number constraint, the topology of the starting CRN model and final MBM models are the same. However, the MBM models is strained because in the process of incorporating the FEM signal via RMC, the bond lengths and bond angles become distorted. Energetically, the CRN configuration is more stable than the MBM configuration. If the total energy of the MBM model is minimized with any potential, it relaxes to the starting CRN and the peaks in the FEM signal completely vanishes. The instability of the signal suggests that the model is unphysical. However, it asserts that there strained disordered networks without nanometer-sized crystallites that produce FEM signal matching experiment.

In table 5.1, we show some structural properties of 1000-atom ECM models generated by method described in section 5.2. ECM$_{Keat}$ is obtained by independently minimizing the total ECM model with the Keating potential. We clearly see that there is significant strain in the original

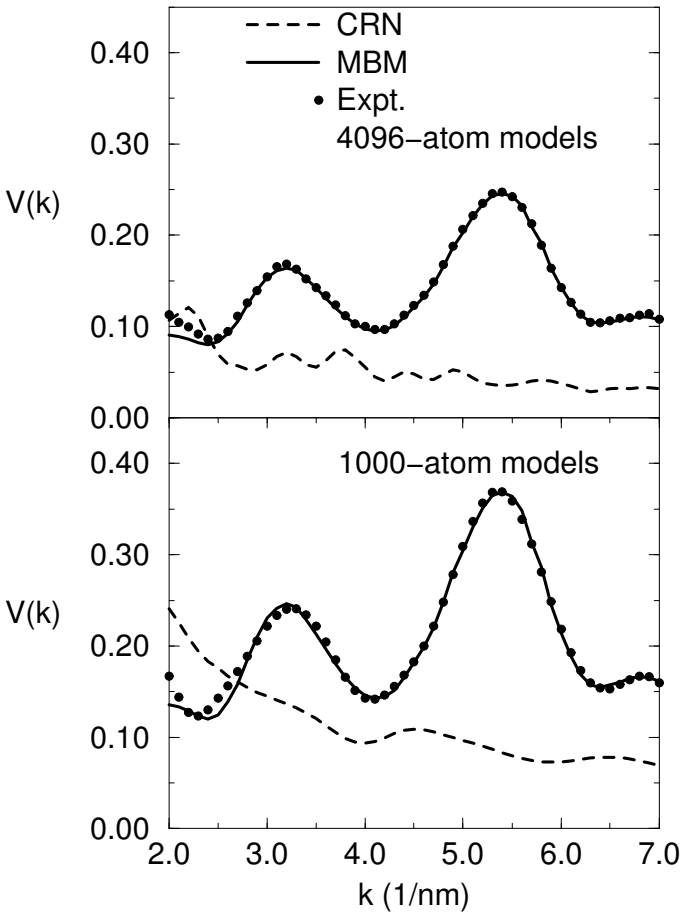

Figure 5.3: Plots of the variance of the diffracted intensities for 4096-atom models (top) and 1000-atom models (bottom).

Table 5.1. Total energy (Keating) per atom and short range properties for the models ECM and $\mathrm{ECM_{Keat}}$.

	ECM	$\mathrm{ECM_{Keat}}$
E_{Keat} (eV/atom)	0.321	0.251
$\langle r \rangle$	2.33	2.35
$\langle \theta \rangle$	109.0^o	109.2^o
$\Delta\theta$	10.9^o	8.9^o
4-fold coordination	98.6%	100%
% of N-member rings		
4	0.0	3.79
5	31.2	36.45
6	52.27	40.02
7	14.53	16.17
8	2.1	2.47

ECM network compared to the geometrically relaxed counterpart. The reason for this is simple: the number of moves that simultaneously reduce the energy and satisfy the constraint are significantly less than the moves that will lower only the energy. The strain is clearly reflected in the spread of the bond angle about the central value. It is not surprising that the energy difference between ECM and $\mathrm{ECM}_{\mathrm{Keat}}$ is 0.07 eV/atom. The overall quality of the ring statistics is good and compare favorably to models generated with other schemes. The well relaxed nature of $\mathrm{ECM}_{\mathrm{Keat}}$ yields more 6-member rings, which is the only irreducible ring present in crystalline diamond. We have checked the dihedral angle distribution for these models and observed that there is not much difference between them.

We now discuss the presence of FEM signal in the ECM models. In figure 5.4, we show plots of the fluctuation in the diffracted intensity $V(k)$ against the scattering vector magnitude k. In the top panel, we show plots for a 1000-atom ECM model, 1000-atom CRN, and experimental data. We see a nice match of $V(k)$ for the ECM model with experiment and then a nearly featureless plot for the CRN. The ECM model has a relatively high energy, so we minimized the potential energy surface independently with the Keating potential and an *ab initio* Hamiltonian SIESTA [74, 75, 76]. We use the energy minimization process to test the stability of the FEM signal. The $V(k)$ plots for the resulting configuration are shown in the lower panel of figure 5.4. It can be observed that the signal reduces by more than a factor of half for these relaxed models but does not vanish completely, indicating that the relaxed models still has some degree of MRO. It is worth noting the reversal of peak heights after relaxation. In the PC model, peak height reversals are related to grain shape, size, and their relative orientations. The reason why we observe this effect in our model is not exactly known. Experiments by Gibson and Treacy [160] have shown that the FEM signal of as-deposited *a*-Si decreases upon annealing. While the structural relaxation we have performed here and the

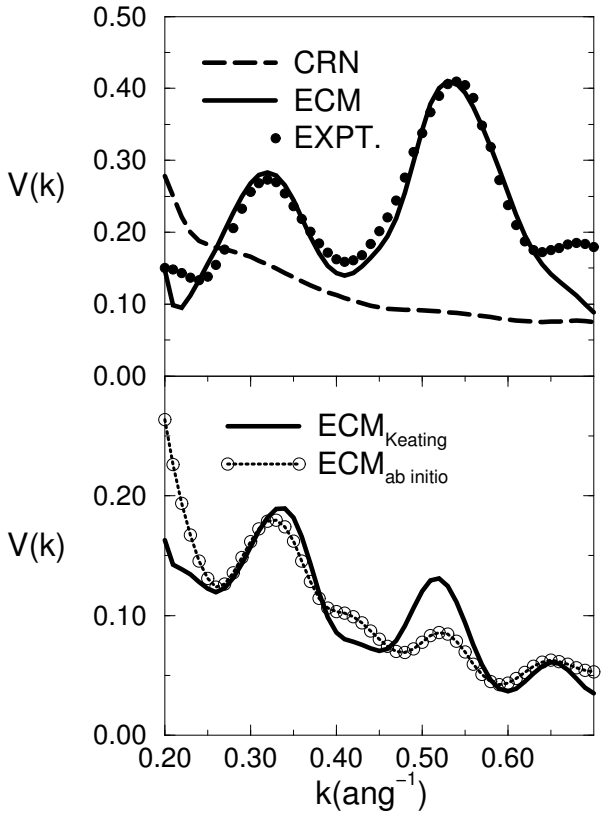

Figure 5.4: Plots of the normalized variance of the intensity against the scattering vector magnitude. The top panel shows the plots the ECM, CRN, and experiment. The top panel shows the reduction in the signal after relaxations with Keating and *ab initio* potential. Intensity is computed using spatial resolution $0.61/Q$ = 0.6 nm.

annealing process in experiment are quite distinct processes, a comparison between the two is not unreasonable.

For the rest of this section, we will dwell on the possible real space atomic arrangements that give rise to the FEM signals the ECM models. In particular, we are interested in structural units (with size proportional to the MRO length scale) in the ECM networks that control the FEM signal and which a conventional CRN lacks. Generally, there is a problem with a very clear description of how specific atomic arrangements in real space give rise to MRO. For real space correlations involving three atoms or more, we are limited. However, we employed an informative topological tool known as the Schlafli cluster analysis [169, 170] to search for MRO. The Schlafli cluster of an atom consists of all the atoms and bonds in all the rings emanating from that atom. Each Schlafli cluster is labeled by its weight and circuit symbol. The weight is just the number of atoms involved in a cluster. The circuit symbol is a convenient way of summarizing all the rings emanating from an atom [171]. For example, diamond-like crystalline silicon has the Schlafli cluster $6_2.6_2.6_2.6_2.6_2.6_2$: 29. indicating that there are two different six-membered rings emanating from each pair of bonds of the central atom and that there are 29 atoms in the cluster. The diameter of this cluster is 0.9 nm. This is a measure of MRO that does not first select a length scale, such as the fourth neighbor distance for the dihedral angle distribution, then go looking for MRO at that length scale. Schlafli clusters can easily be found in the PC model because of *a priori* implantation of crystalline grains. We apply this tool to compute the Schlafli cluster of atoms in all models. Cubic diamond (weight 29), hexagonal lonsdaleite (weight 27) and body-centered γ-Si (weight) all have a common circuit symbol of $6_2.6_2.6_2.6_2.6_2.6_2$. So our search was focused on clusters with this circuit symbol. A $6_2.6_2.6_2.6_2.6_2.6_2$:23 cluster was found in the strained ECM model. The diameter of the cluster is

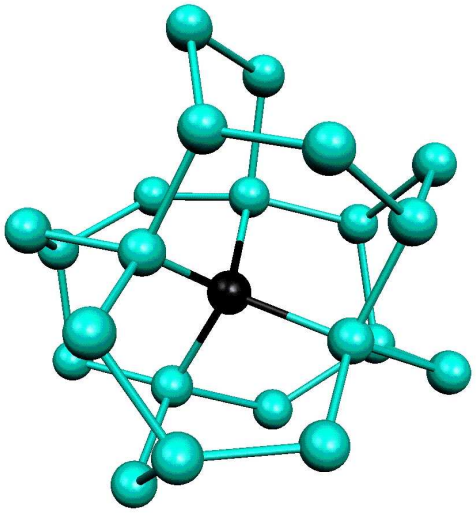

Figure 5.5: $6_2.6_2.6_2.6_2.6_2.6_2$:23 found in the strained ECM model. The black atom is the central atom

about 0.8 nm, which lies in the MRO length scale. This is shown in figure 5.5. The existence of this cluster is rare in CRN models [159, 170]. This observation tends to suggest two key points:

- If the simulation is allowed to run for considerably long time, more Schlafli clusters would be formed and the final model would closely resemble a PC model.
- The PC model is probably the only model that can match FEM data and have a corresponding real space structure.

5.4 Conclusion

We have demonstrated how to model realistic configurations of *a*-Si having MRO as measured by the variance of the diffracted intensity of amorphous thin films in FEM experiments. Models having more MRO are strained and a reduction in the strain implies a reduction in the amount of

order. Our results suggests that the PC model is not the only model that matches FEM experiments. However, the PC connects the FEM signal to real space in a consistent manner (nanometer-sized crystalline grains) whereas our the ECM models do not. Nevertheless, we believe this is a good starting point toward the atomistic interpretation of MRO as measured by FEM experiments.

Chapter 6

Systematic Study of Electron Localization in an Amorphous Semiconductor

The importance of defects in amorphous semiconductors cannot be overemphasized. As we explained in chapter 1, coordination defects in *a*-Si are of key importance to bulk and transport properties. Full understanding of the nature of defects in *a*-Si can help unlock the mystery behind phenomena like the Staebler-Wronski effect [3] and help establish the link between localization of defect states and large electron-phonon coupling. Early theoretical work on defect states in *a*-Si and *a*-Si:H was based on tight-binding methods [172, 173, 174, 175, 176, 177]. For example, Biswas *et al.* [172] and Fedders and Carlsson [174] investigated the electronic structure of dangling and floating bonds in *a*-Si. They showed that gap defect states associated with dangling bonds are strongly localized on the central atom of the dangling bond [172]. More recently, density functional calcula-

tions of dangling bond states using the local density approximation (LDA) have been performed by Fedders and Drabold [178]. They reported a wave function localization of 10-15% on the central atom in supercell models with one defect and far less on supercell models with many defects due to defect band formation. This finding was at variance with electron spin resonance (ESR) experiments, which showed that over 50% of spin density of the gap state is located on the central atom of the dangling bond [179, 180]. However, recent calculations by Fedders *et al.* [181], using the local spin density approximation (LSDA) have shown that a large spin localization of a defect state does not necessarily imply the existence of a corresponding wave function localization.

The Kohn-Sham formulation of density functional theory (DFT), which we have discussed in chapter 2, has enjoyed enormous success in describing the ground state properties and defects for a wide range of materials, and in particular, *a*-Si. We also made mention of the fact that the single-particle Kohn-Sham energies in the theory are mere mathematical entities that facilitates the calculations and therefore do not have no formal justification as quasi-particle energies. However from an empirical point of view, Hybertsen and Louie have shown, using GW calculations [182], that for states close to the fundamental band edges of bulk semiconductors and in particular Si (which are the states that we are interested in), there is a 99.9% overlap of the quasi-particle wave function with the corresponding Kohn-Sham orbital (GW calculations provide post Hartree self-energy corrections to DFT/LDA). This provides some rationale for interpreting the Kohn-Sham orbitals as quasi-particle states.

Because of the importance of defects in semiconductor theory and applications, they have to be accurately represented in electronic structure calculations. If the electronic structure of defect states are performed using DFT, then there are two key issues that need to be examined: basis functions used to expand the eigenstates and approximation for the exchange-correlation functional.

Within plane-wave DFT calculations, there is a general problem for the accurate representation of localized mid-gap and band tail states in amorphous semiconductors. To accurately treat defect states, several plane wave basis functions needed to expand localized eigenstate whereas only few localized orbitals are needed. For local basis DFT methods, there is the general problem of lack of systematic convergence of the basis set. This means that unlike plane waves, merely increasing the size and of local basis set and changing the shape of the individual basis functions does not necessarily guarantee that high quality results. However, localized basis provides a natural way to quantify wave-function localization as opposed to plane waves. The second important issue we raised above was the reliability of approximations to the exchange-correlation functionals to correctly reveal the localized behavior of electronic states with respect to its wave function and spin. The fact is that the generalized gradient approximation (GGA) and the local density approximation (LDA) sometimes make different predictions. For example, recent density functional and GW studies of the metal-insulator transition of bcc hydrogen showed that eigenfunctions obtained using GGA are more localized [183, 184] and closer to quasi-particle energies and states compared to using LSDA. Also, it was observed that GGA band gap was systematically larger than LSDA gap [183]. This does not necessarily make GGA methods better than LDA methods. There are situations in which LDA methods have outperformed GGA. Finally, there is the question of spin localization and whether or not it also basis set or exchange-correlation functional dependent. Spin localization is important because recent results have shown that to accurate study defects, spin density functionals calculations should be employed to interpret electron spin resonance (ESR) experiments [181].

In this Chapter, we systematically examine the dependence of the wave-function and spin localization of band tail and gap states on basis sets, density functionals for three defected models: two 216-atom supercell models of amorphous Si and a 218-atom supercell model of crystalline

Si:H with a vacancy. The crystalline model will serve as a benchmark to compare the nature of a dangling bond defect state in an amorphous environment with that of the crystalline phase. We compare the localization of gap and tail states within the LDA, LSDA and GGA for frozen static lattices (unrelaxed samples) as well as for fully relaxed static lattices for a given Hamiltonian. Our motivation for performing the frozen lattice calculation is to systematically investigate the sole effect of different basis sets and density functionals on the localization of states for a fixed configuration. We study the relaxation effects to see the dependence of the local geometry of the defect sites on the different basis orbitals and density functionals and the behavior of localized defect states in a relaxed environment compared to the frozen one. We computed spin and wave function localization for defect states, to determine the correlation between spin density and charge density.

6.1 Models and Calculations

Two of the defect models used here were generated by Barkema and Mousseau [132] using an improved version of WWW algorithm [131] discussed in chapter 3. Each model contain a pair of dangling bond defect. To create a dangling bond, a ghost bond (bond with zero energy) is placed between two atoms. The two atoms are chosen in such a way that they are quite close in one model and reasonably far in another model. In this work we refer these two 216-atom supercell models of *a*-Si as CLOSE and FAR. In model CLOSE the atoms with the dangling bonds are separated by a distance of 4.6 Å, whereas in the model FAR they are 7.7 Å apart. The third model, which we will refer to as *c*-Si:H, is generated as follows: starting with a 216-atom cell of silicon in the diamond structure, an atom is removed resulting in the creation of a vacancy with four

dangling bonds. Three of the dangling bonds were terminated by placing an H atom at a distance of about 1.5 Å from each of them. This results in one isolated dangling bond.

Our DFT calculations were performed within the LDA (with and without spin polarization) and the GGA using the code SIESTA [74, 75, 76]. We used the parameterization of Perdew and Zunger [49] for the exchange-correlation functional in all LDA calculations and the parametrization of Perdew, Burke and Ernzerhof [52] for the exchange-correlation functional in all GGA calculations. Norm conserving Troullier-Martins pseudopotentials [70] factorized in the Kleinman-Bylander form [72] were used to remove core electrons. To describe the valence electrons, we used atomic orbitals basis set consisting of finite-range numerical pseudo-atomic wave functions of the form discussed in section 2.10. We employ single-ζ (SZ) and double-ζ with polarization functions (DZP) basis sets on all the atoms. We solved the self consistent Kohn-Sham equations by direct diagonalization of the Hamiltonian. The Γ point was used to sample the Brillouin zone in all calculations.

To characterize the spatial extent of an electronic state, we employ three measures of localization discussed in section 2.11.2.

6.2 Results

6.2.1 Frozen Static Lattice Calculations I: Wave Function Localization

In order to understand the effect of basis set and density functional on the localization of defects states, we first compute the electronic structure of the three models with fixed geometry, using the different basis sets and density functionals.

We begin our discussion with simplest of the three models, *c*-Si:H. The position and localization for individual states are reported in Figs. 6.1 and 6.2, where the IPR and $\frac{1}{\mathrm{W(E)}}$, respectively, are used to measure localization. Each spike indicates an energy eigenvalue. With both measures of localization, and for all basis sets and density functionals considered, we observe a highly localized state near at Fermi level, with all the other states being extended. This state is centered on the dangling bond atom, which contributes with a Mulliken charge of $0.54e$ and $0.57e$ in the SZ-LDA and SZ-GGA cases respectively, and $0.27e$ and $0.31e$ in the DZP-LDA and DZP-GGA cases respectively. The rest of the state is mainly localized in the neighbor atoms to the dangling bond. Both the IPR and $\frac{1}{\mathrm{W(E)}}$ provide the same qualitative picture of the localized state, and its evolution with basis set and density functional. In particular, localization decreases strongly when we go from the SZ basis to a more complete DZP basis. It also increases but only by a small amount when using GGA instead of LDA. These are general trends which we will also observe with the other models, as we will see below. We note that the decrease in the measure of localization is not proportional to the decrease in the Mulliken charges at the defect state, as both definitions of localization are nonlinear. While the charge at the dangling bond site is reduced by a factor of two when moving from SZ to DZP, the localization measure decreases by roughly a factor of three, both for IPR and $\frac{1}{\mathrm{W(E)}}$.

Unlike $\mathrm{I(E)}$ and $\mathrm{W(E)}$ that are point estimators of the localization (in the sense that they use only the Mulliken charge at each atomic site), σ^2 is a more physical representation. The results obtained for σ^2 for *c*-Si:H with the lattice frozen are presented in Table 6.1. The spread for the localized state M is simple to compute since it is unimodal (peaked at only one site and therefore having a well defined center). From σ^2, we also see monotonic increase in the spread and a decrease in the total charge in the localization volume as the basis sets are increased (see Table 6.1). The GGA states show a slightly smaller spread than the corresponding LDA states. In Fig. 6.3 we show

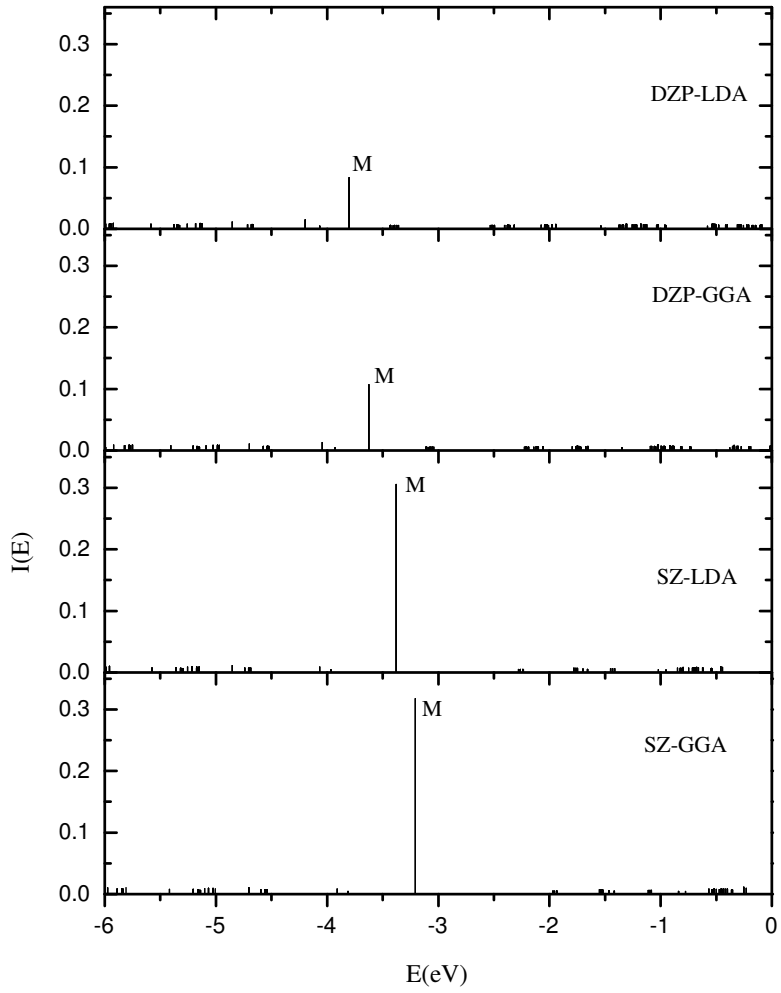

Figure 6.1: IPR for the model c-Si:H computed using frozen coordinates. The highly localized mid-gap state (labeled M) sits on the central atom of the dangling bond. For the SZ basis functions the charge localization on the central atom within the LDA and GGA are respectively $0.54e$ and $0.57e$. For the DZP basis sets, the charge localization on the atom reduce to $0.27e$ and $0.31e$ respectively within the LDA and GGA.

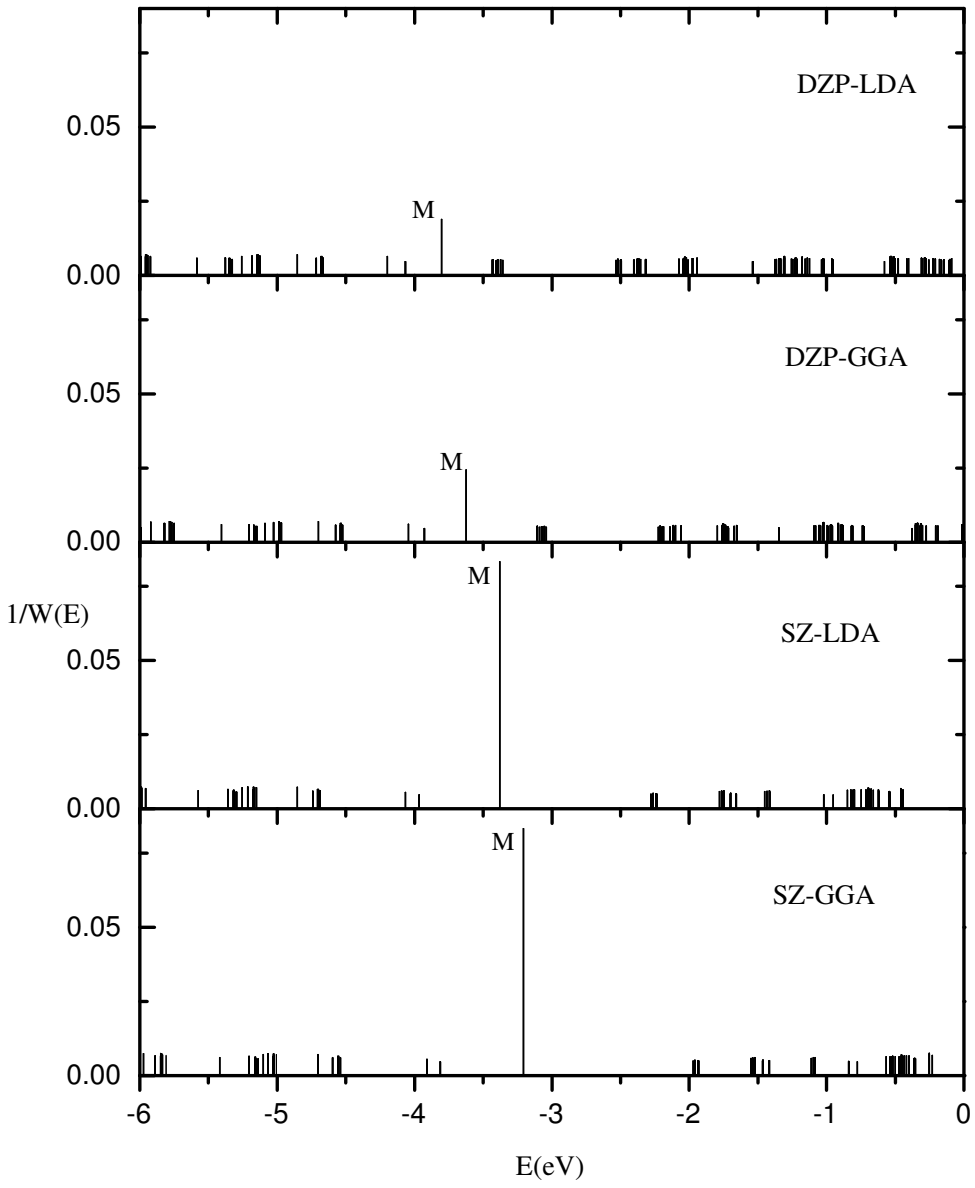

Figure 6.2: Localization of states (as measured using $\frac{1}{W(E)}$) for the model *c*-Si:H using the frozen lattice. The only localized mid-gap state (labeled M) sits on the central atom of the dangling bond.

Table 6.1. The frozen lattice results for the spread σ^2 and the charge integrated over a sphere q_Ω for the localized mid-gap state M for the supercell c-Si:H. Unit for σ^2 is Å^2.

Functional	Basis	σ^2	q_Ω
LDA	DZP	32.14	0.80
GGA	DZP	28.46	0.83
LDA	SZ	19.44	0.90
GGA	SZ	18.55	0.91

snapshots of the isosurface of the wave function for the localized mid-gap state of c-Si:H within two approximations. We see a dangling bond orbital confined to a small region in space in the SZ case implying that the state is well localized. In the DZP case, we see pieces of the surface in the vacancy and other neighboring atoms besides the dangling bond orbital making it less localized compared to the SZ case.

Next, we analyze the localized nature of the states for the amorphous model CLOSE. In this system, we expect to see two localized states in the gap, corresponding to the two dangling bonds present in the structure. Indeed, we see two highly localized gap states H and L in the IPR shown in Fig. 6.4, with an energy splitting which is just over a tenth of an eV. The state H is the highest occupied molecular orbital (HOMO) and L is the lowest unoccupied molecular orbital (LUMO). Due to the small distance between the two dangling bonds, the localized H and L states are bonding and anti-bonding combinations of the dangling bond states, and therefore both H and L have nearly equal weights on the two defect sites. In the SZ case, the total Mulliken charge

contributions from the two dangling bond sites for each of the two localized wave functions range between $0.52e$ and $0.64e$. The charge concentrations drop to the range $0.29e$-$0.40e$ in the DZP case. In Fig. 6.5, we show the results for $\frac{1}{\mathrm{W(E)}}$. We see a sharp drop in the number of atoms a given eigenstate can reach as the energy changes from the edge of either the valence or conduction band into the gap. We again see that all the features of $\mathrm{I(E)}$ are reproduced in $\frac{1}{\mathrm{W(E)}}$. For each localized state, both measures decrease by approximately a factor of 2 when the basis sets are increased from SZ to DZP.

In Fig. 6.6, we show the IPR for the model FAR. As in the CLOSE case, both the HOMO and LUMO states are localized. Now, however, since the distance between the two dangling bond sites is larger, the HOMO-LUMO splitting is much smaller, only ~ 6 meV. The HOMO has now most of its weight on one of the dangling bonds, whereas the LUMO is mostly localized in the other. The trend in localization of the gap states is similar to the the trend observed in the other two two models, decreasing strongly with more complete basis sets.

Since the wave functions associated with the gap states in the unrelaxed CLOSE and FAR models do not have a single center, but are peaked at the two dangling bonds, we will not analyze the localization by means of the spread in these cases. However, as we show in next section, relaxation leads in some cases to localization of the wave functions around one of the dangling bonds, and this will allow us to use the spread in such cases to quantify localization.

Our frozen lattice calculations consistently show that increasing the basis set decreases the localization. Although this is not unexpected, the huge decrease in the localization of the wave function as the size of the basis functions increase from SZ to DZP is quite surprising. The fact that the results are consistent in both the amorphous and crystalline system makes it even more interesting and general. A plausible explanation for this effect is that the energy gap is clearly

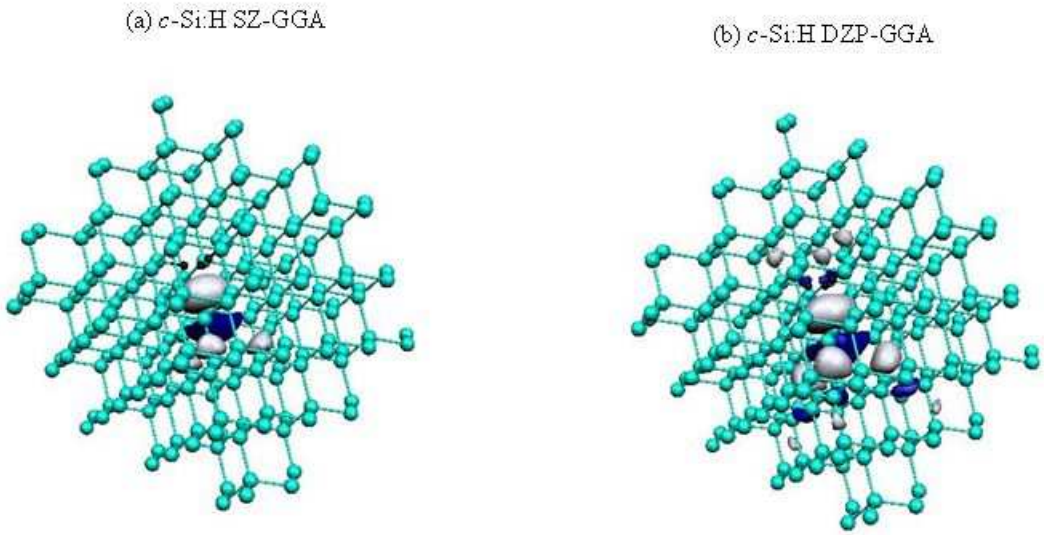

Figure 6.3: Isosurface plots for localized wave functions corresponding to a *c*-Si:H dangling bond defect state. The wave functions were generated with the same cut-off. Each surface is labeled according to the basis set and functional used. The surface is confined to a small region in space in the simple SZ case. For the DZP basis we see a localized dangling bond orbital with pieces of the surface in the vacancy and other neighboring sites. H atoms are colored black.

reduced as the basis set is more complete. The localized states are then closer to the band edges, and therefore are more able to mix with the extended bulk states, becoming more delocalized. Obviously, the delocalization will not proceed indefinitely upon improvement of the basis set, but will converge as the basis approaches completeness.

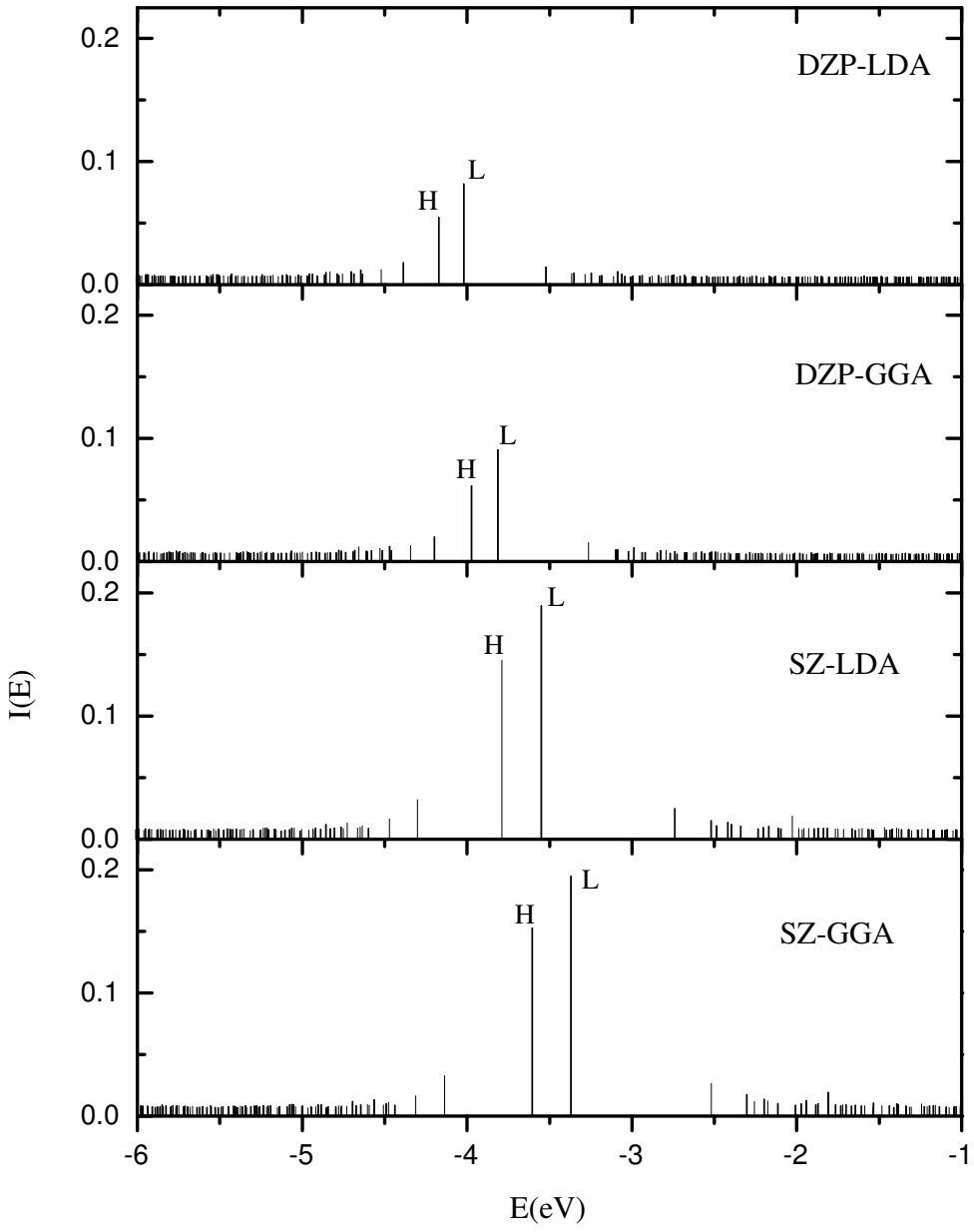

Figure 6.4: IPR for the model CLOSE using frozen coordinates. The two highly localized mid-gap states sit on the central atoms of the two dangling bonds. The state labeled H is the highest occupied molecular orbital (HOMO) and the state labeled L is the lowest unoccupied molecular orbital (LUMO).

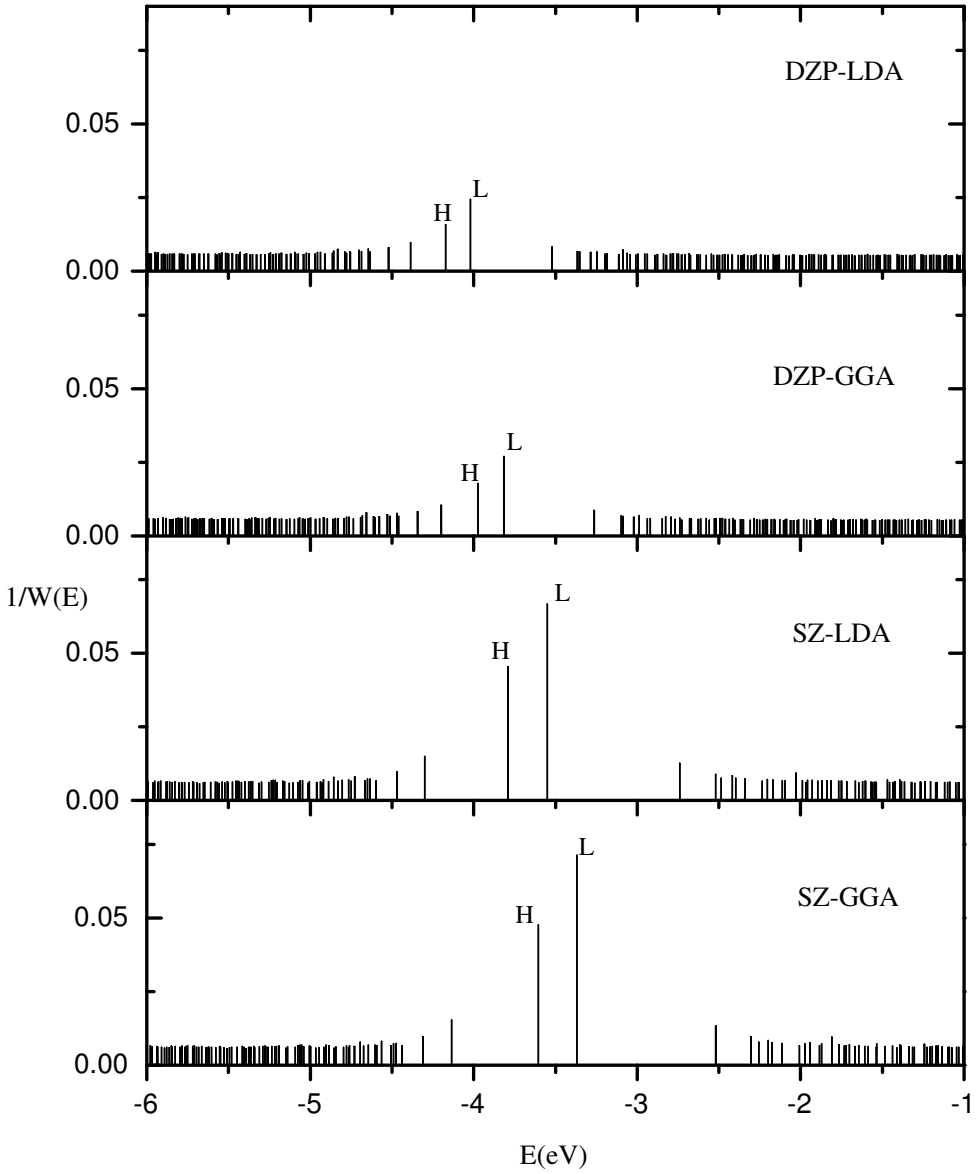

Figure 6.5: $\frac{1}{W(E)}$ for the model CLOSE with the atoms frozen.

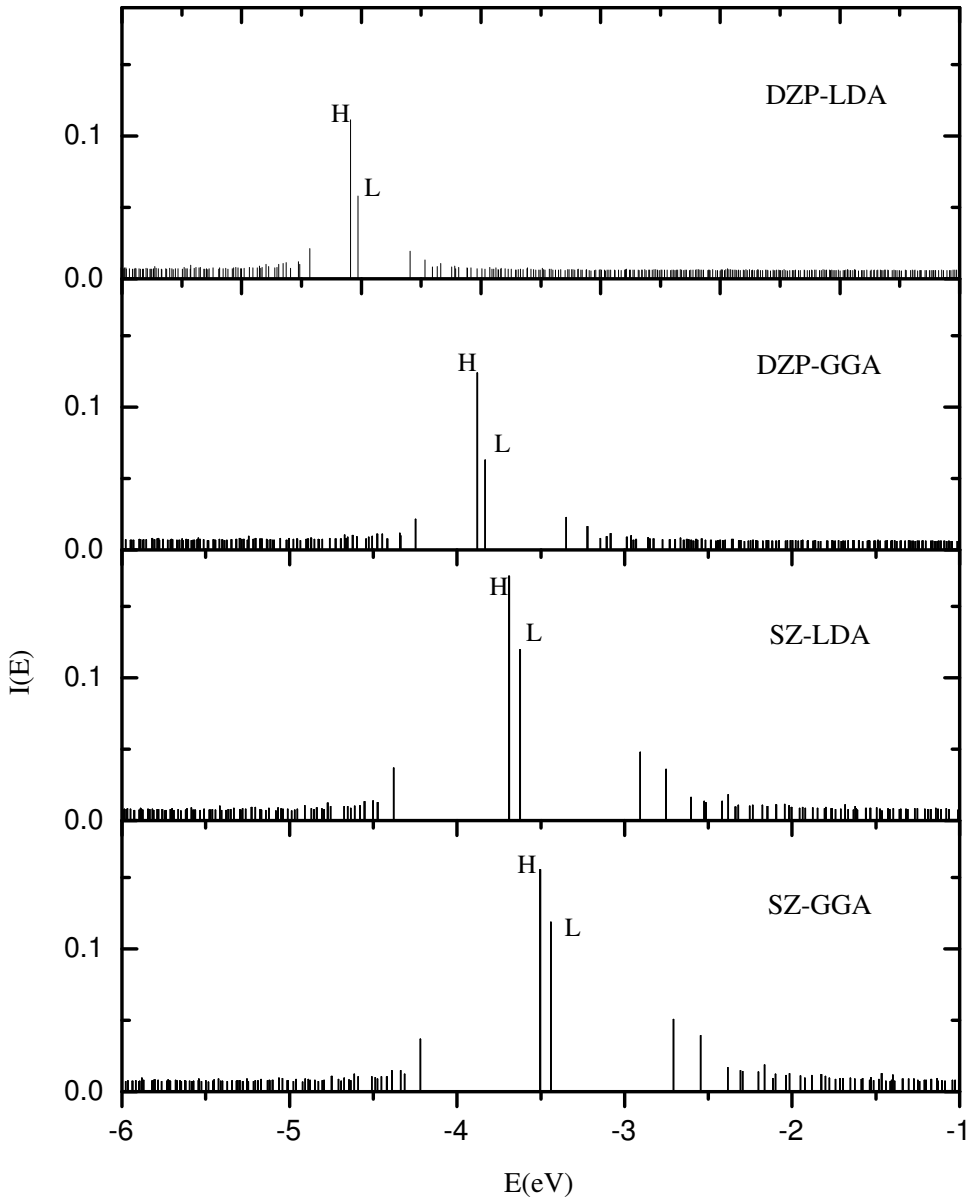

Figure 6.6: IPR for the model FAR using frozen coordinates. The two highly localized mid-gap states (H for HOMO and L for LUMO) are nearly degenerate.

6.2.2 Frozen Static Lattice Calculations II: Spin Localization versus Wave Function Localization

Fedders and co-workers [181] have shown that, in order to correlate the degree of localization from dangling bond states with ESR experiments, it is not enough to look at the wave functions, but to the net spin polarization near the dangling bond. The reason is that the spin density has also contributions from electronic states other than the localized defect wave function, which contribute to make the spin polarization more localized than the specific localized state wave function. In order to confirm this result (obtained by Fedders *et al.* on cells of a-Si:H) in our structural models, we performed calculations allowing for spin polarization in our frozen lattice models, using the DZP basis set. Except for the *c*-Si:H case, where there is one unpaired electron that yields a net spin polarization, we were not able to find a spin polarized solution for any of the amorphous cells. The reason is the existence of two interacting dangling bonds, which favors the formation of a spin singlet with two electrons paired. In order to force the appearance of a spin moment in our models, we introduce an unpaired spin by removing a single electron from the system.

In the model *c*-Si:H with, we find a contribution of almost 50% to net spin by the central dangling bond and its neighbors (the central atom alone contributing 38%). However, the Mulliken charge contribution to the wave function of the corresponding localized state from the defect site is only $0.29e$. The hydrogen-terminated dangling bond sites also contribute about 10% of the net spin. The remainder is somewhat distributed uniformly at the other sites. In CLOSE, about 57% of the net spin polarization was due the one dangling bond and its three neighbors. The other dangling bond contributed only 6% to the net spin with essentially zero contribution coming from the neighbors. In FAR, about 54% of the net spin localization sits on the isolated dangling bond and its nearest

neighbors. Our results are in good agreement with the previous LSDA calculations by Fedders *et al.*[181], and in reasonable agreement with the experiment [179, 180].

Our results confirm that, for a dangling bond defect state, there is a rather large difference between spin localization and wave function localization. In particular, the degree of spin localization is greater than that of the wave function localization at the dangling bond site. To our knowledge, no experimental methods exist for measuring the extent wave function localization on the dangling bond orbital as opposed to spin.

6.2.3 Relaxation Effects I: Geometry of Defect Sites, Density Functionals and Basis Sets

In this section we discuss the geometry around the defect sites before and after relaxation. The details of the local geometry are very important in determining the local electronic structure and the strain around the defect site, and here we study the dependence with varying basis and density functional. We relax all the models using a conjugate gradient optimization until the largest atomic force is smaller than 0.04 eV/Å.

Relaxation effects for the simple dangling bond defect in *c*-Si:H are relatively small. There is no major rearrangements in bonding, but only a relaxation of the surroundings of the vacancy site.

In the unrelaxed CLOSE model, the dangling bonds were originally separated by a distance of 4.6 Å. After SZ basis relaxations, both with LDA and GGA, the defect sites come closer, to a distance of about 2.6 Å, to form a highly strained bond. For the more complete DZP basis set, the two defect sites also approach each other, but they continue being under-coordinated, so the two distinct dangling bonds remain present.

In the case of the FAR model, one of the two well separated dangling bonds forms a strained bond with a neighboring atom after a SZ-LDA relaxation. The dangling bond is therefore terminated, and a floating bond is introduced. The other dangling bond remains present. For the SZ-GGA relaxation, the two dangling bonds still continue to exist, but the one that was terminated in the SZ-LDA case also approaches a neighbor and tries to form a bond. After the DZP relaxations, the dangling bonds still exist both in LDA and GGA.

Our results indicate that SZ basis tends to favor tetrahedral bond formation whereas DZP qualitatively preserves the original structure with the dangling bonds present. Also, the SZ tends to favor the transformation of dangling bonds into floating bonds. The results therefore suggest that the richer DZP basis set is necessary for an accurate description of the geometry of both isolated and clustered dangling bonds in amorphous silicon. The SZ basis is not flexible enough to provide sufficient freedom to describe the different shape of the wave function at the dangling bonds compared to covalent sp^3 bonds (for which the basis is ideally suited). Therefore, it tends to favor the disappearance of the dangling bonds through annihilation with other dangling bonds or formation of floating bonds with other already fourfold coordinated atoms.

6.2.4 Relaxation Effects II: Localization, Density Functionals and Basis Sets

We now consider in detailed the trends in the localization behavior of the electronic states in the gap after full structural relaxations. We also examine the role the defect site in a relaxed environment plays in the localization of gap states. We study the localization using the IPR for each of the three fully relaxed models within the different approximations.

We first studied the simple dangling bond defect in *c*-Si:H. Fig. 6.7 shows that, for the relaxed structures, the localization behavior of the mid-gap state is density functional dependent

Table 6.2. The spread and integrated charge for the localized mid-gap state M for the relaxed c-Si:H model.

Functional	Basis	σ^2	q_Ω
LDA	DZP	34.24	0.79
GGA	DZP	27.14	0.84
LDA	SZ	38.42	0.76
GGA	SZ	32.07	0.81

but rather basis set independent, contrary to the results obtained in the frozen lattice calculations (Fig. 6.1). In other words, within the same density functional approximation, SZ yields a similar wave function localization as DZP for the simple defect in the relaxed crystalline environment. We also see that the GGA defect state is more localized than the LDA defect state for a given basis set. The IPR values obtained with DZP are, nevertheless, almost unchanged upon relaxation, the difference between the unrelaxed and relaxed geometries occurring mainly for the SZ basis. The analysis from the real space spread, shown in Table 6.2, confirms the results obtained via the IPR. We see again the similarity between SZ and DZP for a given functional, with the GGA states having a smaller spread than LDA.

The IPR for the fully relaxed CLOSE is shown in Fig. 6.8. As discussed in the previous subsection, structural relaxation for this model is basis dependent. We first see that the splitting between the HOMO and LUMO levels is now much larger than in the unrelaxed case. This is attributed to the fact that, in order to minimize the total energy, occupied defect states move down

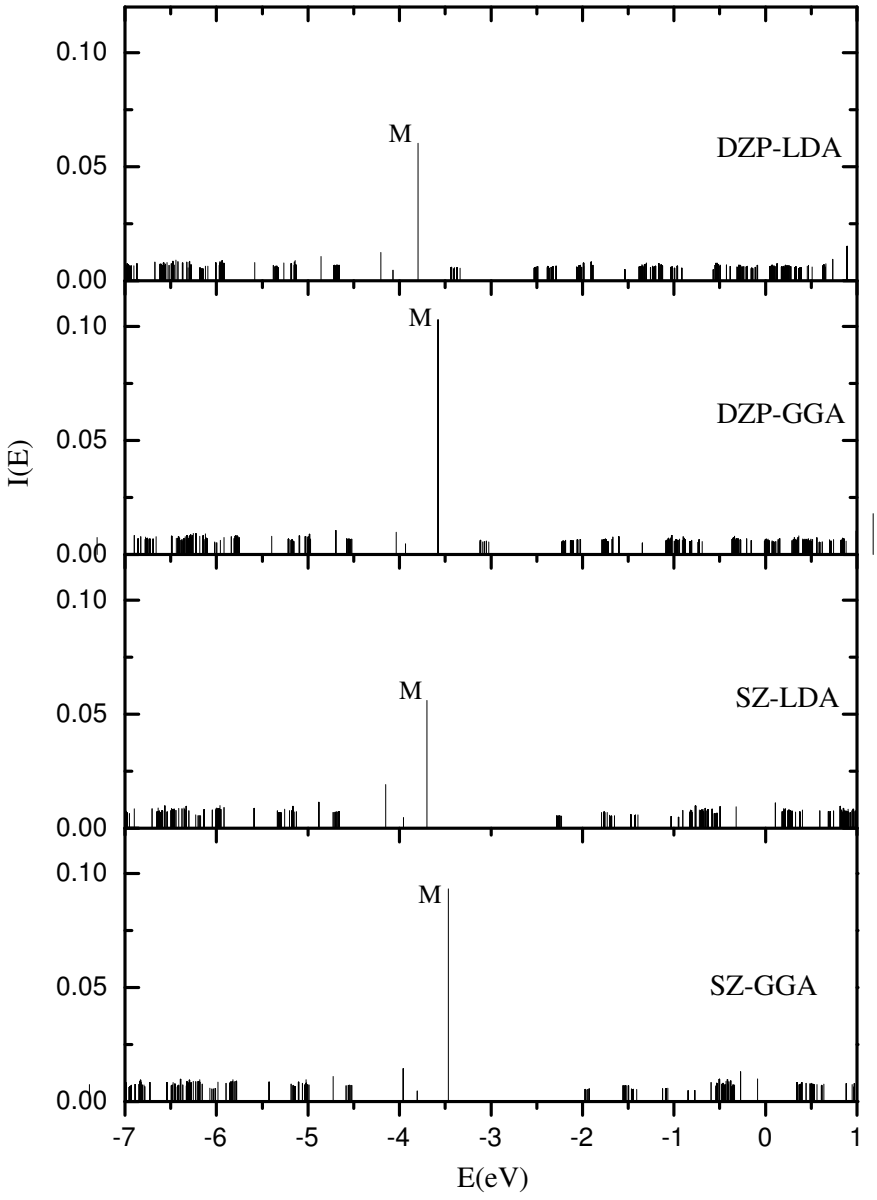

Figure 6.7: IPR for the fully relaxed model *c*-Si:H.

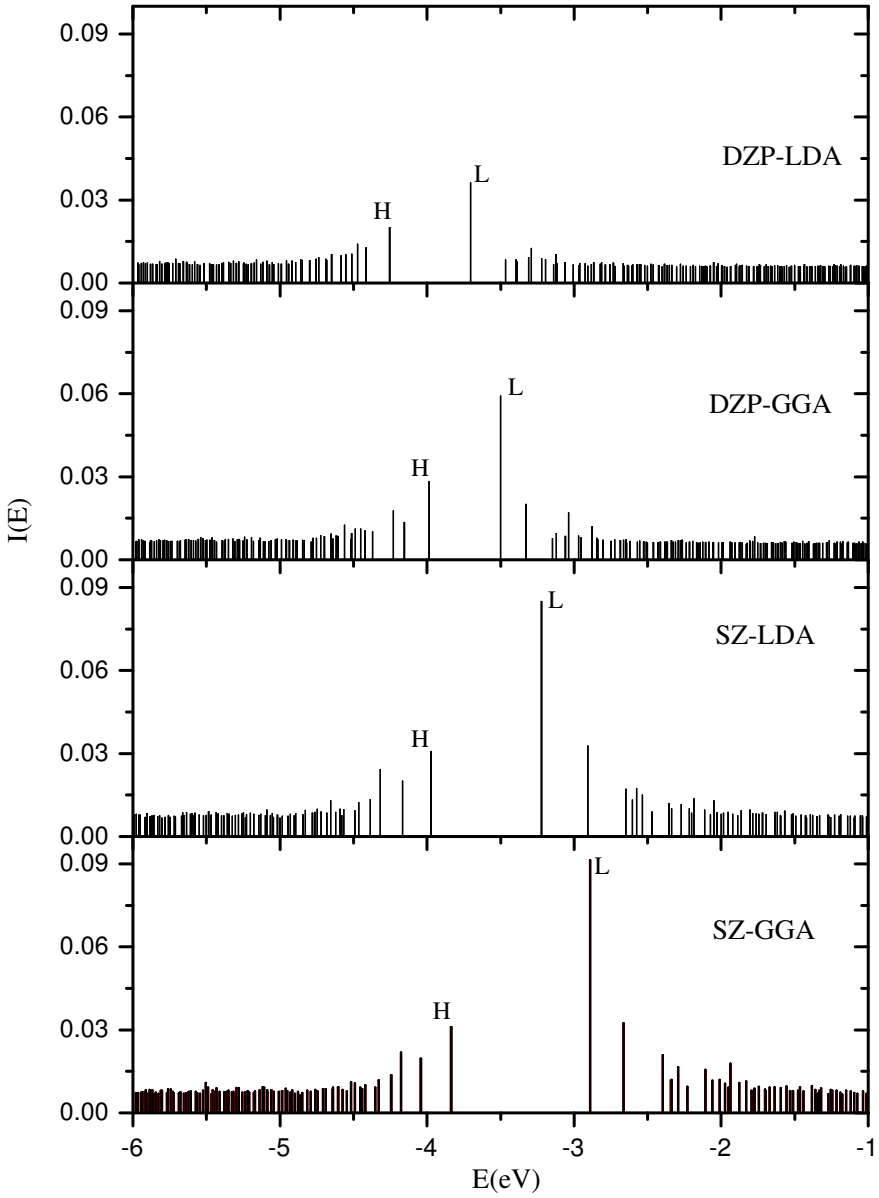

Figure 6.8: IPR for the fully relaxed model CLOSE.

in energy and closer into the valence band edge, whereas the unoccupied states do not affect the energy and thus can move toward the conduction band edge. The second observation is that now the localization of both HOMO and LUMO has decreased considerably compared to the unrelaxed case. Again, this is a consequence of the levels being closer to the band edges, mixing more strongly with the delocalized bulk states and therefore becoming less localized. The effect is larger for the HOMO, which is the one that adjusts its shape to optimize the total energy. A third observation is the appearance of increasingly localized states in the band edges, corresponding to bulk states which start becoming localized and form the precursor of band tails. This effect is originated from the strain field imposed by the relaxation of the sites around the defects. Therefore, the presence of defects like dangling bonds in amorphous silicon also brings the appearance of band tails of weakly localized states due to the introduction of stress in the surroundings of the defect. This supports results from previous work [86] that there is not a one to one correspondence between spectral and geometric defects. The localization of the tail states decays as one moves deeper into the conduction and valence regions, as was previously observed by Dong and Drabold [134] using a simple orthogonal tight-binding Hamiltonian on a much larger 4096-atom model of a-Si.

We now focus on the evolution of localization with basis set and density functional for the relaxed CLOSE model. As we observed with the c-Si:H case, the HOMO level becomes less localized upon relaxation, specially in the case of the SZ basis, for which the IPR is reduced by more than a factor of two. The degree of HOMO localization predicted by the SZ and DZP bases for the relaxed structure is therefore very similar. For the LUMO, the difference between SZ and DZP is still large, as in the unrelaxed case. The high IPR values associated with the LUMO in the SZ cases are primarily due to strain (as a result of the bond between the two neighboring dangling bonds atoms). The spread of the LUMO is reported in Table 6.3. We again observe the common

Table 6.3. The results for σ^2 and q_Ω corresponding to the localized LUMO state for the relaxed CLOSE model.

Functional	Basis	σ^2	q_Ω
LDA	DZP	31.85	0.76
GGA	DZP	28.14	0.79
LDA	SZ	23.50	0.84
GGA	SZ	22.45	0.87

trends in the spread: GGA states show slightly less spread compared to LDA, and SZ basis set yield more localized states than DZP basis set.

In order to get a pictorial representation of the localized states in this relaxed model and the evolution with basis set and density functional, we assign different colors to each site according to its Mulliken charge contribution to the given eigenstate. We depict this spatial feature by showing only 65% of the total charge for the LUMO in Fig. 6.9. We observe a small network connection of atoms for the localized states in Figs. 6.9(a) and 6.9(b) but the connectivity spreads out in a rather 1D fashion, mimicking a chain in Figs. 6.9(c) and 6.9(d). The small size of our cell does not allow us to immediately visualize a localized region containing completely the cluster of atoms, which can be done in larger supercell containing thousands of atoms [134].

We now turn to the relaxed FAR model. In Fig. 6.10, we have plotted the IPR for the this model, for different density functionals and basis sets. Again, we see the same general features that we pointed out in the relaxed CLOSE case. First, the splitting between HOMO and LUMO is much

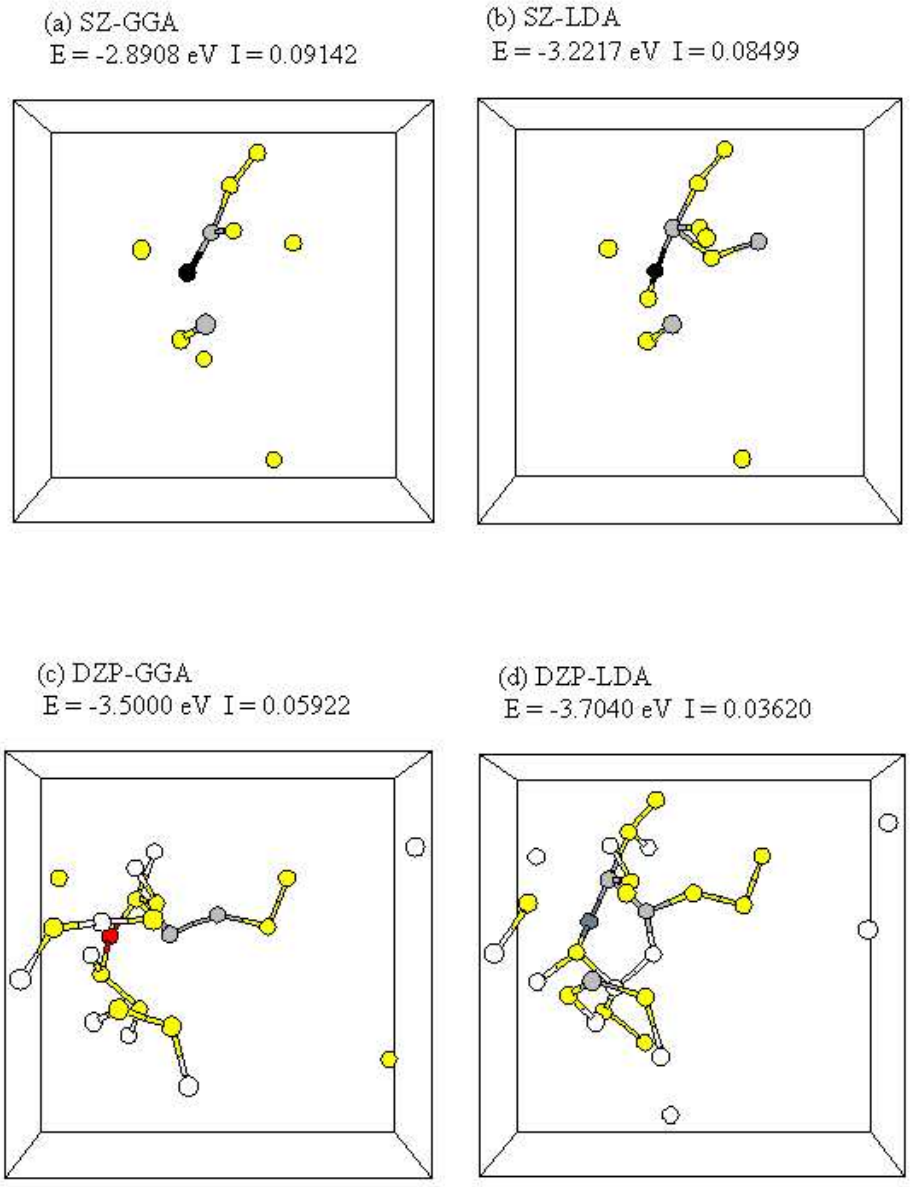

Figure 6.9: Spatial character of localized eigenstates for the LUMO state for the relaxed CLOSE model. The energy and its corresponding IPR localization are indicated in each picture. We use the following color code to depict the fraction of the Mulliken charge q for the localized state at each atomic site: black (q $>$ 0.25), red (0.15 $<$ q $<$ 0.25), slategray (0.10 $<$ q $<$ 0.15), gray (0.05 $<$ q $<$ 0.10), yellow (0.01 $<$ q $<$ 0.05) and white (q $<$ 0.01). Only 65% of the total charge is shown.

larger than in the unrelaxed case. Both states, and specially the HOMO, become more delocalized upon relaxation, with the notable exception of the SZ-GGA case, which yields a strongly localized HOMO state. We also see the formation of band tail states, and even the presence of strongly localized states in the gap above the LUMO, due to strong relaxation induced strain fields. Again, in this model, we observe that localization is stronger for GGA than with LDA, and that the difference in localization between SZ and DZP bases is much reduced upon relaxation. The difference in localization for the SZ-GGA and SZ-LDA cases can be explained by the fact that, as mentioned previously, SZ-LDA relaxation results in the disappearance of a dangling bond and the formation of a floating bond, which are known to be less localized than dangling bond defects.

Finally, we visualize a chosen state (HOMO in this case) for the relaxed FAR model using a color coding in Fig. 6.11 The Mulliken charge concentrations on the atoms changes from confined cluster-like character (or equivalently short 1D strings) in Fig. 6.11(a) to long string-like character of atoms in Fig. 6.11(d) as one tunes the basis and functional from SZ-GGA through to DZP-LDA . The ”tinker-toy” character can be attributed to less localized states and it is mainly due to weak quantum mechanical mixing. This behavior has been observed earlier by Drabold *et. al*. [134, 185]

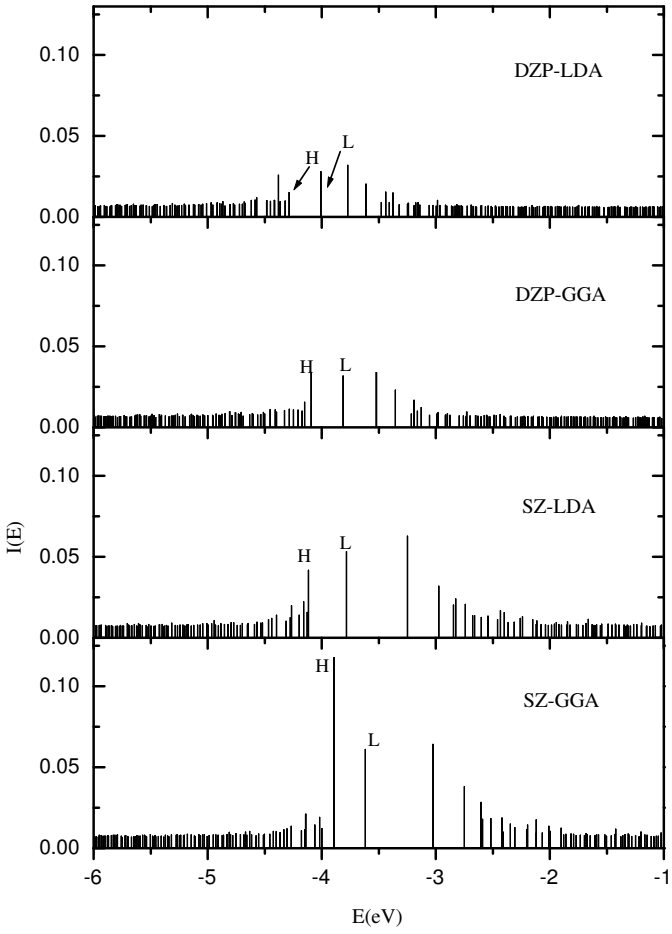

Figure 6.10: IPR for the fully relaxed model FAR.

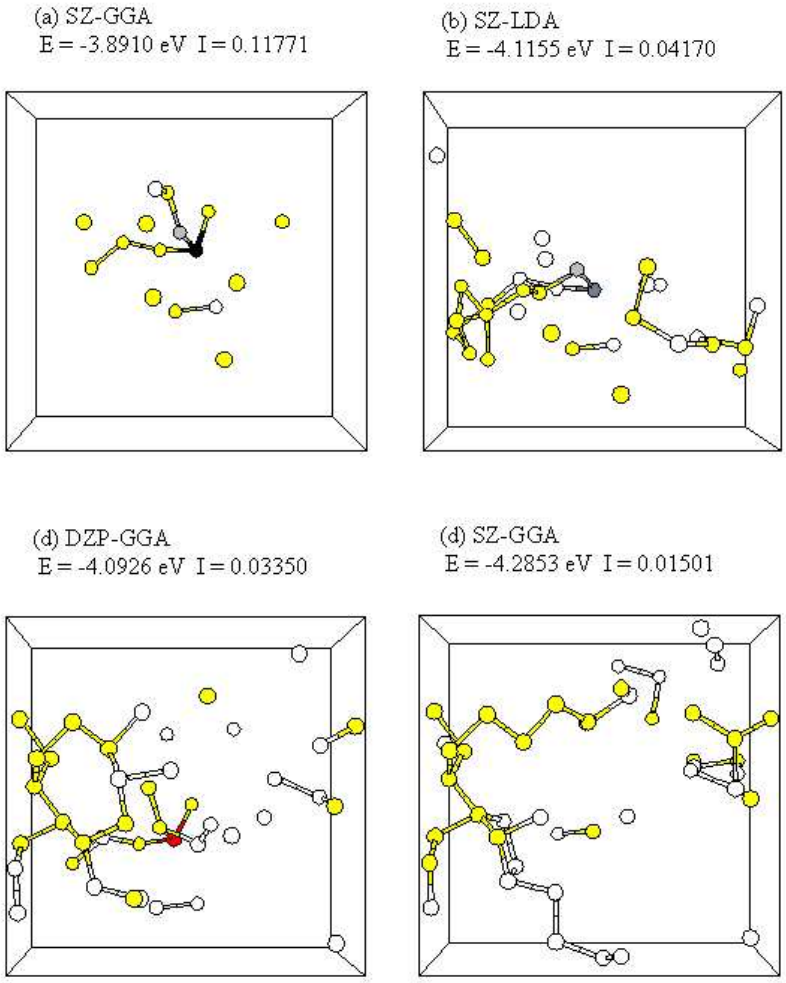

Figure 6.11: Spatial character of localized eigenstates for the HOMO state for the relaxed FAR model. The color coding is the same as the CLOSE case in Fig. 6.9.

6.3 Conclusion

We have performed a first principles electronic structure calculation on three Si supercells: two 216-atom supercells of amorphous silicon with two dangling bonds and one 218-atom supercell of hydrogenated crystalline silicon with a void. Depending on the initial distance between the dangling bonds, the two a-Si models have been classified as CLOSE and FAR. We examined the nature of localized band tail and gaps states within the LDA and GGA using both minimal SZ basis and more complete DZP basis with particular attention paid to relaxation effects. Spin localization and wave function localization for dangling bond defect states states has also been studied. We computed the wave function localization as the spread of the wave function in real space and via other measures that utilize the Mulliken charges.

For the frozen lattice calculations, we find that the localization of wave functions associated with defect states decrease with larger basis sets and has enhancement of localization using GGA compared to LDA for all the models. The reduction in charges at the atomic sites for a larger basis set can be attributed to the hybridization between the atomic orbitals, providing the electronic charges more degrees of freedom to redistribute themselves. This is reflected in a smaller distance between the defect states and the band gap edges, that also favors delocalization. Unpaired spin LSDA calculations performed on frozen lattices showed that the degrees of spin and wave function localization are different. In particular, degree of spin localization at a dangling bond site is far greater than the degree of wave function localization. The difference between the localization of a defect state in a fully relaxed and frozen systems is non-trivial, especially in the minimal basis calculations. In particular, there was a considerable reduction in localization (as measured using the Mulliken charge) for the relaxed systems compared to the frozen lattices. We also conclude that a

large basis set (DZP in this case) is necessary for an accurate description of both the geometry and localized states associated with defect sites.

Chapter 7

The Electron-Phonon Coupling is Large for Localized States

The electron-phonon (e-p) coupling, the interaction term connecting the electronic and lattice systems, is a physical quantity of key importance. Perhaps most spectacularly, the e-p coupling is the origin of superconductivity as expressed in BCS theory [186]. Phillips [187] has shown that large e-p couplings in the cuprate superconductors can lead to a successful model of high T_c superconductivity [188] within the framework of conventional BCS superconductivity. The e-p coupling is also the mediator of all light-induced structural changes in materials. In amorphous silicon, the greatest outstanding problem of the material, the Staebler-Wronski effect [3], depends critically upon the electron-lattice interaction. A zoo of analogous effects is studied in glasses; perhaps the most important example is reversible photo-amorphization and photo-crystallization used in the GeSbTe phase-change materials used in current writable CD and DVD technology.

The electron-phonon interaction and localized electron states in disordered systems have a well-defined relationship. First, this interaction is the root of thermally driven hopping between

localized states [189]. Thermally driven hopping is the principal mechanism for conduction at sufficiently low temperatures in disordered systems. In a seminal paper by Anderson [190], he showed, using a random lattice model, that mobiles electrons will be localized unless they are thermally activated and that at low enough defect concentrations, no diffusion takes place in the network. Second, Overhof and Thomas [191] have argued that an electron, starting in some localized state, will suffer scattering with the phonons and become progressively delocalized. Using the Anderson model [190], they showed that delocalization of electron states induced by phonons is necessary to understanding transport in disordered systems. The two factors enumerated above suggest that for localized states, the interaction of electrons with the lattice is expected to be strong. If the interaction is quantifiable, then its magnitude should be reasonably large for localized electron states.

Previous thermal simulations with Bohn-Oppenheimer dynamics have indicated that there exists a large electron-phonon coupling for the localized states in the band tails and in the optical gap [192, 189]. Earlier works on chalcogenide glasses by Cobb and Drabold [193] have emphasized a strong correlation between the thermal fluctuations as gauged by root mean square (RMS) variation in the LDA eigenvalues and wave function localization of a gap or tail state measured by inverse participation ratio, a simple measure of localization which we have discussed in chapter 2. Drabold *et al*. [86, 178] have also shown that localized eigenvectors may fluctuate dramatically even at room temperature. Recently, Li and Drabold [192] relaxed the adiabatic (Born-Oppenheimer) approximation to track the time-development of electron packets scattered by lattice vibrations. In this chapter, we examine the electron-phonon coupling and provide a heuristic analysis of the e-p coupling for localized electron states. We explore the e-p coupling in some detail for a particular model system (amorphous silicon) which provides us with a convenient variety of localized, partly localized "bandtail" and extended states. The e-p coupling is determined by computing the defor-

mation potential (which measures the response of a selected electron state to a particular phonon). We also track thermally-induced fluctuations of electronic eigenvalues. We find that localized states always exhibit a large e-p coupling. We perform Born-Oppenheimer molecular dynamics using the first principles code SIESTA [74, 75, 76], and the eigenvalues and states that we study are from the Kohn-Sham equations with a rich local basis orbitals. A rationale for the study of the Kohn-Sham states is given elsewhere [194]. We emphasize that the results that we give are qualitatively general – not just an artifact of studying a disordered phase of silicon (we have, for example, seen exactly the same effects in various binary glasses which exhibit very different topological and chemical disorder).

7.1 Theory

To establish a connection between electron-phonon coupling and wave function localization for the electrons, we consider an electronic eigenvalue λ_n near the band gap of a-Si. The sensitivity of λ_n due to an arbitrary small displacement of an atom (possibly thermally induced) can be estimated using the Hellmann-Feynman theorem [110],

$$\frac{\partial \lambda_n}{\partial \mathbf{R}_\alpha} = \langle \psi_n | \frac{\partial \mathbf{H}}{\partial \mathbf{R}_\alpha} | \psi_n \rangle.$$

Here we have assumed that the basis functions are fixed and $|\psi_n\rangle$ are the eigenvectors of the Hamiltonian $\mathbf{H}$. For small lattice distortion $\{\delta \mathbf{R}_\alpha\}$, the corresponding change in $\delta\lambda_n$ is,

$$\delta \lambda_n \approx \sum_{\alpha=1}^{3N} \langle \psi_n | \frac{\partial \mathbf{H}}{\partial \mathbf{R}_\alpha} | \psi_n \rangle \, \delta \mathbf{R}_\alpha \tag{7.1}$$

where N is the total number of atoms in the model. If the displacement $\delta\mathbf{R}_\alpha(\mathrm{t})$ arises from classical vibrations, one can write [1],

$$\delta\mathbf{R}_\alpha(\mathrm{t}) = \sum_{\mathrm{s}=1}^{3\mathrm{N}} \mathrm{A}(\mathrm{T},\omega_\mathrm{s})\ \cos(\omega_\mathrm{s}\,\mathrm{t} + \phi_{\omega_\mathrm{s}})\,\chi_\alpha(\omega_\mathrm{s}), \tag{7.2}$$

where s indexes the normal mode frequencies ω ($\neq 0$), $\mathrm{A}(\mathrm{T},\omega)$ is the temperature dependent amplitude of the mode with frequency ω, ϕ_ω is an arbitrary phase, $\chi_\alpha(\omega)$ is a normal mode with frequency ω and vibrational displacement index α. Using the temperature dependent squared amplitude $\mathrm{A}^2 = 3\mathrm{k_B T}/\mathrm{M}\omega^2$, the trajectory (long time) average of $\delta\lambda_\mathrm{n}^2$ can be written (using Eq. 7.1 and 7.2) as,

$$\langle\delta\lambda_n^2\rangle = \lim_{\tau\to\infty}\frac{1}{\tau}\int_0^\tau dt\,\delta\lambda_n^2(t) \approx \left(\frac{3\mathrm{k_B T}}{2\mathrm{M}}\right)\sum_{\mathrm{s}=1}^{3\mathrm{N}}\frac{\Xi_\mathrm{n}^2(\omega_\mathrm{s})}{{\omega_\mathrm{s}}^2}, \tag{7.3}$$

where the electron-phonon coupling $\Xi_\mathrm{n}(\omega)$ is given by,

$$\Xi_n(\omega) = \sum_{\alpha=1}^{3N}\langle\psi_n|\frac{\partial\mathbf{H}}{\partial\mathbf{R}_\alpha}|\psi_n\rangle\,\chi_\alpha(\omega). \tag{7.4}$$

One can infer from Eq. 7.3 that thermally induced fluctuation in the energy eigenvalues is a consequence of electron-phonon coupling. Note that for a given electronic eigenvalue, the contribution to the coupling comes from the entire vibrational spectrum involving all the atoms in the systems. Since the normalized eigenstate can be written as $|\psi_n\rangle = \sum_i a_{ni}\,|i\rangle$ where $|i\rangle$ are the basis orbitals, it follows from Eq. 7.4 that,

$$\begin{aligned}
\Xi_n^2(\omega) &= \sum_{\alpha,\beta,i,j,k,l} a_{ni}^* a_{nj} a_{nk}^* a_{nl}\langle i|\frac{\partial\mathbf{H}}{\partial\mathbf{R}_\alpha}|j\rangle\,\langle k|\frac{\partial\mathbf{H}}{\partial\mathbf{R}_\beta}|l\rangle\,\chi_\alpha(\omega)\,\chi_\beta(\omega) \\
&= \sum_{i,\alpha}|a_{ni}|^4\,[\langle i|\frac{\partial\mathbf{H}}{\partial\mathbf{R}_\alpha}|i\rangle]^2\,\chi_\alpha^2(\omega) \\
&+ {\sum_{ijkl\alpha\beta}}' a_{ni}^* a_{nj} a_{nk}^* a_{nl}\langle i|\frac{\partial\mathbf{H}}{\partial\mathbf{R}_\alpha}|j\rangle\,\langle k|\frac{\partial\mathbf{H}}{\partial\mathbf{R}_\beta}|l\rangle\,\chi_\alpha(\omega)\,\chi_\beta(\omega)
\end{aligned} \tag{7.5}$$

[1] Here we have only considered the Γ point in our calculations which corresponds to $\vec{k}$=0

The first term in the second line of Eq. 7.5 is positive definite (diagonal) while the second one, the off-diagonal term (indicated by the prime), is not of a single sign. In the event that only a few a_{ni} dominate (the case for localized states), then the leading contribution to the electron-phonon coupling originates largely from the diagonal term. The addition of a large number of terms of mixed sign and small magnitude leads to cancellations in the off-diagonal term leaving behind a small contribution to electron-phonon coupling. By comparing to direct calculations with the full Eq. 7.5, we show that dropping the second term appears to be reasonable for well-localized electron states. The approximate "diagonal" electron-phonon coupling can be written as,

$$\begin{aligned} \Xi_n^2(\omega) &\approx \sum_{\alpha=1}^{3N}\sum_{i=1}^{N_b} |a_{ni}|^4 [\langle i|\frac{\partial \mathbf{H}}{\partial \mathbf{R}_\alpha}|i\rangle]^2 \chi_\alpha^2(\omega) \\ &= \sum_{\alpha=1}^{3N}\sum_{i=1}^{N_b} q_{ni}^2 [\langle i|\frac{\partial \mathbf{H}}{\partial \mathbf{R}_\alpha}|i\rangle]^2 \chi_\alpha^2(\omega) \end{aligned} \tag{7.6}$$

where N_b is the number of basis orbitals and $q_{ni} = |a_{n,i}|^2$ is the charge sitting on the ith orbital for a given normalized eigenstate $|\psi_n\rangle$. The degree of wave function localization can be measured by defining inverse participation ratio $\mathcal{I}$ for the eigenstates $|\psi_n\rangle$,

$$\mathcal{I}(n) = \sum_{i=1}^{\mathrm{N_b}} q_{ni}^2. \tag{7.7}$$

Equation 7.6 leads to an approximate but analytic connection between $\mathcal{I}$ and electron-phonon coupling. Since $\mathcal{I}$ is large for localized states, one expects $\Xi_n(\omega)$ (and therefore $\langle \delta\lambda_n^2 \rangle$) to be large for a localized state. If we further assume that $\gamma^2(\omega, i) = \sum_{\alpha=1}^{3N} [\langle i|\partial \mathbf{H}/\partial \mathbf{R}_\alpha|i\rangle]^2 \chi_\alpha^2(\omega)$ is weakly dependent upon site/orbital index i, then

$$\Xi_n^2(\omega) \sim \mathcal{I}_n \times f(\omega), \tag{7.8}$$

where $f(\omega)$ is defined from γ^2 and Eq. 7.6. In this "separable" approximation, it is also the case that $\langle \delta\lambda_n^2 \rangle \propto \mathcal{I}_n$.

7.2 Methodology

The model of *a*-Si we have used in our calculations was generated by Barkema and Mousseau [132] using an improved version of the Wooten, Winer and Weaire (WWW) algorithm [131]. The details of the construction was reported in Ref. [132]. The model consists of 216 atoms of Si packed inside a cubic box of length 16.282 Å and has two 3-fold coordinated atoms. The average bond angle is 109.5° with a root mean square deviation of 11.0°. The density functional calculations were performed within the local density approximation (LDA) using the first principles code SIESTA [74, 75, 76]. We have used a non self-consistent version of density functional theory based on the linearization of the Kohn-Sham equation, the Harris functional approximation [53] described in chapter 2. The parameterization of Perdew and Zunger [49] for the exchange-correlation functional. Norm conserving Troullier-Martins pseudopotentials [70] factorized in the Kleinman-Bylander form [72] were used to remove the core electrons. The choice of an appropriate basis is found be very important and has been discussed at length in Chapter 6. While the minimal basis consisting of one s and three p electrons can adequately describe the electronic structure of amorphous silicon in general, there is some concern about the applicability of these minimal basis in describing deeply localized and low lying excited states in the conduction bands accurately. We have therefore employed a larger single ζ basis with polarization (d) orbitals (SZP) in the present work. Throughout the calculation we have used only the Γ point to sample the Brillouin zone.

Starting with a fully relaxed configuration, we construct the dynamical matrix elements by successively displacing each atom in the supercell along three orthogonal directions (x, y and z) by a small displacement δd and computing the forces for each configuration. The construction of the dynamical matrix within the harmonic approximation was described in section 1.4. We have checked

the convergence of the matrix elements by using a different set of values for atomic displacement and for our calculations, we have used $\delta d = 0.01$ Å.

To explore the validity of our analysis and to elucidate the connection between the localization (IPR) ($\mathcal{I}$) of electronic eigenstates and fluctuation of the conjugate eigenvalues, we performed thermal MD simulations at constant temperatures using the Noŝe-Hoover [114] method. The Noŝe-Hoover method is a simulation in the (N,V,T) ensemble in which a heat bath is coupled to the physical system, and heat enters or leaves the system to maintain thermal equilibrium. The simulations were performed at temperatures 150 K, 300 K, 500 K and 700 K with a time step of 2.5 femtoseconds for a total period of 2.5 picoseconds. For a given temperature, the mean square fluctuations were computed by tracking the eigenvalues at each time step and averaging over the total time of simulation excepting the first few hundred time steps to exclude particle energies before equilibration. The mean square fluctuation ($\mathcal{R}$) for a set of energy eigenvalues $\{\lambda_{\mathrm{n}}(\mathrm{t})\}$ is defined as :

$$\mathcal{R}_n = \langle(\lambda_n(t) - \langle\lambda_n\rangle)^2\rangle, \tag{7.9}$$

where $\langle\,\rangle$ denotes the average over time. We study the fluctuations of $\{\lambda_{\mathrm{n}}(\mathrm{t})\}$ by plotting against time at a given temperature and compare it with the $\mathcal{I}$ obtained for the corresponding eigenvalues. For illustrations of such adiabatic evolution of Kohn-Sham eigenvalues, see Refs. [86] and [178].

7.3 Results

In Fig. 7.1, we have plotted the electron-phonon coupling for the states near the band gap obtained directly from Eq. 7.4. It is clear from the figure that the e-p coupling is large only in the vicinity of conduction and valence band tails. The largest e-p coupling in the plot corresponds

to the highest occupied molecular orbital (HOMO) in the optical frequency regime around 415 cm^{-1}. The lowest occupied molecular orbital (LUMO) also has a large feature around the same frequency. A Mulliken charge analysis and inverse participation ratio calculation of the electronic eigenfunctions have shown that both these two states – the HOMO and LUMO are highly localized and are centered around the dangling bonds present in the model. On moving further from the band tails in either direction along the energy axis, the e-p coupling drops quickly and the surface becomes featureless for a given eigenvalue. This behavior of e-p coupling can be understood from the arguments presented in section 7.1 where we have shown that the e-p coupling for localized states is directly proportional to the inverse participation ratio. For a localized state, therefore, the large value of electron-phonon coupling can be attributed to the large value of inverse participation ratio associated with that state. Since HOMO and LUMO are the two most localized states in the spectrum, the e-p coupling is large for these states and as we move toward the tail states, the coupling decreases. It is important to note that the plot in the Fig. 7.1 has been obtained from Eq. 7.4 without making any approximation and is exact inasmuch as the matrix elements obtained from the density functional Hamiltonian are correct. This observation supports our assumptions that the dominant contribution to e-p coupling comes from the diagonal term in Eq. 7.5 and that $\gamma^2(\omega, i)$ is weakly dependent upon site/orbital index i and also indicates from direct simulation there exists a linear relationship between mean square fluctuation of electronic eigenvalues and the corresponding inverse participation ratio for localized states.

In order to justify our arguments further presented in section 7.1, we now give a look at the mean square fluctuation of energy eigenvalues. As outlined in section 7.2, we have computed the mean square fluctuations at four different temperatures (150 K, 300 K, 500 K and 700 K) from MD runs over a period of 2.5 picoseconds and plotted in the Fig. 7.2. The fluctuation obtained this way

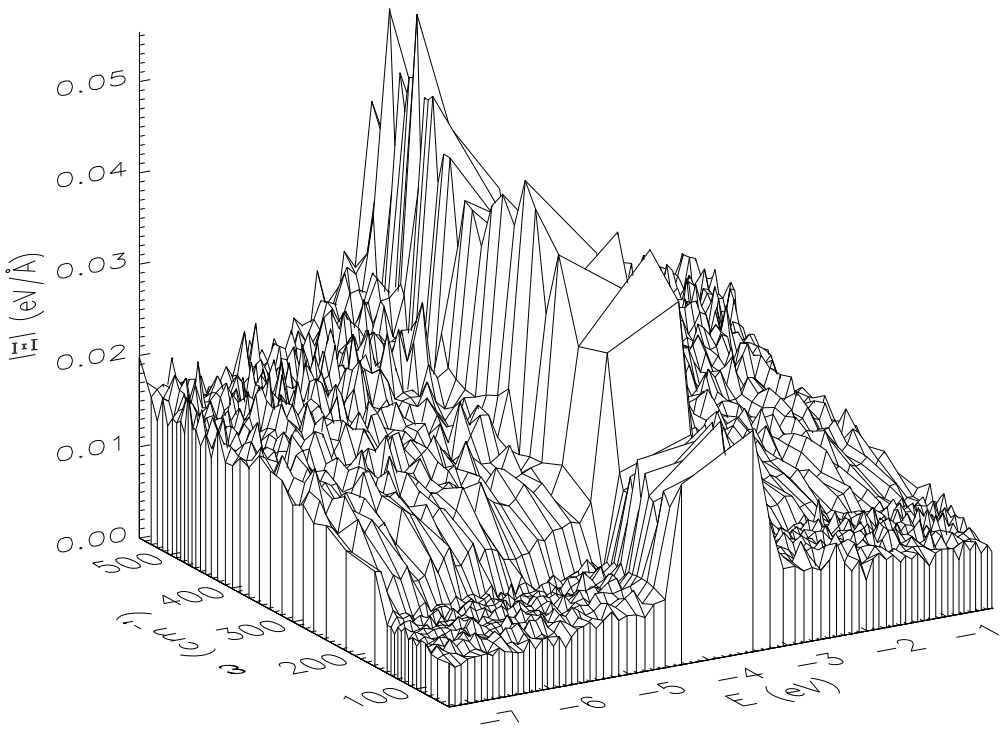

Figure 7.1: Electron-phonon coupling surface plot for a 216-atom model of *a*-Si. The absolute value of electron-phonon coupling $|\Xi|$ (cf. Eq. 7.4) is plotted as a function of phonon frequency ω and energy eigenvalues near the gap. The largest value of $|\Xi|$ in the plot corresponds to the eigenvalue for HOMO, which is the most localized state in the spectrum.

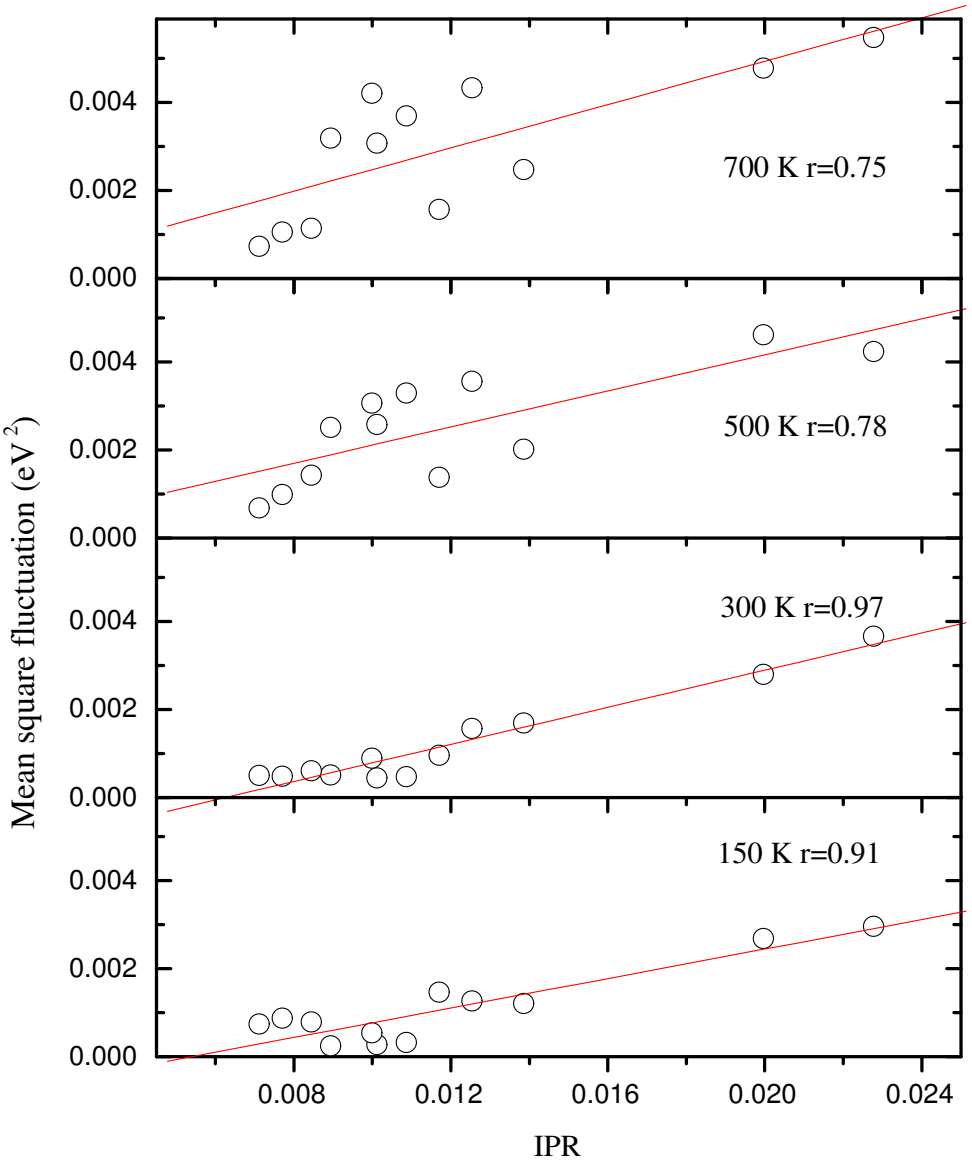

Figure 7.2: Mean square fluctuations of electronic eigenvalues versus inverse participation ratio plot at different temperature. The fluctuations at temperature 150 K and 300 K are found to be linearly correlated with the participation ratio for the corresponding eigenstates as predicted in section 7.1. The correlation coefficient (r) for different temperature is indicated in the plot.

provides a dynamical characteristic of the band tails states and is compared with a static property, the inverse participation ratio of the same states. A simple linear fit reveals a strong correlation between the eigenvalue fluctuation and the corresponding inverse participation ratio for the states. The correlation is found to be as high as ≈ 0.95 for T = 150 K and 300 K and falls to ≈ 0.8 at high temperature. The value of the correlation coefficient for different temperature is indicated in the Fig. 7.2. A reduction in the correlation coefficient at high temperature is not surprising since the harmonic approximation begins to break down. Once again, we see that the result is in accordance with our prediction in section 7.1 and provides a simple physical picture for having a large electron-phonon coupling for the localized states.

In Fig. 7.3, we have plotted the time averaged electronic density of states for four different temperature in order to study the effect of thermal disorder on the tail states. It is quite clear from the figure that the effect of thermal broadening is quite significant on both sides of the gap. Photo-electron spectroscopic studies on *a*-Si:H by Aljishi *et al*. [195] have shown that the conduction tail is indeed more susceptible to thermal disorder than the valence tail. The temperature dependence can be conveniently expressed by introducing a characteristic energy $\mathrm{E_0}$ and fitting the electronic density of states to $\rho(\mathrm{E}) \approx \exp(|\mathrm{E} - \mathrm{E_f}|/\mathrm{E_o}(\mathrm{T}))$. Aljishi *et al.* expressed the temperature dependence of the tail states by the slope of the $\mathrm{E_o}(\mathrm{T})$ vs. T plot and obtained a smaller value for the conduction band tail. We have observed a qualitative agreement of our results with experiment. The key observation that one should note from Fig. 7.3 is the following: the shape of the tail in the conduction band rapidly changes as the temperature rises from 150K to 700K. The corresponding change in the valence tail for the same range of energy (0.4eV) is however much less and is rather smooth compared to the conduction tail. Since the localized defect states (coming from the two dangling bonds) have been removed before plotting, this observation qualitatively suggests that the conduction tail

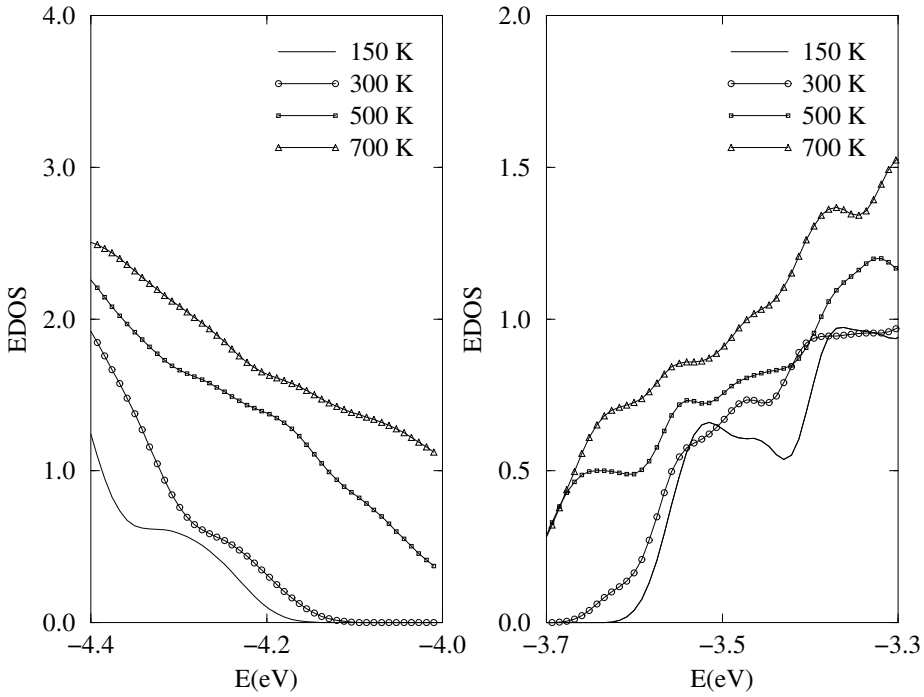

Figure 7.3: The average electronic density of band tails states for four different temperature T = 150 K, 300 K, 500 K and 700 K. Note that the conduction band tails (right) near the Fermi level (which is between the two tails) are more sensitive to thermal disorder than the valence band tails (left) providing a qualitative agreement with experimental result in Ref. [195].

states are more susceptible to lattice motion. It is tempting to attempt to estimate decay parameters for a direct comparison to experiment [195], but the sparse sampling of tail states for this 216-atom model makes this a dangerous exercise. The basic features do appear to be represented in our study, however.

7.4 Conclusion

Using accurate methods and a reasonable model of a-Si, we showed that there is

1. a large e-p coupling for localized states, and

2. a significant correlation between thermal fluctuation of electron energy eigenvalues conjugate to localized states and the IPR of the model at $\mathrm{T} = 0\mathrm{K}$.

We found a qualitative agreement with photoemission experiments [195], and provided a simple analytic argument for the origin of these effects. Identical experience with models of other amorphous materials has convinced us that the results are correct in at least a qualitative way for binary glasses and amorphous materials, and perhaps other systems beside.

Chapter 8

Concluding Remarks

8.1 Summary

In this dissertation, we have demonstrated, using a reverse Monte Carlo (RMC) method, how experimental information (and topological constraints) can be used to model a continuous random network (CRN) of amorphous silicon. Although our RMC model does not have the high quality possessed by the WWW models, it is improves dramatically over old RMC and molecular dynamics models and has fewer defects. Other than certain structural properties, old RMC methods did not critically address the electronic properties of the models. Our work took RMC a step further by examining the electronic spectrum as well. RMC has the flexibility which allows it to be applied to other systems. We have exploited the said flexibility to include fluctuation electron microscopy (FEM) experimental information in models of *a*-Si. We demonstrated how the experimental FEM signal, which is otherwise absent in the conventional CRN model, can be included in a CRN using constrained RMC. This suggests that there are disordered networks without crystallites that yield

FEM signals matching with experiments, implying that the FEM signal is not unique to the paracrystalline (PC) model.

For our study on the localization of defect electron states, we investigated the gap and band tail states in models of amorphous silicon. Starting with two 216-atom models of amorphous silicon with defect concentration close to the experiments, we systematically studied the dependence of electron localization on basis set, density functional and spin polarization using the first principles density functional code SIESTA. We compared three different schemes for characterizing localization: information entropy, inverse participation ratio and spatial variance. All calculations were carried out for static lattices. Our results showed that to accurately describe defected structures within self consistent density functional theory, a rich basis set is necessary. Our study revealed that the localization of the wave function associated with the defect states decreases with larger basis sets and there is some enhancement of localization from GGA relative to LDA. Spin localization results obtained via LSDA calculations, are in reasonable agreement with experiment and with previous LSDA calculations on *a*-Si:H models.

Finally, from density functional calculations, we showed that localized states stemming from defects or topological disorder exhibit an anomalously large electron-phonon coupling. We provided a simple analysis to explain the observation and performed a detailed study on amorphous silicon. We computed first principles deformation potentials (that is, the sensitivity of specific electronic eigenstates to individual classical normal modes of vibration). We also probed thermal fluctuations in electronic eigenvalues by performing first principles thermal simulations. We found a strong correlation between a static property of the network [localization, as gauged by inverse participation ratio (IPR)] and a dynamical property [the amplitude of thermal fluctuations of electron energy eigenvalues] for localized electron states. In particular, both the electron-phonon coupling

and the variance of energy eigenvalues are proportional to the IPR of the localized state. Our results indicated that the conduction band tails are more susceptible to thermal disorder than the valence band tails, which is in qualitative agreement with photo-emission experiments [195]. While the computations are carried out for silicon, very similar effects have been seen in other systems with disorder.

8.2 Further Considerations

Realistic atomistic configurations are a necessary ingredient for electronic structure computations. The RMC method has the ability to model atomistic systems using *a priori* information from experiments and network topological constraints. However, we believe that the RMC algorithm can be improved. One way to optimize the RMC algorithm to produce high quality models is to devise a scheme for selecting the constraint coupling parameters that will push the cost function very close to the minimum. At the moment, they are chosen using trial and error. Another factor that facilitates the convergence of the calculation is the rate at which the RMC simulation temperature is decreased. This should be thoroughly investigated and applied to a variety of more complex systems.

Although we have successfully used RMC and the modified WWW algorithm to generate *a*-Si configurations that matches FEM data, the work lacks a physical picture connecting the real space arrangement of atoms to the presence of the FEM signal in the model, which measures MRO. On the other hand, the PC model, which is indistinguishable from the CRN model as far as structural, electronic, and vibrational properties are concerned, has a corresponding real space picture (in the form of nanometer-sized crystalline regions) to back the FEM signal. Much work is therefore

needed on the atomistic interpretation of the signal in ECM models without crystallites. One way to address the problem, for example, is to devise an order parameter that is sensitive to the MRO.

For the systematic study on the localization of dangling bond defect states, we empirically observed how the single ζ (SZ) basis tends overestimate the measures of localization (by about a factor of 2 in certain cases) when compared double ζ with polarization (DZP) orbitals. Besides the fact the DZP basis are more complete than SZ basis, it is quite surprising that there such large differences in localization with respect to basis set size. This merits further study, especially as most previous calculations were performed with limited basis sets. A systematic study for a dynamical lattice would be a good starting point. It would also worth performing calculations on defected binary systems to see if similar effects exist.

Bibliography

[1] P. Voyles, *Fluctuation Electron Microscopy of Medium-Range Order in Amorphous Silicon Thin Films*, Ph.D. thesis, University Illinois at Urbana-Champagne, (2001).

[2] Cameron Walker, *The Future of Alternative Energy*, National Geographic News October 28, 2004; Michael Parfit, *Freedom*, National Geographic Magazine, August 2005.

[3] D. Staebler and C. R. Wronski, Appl. Phys. Lett. **31**, 292 (1977).

[4] S. R. Elliott, *Physics of Amorphous Materials* (Longman Group Limited, UK, 1984).

[5] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Holt Saunders, Philadelphia, 1976).

[6] W.H.Zachariasen, Jour. Amer. Chem. Soc. **54**, 3841 (1932).

[7] N. F. Mott, Adv. in Phys. **16**, 49 (1967).

[8] R. J. Bell and P. Dean, Nature (London) **212**, 1354 (1966); D. E. Polk, J.Non-Cryst. Solids **5**, 365 (1971).

[9] L. Guttman, AIP Conf. Proc. **31**, 268 (1976).

[10] J. Hansen and I. R. McDonald, *Theory of Simple Liquids* (Academic Press, London, 1986).

[11] K. Laaziri, S. Kycia, S. Roorda, M. Chicoine, J. L. Robertson, J. Wang, and S. C. Moss, Phys. Rev. Lett. **60**, 3460 (1999).

[12] D. Beeman, R. Tsu, and M. F. Thorpe, Phys. Rev. B **32**, 874 (1985).

[13] D. A. Drabold, *Molecular Dynamics Simulations of Network Glasses* in: P. Boolchand (ed.), *Insulating and Semiconducting Glasses* (World Scientific, Singapore 2000), pp. 607-651.

[14] P. Biswas, Phys. Rev. B **65**, 125208 (2002).

[15] P. Ordejón, D. A. Drabold, R. M. Martin, and S. Itoh, Phys. Rev. Lett. **75**, 1324 (1995).

[16] W. A. Kamitakahara, C. M. Soukoulis, H. R. Sanks, U. Buchenau, G. S. Grest, Phys. Rev. B **36**, 6539 (1987).

[17] L. Van Hove, Phys. Rev. **89**, 1189 (1953).

[18] S. Öberg, www.hpc2n.umu.se/projects/focus

[19] D. A. Drabold and T. A. Abtew, to be published.

[20] R. Martin, *Electronic Structure*: *Basic Theory and Practical Methods* (Cambridge University Press, Cambridge, 2004).

[21] M. Born and R. Oppenheimer, Ann. Phys. **84**, 457 (1927).

[22] Jan K. Labanowski, http://www.ccl.net/cca/documents/DFT/dft-overview/previous-version/dft.html.shtml.

[23] J. M. Combes, P. Duclos, R. Seiler, *The Born-Oppenheimer Approximation* in: G. Velo, A. S. Wightman (eds.), *Rigorous Atomic and Molecular Physics* (Plenum Press, New York, 1981) pp. 185-212.

[24] B. T. Sutcliffe, *The Nuclear Motion Problem in Molecular Physics*, Advances in Quantum Chemistry **28**, 65 (1997).

[25] F. Gemperle and F. X. Gadea, Europhysics Letters **48**, 513 (1999).

[26] L. J. Butler, Annu. Rev. Phys. Chem. **49**, 125 (1998).

[27] Robert G. Parr and Weitao Yang, *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989).

[28] A. Szabo, and N. S. Ostlund, *Modern Quantum Chemistry*: *Introduction to Advanced Electronic Structure Theory* (McGraw-Hill, New York, 1989).

[29] George Arfken, *Mathematical Methods of Physics* (Academic Press Inc., Orlando, 1985).

[30] V. Fock, Z. Physik **61** 126; **62**, 795 (1930).

[31] J. C. Slater, Phys. Rev. **35**, 210 (1930).

[32] http://www.hpc.susx.ac.uk/ venables/qmms4.html.

[33] D. R. Hartree, Proc. R. Soc. London A113, 621 (1928).

[34] C. C. J. Roothan, Rev. Mod. Phys. **23**, 69 (1951).

[35] C. C. J. Roothan, Rev. Mod. Phys. **32**, 179 (1960).

[36] P.W. Atkins and R.S. Friedman, *Molecular Quantum Mechanics* (Oxford University Press, Oxford, 1997).

[37] A. P. Sutton, *Electronic Structure of Materials* (Clarendon Press, Oxford, 1993).

[38] P. Fulde, *Electron Correlations in Molecules and Solids*, Springer Series in Solid State Sciences, Vol. 100, 3rd edition (Springer, Berlin, Heidelberg 1995).

[39] L. H. Thomas, Proc. Cambridge Philos. Soc. **23**, 542 (1927).

[40] E. Fermi, Z. Phys. **48**, 73 (1928).

[41] P. A. M. Dirac, Proc. Cambridge Philos. Soc. **26**, 376 (1930).

[42] P. Hohenberg and W. Kohn, Phys. Rev. 136, B864 (1964).

[43] W. Kohn, *Overwiew of Density Functional Theory*, in: E. K. U. Gross and Reiner M. Dreizler (eds.) *Density Functional Theory* (NATO ASI Series, Plenum Press, New York, 1995).

[44] W. Kohn and L. J. Sham Phys. Rev. **140** (4A), 1133 (1965).

[45] M. Levy, J. P. Perdew, and V. Sahni, Phys. Rev. A **30**, 2745 (1984).

[46] David Drabold, private communication, August 2005.

[47] D. M. Ceperly and B. M. Alder, Phys. Rev. Lett. **45**, 566 (1980).

[48] S. H. Vosko, L. Wilk, and M. Nussair, Can. J. Phys. **58**, 1200 (1980).

[49] J. P. Perdew and A. Zunger, Phy. Rev. B **23**, 5048 (1981).

[50] A. D. Becke, Phys. Rev. A **38**, 3098 (1988).

[51] J. P. Perdew and Y. Wang, Phys. Rev. B **45**, 13244 (1992).

[52] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[53] J. Harris, Phys. Rev. B **31**, 1770 (1985).

[54] V. I. Anisimov, F. Aryasetiawan, and A. I. Lichtenstien, J. Phys. Condensed Matter **9**, 767 (1997).

[55] V. I. Anisimov, J. Zaanen, and O. K. Andersen, Phys. Rev. B **44**, 943 (1991).

[56] F. W. Averill and G. S. Painter, Phys. Rev. B **41**, 10344 (1990).

[57] D. Sánchez-Portal, E. Artacho and J. M. Soler, Solid State Commun. **95**, 685 (1995).

[58] R. S. Mulliken, J. Chem. Phys. **23**, 1833 (1955).

[59] J. C. Phillips, Phys. Rev. **112**, 685 (1958).

[60] J. C. Phillips and Leonard Kleinman Phys. Rev. **116**, 287 (1959).

[61] http://ermes.physics.ncsu.edu/pseudopotentials.ppt

[62] A. Redondo, W. A. Goddard III and T. C. McGill, Phys. Rev. B **15**, 5038 (1977).

[63] D. R. Hamann, M. Schlüter and C. Chiang, Phys. Rev. Lett. **43**, 1494 (1979).

[64] A. Zunger and M. L. Cohen, Phys. Rev. B **20**, 4082 (1979).

[65] G. P. Kerker, J. Phys. C **13**, L189 (1980).

[66] G. B. Bachelet, D. R. Hamann and M. Schlüter, Phys. Rev. B **26**, 4199 (1982).

[67] D. R. Hamann, Phys. Rev. B **40**, 2980 (1989).

[68] A. M. Rappe, K. M. Rabe, E. Kaxiras and J. D. Joannopoulos, Phys. Rev. B **41**, 1227 (1990).

[69] J. S. Lin, A. Qteish, M. C. Payne and V. Heine, Phys. Rev. B **47**, 4174 (1993).

[70] N. Troullier and J. L. Martins, Phys. Rev. B 43, 1993 (1991).

[71] David Vanderbilt, Phys. Rev. B **41**, 7892 (1990).

[72] L. Kleinman and D. M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982).

[73] X. Gonze, R. Stumpf, and M. Scheffler, Phys. Rev. B**44**, 8503 (1991).

[74] P. Ordejón, E. Artacho and J. M. Soler, Phys. Rev. B **53**, 10441, (1996).

[75] D. Sánchez-Portal, P. Ordejón, E. Artacho and J. M. Soler, Int. J. Quantum Chem. **65**, 453 (1997)

[76] J. M. Soler, E. Artacho, J. D. Gale, A. García, J. Junquera, P. Ordejón, and D. Sanchez-Portal, J. Phys. Cond. Matter. **14**, 2745 (2002).

[77] J. Janquera, Ó. Paz, D. Sanchez-Portal, and E. Artacho, Phys. Rev B **64**, 235111 (2001).

[78] Andrew P. Horsefield, Phys. Rev. B **56**, 6594 (1997).

[79] O. F. Sankey and D. J. Niklewski, Phys. Rev. B **40**, 3979 (1989).

[80] E. Artacho, D. Sanchez-Portal, P. Ordejón, A. García, and J. M. Soler, Phys. Stat. Sol. **215**, 809 (1999).

[81] S. Huzinaga, Comput. Phys. Rep. **2**, 279 (1985); S. Huzinaga *et al*., *Gaussian Basis Sets for Molecular Calculations* (Elsevier Science, Amsterdam, 1984).

[82] P.O. Löwdin, J. Chem. Phys., **18**, 365 (1950).

[83] R. J. Bell and P. Dean, Discuss. Faraday Soc. **50**, 55 (1970).

[84] J. Pipek, Int. J. Quantum Chem. **36**, 487 (1989).

[85] J. Pipek, I. Varga, and T. Nagy, Int. J. Quantum Chem. **37**, 529 (1990).

[86] D. A. Drabold, P. A. Fedders, S. Klemm, and O. F. Sankey, Phys. Rev. Lett. **67**, 2179 (1991).

[87] C. Shannon, *The Mathematical Theory of Communication*, (University of Illinois Press, Urbana, 1949).

[88] E. T. Jaynes, *Probability Theory: The Logic of Science*, Cambridge University Press, Cambridge (2003).

[89] J. P. Lewis, J. Pikus, TH. E. Cheatham, E. B. Starikov, Hao Wang, J. Tomfor, and O. F. Sankey, Phys. Stat. Sol. (b) **233**, 90 (2002).

[90] J. P. Lewis, J. Pikus, TH. E. Cheatham, E. B. Starikov, Hao Wang, J. Tomfor, and O. F. Sankey, J. Phys. Chem. B **107**, 11, (2003).

[91] S. F. Boys, in *Quantum Theory of Atoms, Molecules and the Solid State*, (ed.) P.-O. Löwdin (Academic Press, New York, 1966), p. 253.

[92] N. Mazari and D. Vanderbilt, Phys. Rev. B **56**, 12847 (1997).

[93] P. L. Silvestrelli, N. Marzari, D. Vanderbilt, and M. Parrinello, Solid State Communications **107**, 7 (1998); U. Stephan and D. A. Drabold, Phys. Rev. B **57** 6391 (1998), U. Stephan, R. M. Martin and D. A. Drabold, Phys. Rev. B **62** 6885 (2000).

[94] D. W. Heermann, *Computer Simulation Methods* (Springer-Verlag, Berlin, 1989).

[95] D. C. Young, *Computational Chemistry* (Wiley-Interscience, New York, 2001).

[96] D. Frenkel and B. Smith, *Understanding Molecular Simulations from Algorithms to Applications* (Academic Press, San Diego, 1996).

[97] P. N. Keating, Phys. Rev. **145**, 637 (1966).

[98] F. H. Stillinger and T. A. Weber, Phys. Rev. B **31**, 5262 (1985).

[99] R. L. C. Vink, G. T. Barkema, W. F. van der Weg, and N. Mousseau, J. Non-Cryst. Solids **282**, 248 (2001).

[100] J. C. Slater and G. F. Koster, Phys. Rev. B **94**, 1498 (1954).

[101] D. J. Chadi and M. L. Cohen, Phys. Stat. Sol. (b) **68**, 405 (1975); D. J. Chadi, Phys. Rev. B **29**, 785 (1984) and references therein.

[102] L. Goodwin, A. J. Skinner, and D. J. Pettifor, Europhys. Lett. **9**, 701 (1989).

[103] J. L. Mercer Jr. and M. Y. Chou, Phys. Rev. B **49**, 8506 (1994).

[104] I. Kwon *et al.*, Phys Rev. B **49**, 7242 (1994).

[105] C. Z. Wang, C. T. Khan, and K. M. Ho, Phys. Rev. B **39**, 8586 (1989).

[106] T. J. Lenosky *et al.*, Phys. Rev. B **55**, 1528 (1997).

[107] M. Menon and K. R. Subbaswamy, Phys. Rev. B **50**, 11577 (1994).

[108] Th. Fraueheim *et al.*, Phys. Rev. B **52**, 11492 (1995).

[109] N. Bernstein *et al.*, Phys. Rev. B **62**, 4477 (2000).

[110] R. P. Feynman, Phys. Rev. **56**, 340 (1939); H. Hellman, *Einfuhrung in die Quantumchemie* (Franz Deutsche, Leipzig, 1937).

[111] C. M. Goringe, D. R. Bowler, and E. Hernandez, Rep. Prog. Phys. **60**, 1447 (1997).

[112] R. Car and M. Parrinello, Phy. Rev. Lett. **55**, 2471 (1985).

[113] R. Car, *Modeling Materials by Ab-Initio Molecular Dynamics* and associated references in: J. R. Chelikowsky and S. G. Louie (ed.) *Quantum Theory of Real Materials*, (Kluwer Academic Publishers, Boston, 1996).

[114] S. Nośe, *Molecular Dynamics Simulations at Constant Temperature and Pressure* in: M. Meyer and V. Pontikis (eds.) *Computer Simulations in Material Science*, (Kluwer Academic Publishers, Netherlands, 1991).

[115] H. B. Callen, *Thermodynamics and an Introduction into Thermostatistics*, (Wiley, 1985); L. E. Reichl, *A Modern Course in Statistical Mechanics* (Wiley, 1995).

[116] M. Pliscke and B. Bergerson, *Equilibrium Statistical Physics* (Prentice Hall, 1989).

[117] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Clarendon Press, Oxford, 1990).

[118] D. Raabe, *Computational Materials Science*: *The Simulation of Materials, Microstructures and Properties* (Wiley-VCH, Weinheim, 1998).

[119] W. G. Hoover, *Computational Statistical Mechanics* (Elsevier Sci. Publ., Amsterdam, 1991).

[120] K. Binder, *Monte Carlo and Molecular Dynamics Simulations in Polymer Sciences* (Oxford University Press, Oxford, 1995).

[121] N. Metropolis, A. W. Rosenbluth, A. N. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953) ; N. Metropolis, *The beginning of the Monte Carlo Method*, Los Alamos Science **12**, 125 (1987).

[122] K. Vollmayr, W. Kob, and K. Binder, Phys. Rev. B **54**, 15808 (1996).

[123] Serge Nakhmanson, *Theoretical Studies of Amorphous and Paracrystalline Silicon*, Ph.D. Thesis, Ohio University, 2001.

[124] I. Štich, R. Car, and M. Parrinello, Phys. Rev. B **44**, 11092 (1991).

[125] N. C. Cooper, C. M. Goringe, and D. R. McKenzie, Comp. Mat. Science **17**, 1 (2000).

[126] E. Kim and Y. H. Lee, Phys. Rev. B **49**, 1743 (1994).

[127] H. Urbassek and P. Klein, Phys. Stat. Sol. (b) **217**, 461 (2000).

[128] M. D. Kluge, J. R. Ray, and A. Rahman, Phys. Rev B **36**, 4234 (1987); M. D. Kluge, J. R. Ray, and A. Rahman, Phys. Rev B **37**, 4132 (1988).

[129] W. D. Luedtke and U. Landman, Phys. Rev. B **37**, 4656 (1988); W. D. Luedtke and U. Landman, Phys. Rev. B **40**, 1164 (1989).

[130] J. F. Justo, M. Z. Bazant, E. Kaxiras, V. V. Bulatov and S. Yip, Phys. Rev. B **58**, 2359 (1998).

[131] F. Wooten, K. Winer, and D. Weaire, Phys. Rev. Lett. **54**, 1392 (1985).

[132] G. T. Barkema and N. Mousseau, Phys. Rev. B **62**, 4985 (2000).

[133] R. L. C. Vink, *Computer Simulations of Amorphous Semiconductors*, Ph.D. Thesis, Utrecht University (2002).

[134] J. J. Dong and D. A. Drabold, Phys. Rev. Lett. **80**, 1928 (1998).

[135] B. R. Djordjević, M. F. Thorpe, and F. Wooten, Phys. Rev. B **52**, 5685 (1995).

[136] M. T. Dove, M. G. Tucker, S. A. Wells, and D. A. Keen, EMU Notes in Mineralogy, **4**, Chapter 4, pp. 59-82 (2002).

[137] O. Gereben and L. Pusztai, Phys. Rev. B **50**, 14 136 (1994).

[138] L. Pusztai and O. Gereben, J. Non-crys. Solids **192 & 193**, 640 (1995).

[139] J. K. Walters and R. J. Newport, Phys. Rev. B **53**, 2405 (1996).

[140] R. L. McGreevy, J. Phys.: Condens. Matter **13**, R877 (2001).

[141] R. L. McGreevy and L. Pusztai, Mol. Simul. **1**, 359 (1988).

[142] S. Gurman and R. L. McGreevy, J. Phys.: Condens. Matter **2**, 9463 (1990).

[143] D. A. Keen and R. L. McGreevy, Nature **344**, 423 (1990).

[144] M. G. Tucker, M. T. Dove, and D. A. Keen, J. Appl. Cryst. **34**, 630 (2001).

[145] F. Wooten and D. Weaire in *Solid State Physics*, Vol. **40**, edited by H. Ehrenreich and D. Turnbull (Academic Press, Inc, 1987).

[146] J. Fortner and J. S. Lanin, Phys. Rev B **39**, 5529 (1989).

[147] J. M. Holender and G.J.Morgan, J. Phys. : Condens. Matter, 3 1947 (1991).

[148] P.M. Voyles, J.E. Gerbi, M.M.J. Treacy, J.M. Gibson, J.R. Abelson, J. Non-Cryst. Solids **293**, 45 (2001).

[149] S.R. Elliott Adv. Phys. **38**, 1 (1989).

[150] N. Maley, D. Beeman, and J.S. Lannin, Phys. Rev B. **38** 10611 (1985).

[151] A.P. Sokolov and A.P. Shebanin, Sov. Phys. Semicond. **24** 720 (1990).

[152] A.A. Langford, M.L. Fleet, B.P. Nelson, W.A. Langford, and N. Maley, Phys. Rev. B. **45** 13367 (1992).

[153] S.R. Elliott, J. Phys. Cond. Matter **4**, 7661 (1992); S.R. Elliott, Nature **354**, 445 (1991).

[154] J.D. Martin, S.J. Goettler, Nathalie Frossé, and L. Iton, Nature **419**, 381 (2002).

[155] J.M. Gibson, M.M.J. Treacy, P.M. Voyles, H-C Jin, and J.R. Abelson.

[156] A.C. Wright, Glass Phys. Chem. **24**, 148 (1998).

[157] P.H. Gaskell, P.H. Eckersley, A.C. Barnes, and P. Chieux, Nature **350**, 675 (1991).

[158] M.M.J. Treacy and J.M. Gibson, Acta Cryst. A **52**, 212 (1996).

[159] J.M. Gibson, M.M.J. Treacy, and P.M. Voyles, Ultramicroscopy **83**, 169 (2000).

[160] J.M. Gibson and M.M.J. Treacy, Phys. Rev. Lett **78**, 1074

[161] M.M.J. Treacy, J.M. Gibson, and P.J. Keblinski, J. Non-Cryst. Solids **231**, 99 (1998).

[162] S.M. Nakhmanson, P.M. Voyles, N. Mousseau, G.T. Barkema, and D.A. Drabold, Phys. Rev. B **63**, 235207 (2001).

[163] S. V. Khare, S. M. Nakhmanson, P. M. Voyles, P. Keblinski, and J. R. Abelson Appl. Phys. Lett. **85**, 745 (2004).

[164] P. Biswas, D. Tafen, and D. A. Drabold, Phys Rev B **71** 054204 (2005).

[165] D. Tafen, *Topics in the Theory of Glasses*, Ph.D. Thesis, Ohio University, (2005).

[166] R.K. Dash, P.M. Voyles, J.M. Gibson, M.M.J. Treacy, and P. Keblinski, J. Phys. Cond. Matter **15**, S2425 (2003).

[167] P. M. Voyles and J. R. Abelson, *Medium-Range Order in Amorphous Silicon Measured by Fluctuation Electron Microscopy* in: *Critical Review of Amorphous and Microcrystalline*

Materials and Solar Cells, H. Fritzsche and S. Guha (eds.), *Solar Energy Materials and Solar Cells* **78**, 85 (2003).

[168] P.M. Voyles and D.A. Muller, Ultramicroscopy **93**, 147 (2002).

[169] C.S. Marians and L.W. Hobbs, J. Non-Cryst. Solids **124**, 242 (1990); L.W. Hobbs, C.E. Jesurum, V. Pulim, and B. Berger, Philos. Mag. A **78**, 679 (1998)

[170] M.M.J. Treacy, P.M. Voyles, and J.M. Gibson, J. Non-Cryst. Solids **266**, 150 (2000).

[171] M.O. O'Keeffe and B.G. Hyde, *Crystal Structures: 1. Patterns and Symmetry* (Mineralogical Society of America Monograph, Washington DC, 1996).

[172] R. Biswas, C. Z. Wang, C. T. Chan, K. M. Ho, and C. M. Soukoulis, Phys. Rev. Lett. **63**, 1491 (1989).

[173] D. C. Allan and E. J. Melle, Phys. Rev. B **31**,5565 (1985).

[174] P. A. Fedders and A. E. Carlsson, Phys. Rev. B **39**, 1134 (1989).

[175] J. Holender and G. Morgan, J. Phys.: Condens. Matter **4**, 4473 (1992).

[176] S. Kneif and W. von Neissen, Phys. Rev. B **58**, 4459 (1998).

[177] B. J. Min, Y. H. Lee, C. Z. Wang, C. T. Chan, and K. M. Ho, Phys. Rev. B **45**, 6839 (1992).

[178] P. A. Fedders and D. A. Drabold, Phys. Rev. B **47**, 13 277 (1993).

[179] D. K. Biegelsen and M. Stutzmann, Phys. Rev. B **33**, 3006 (1986).

[180] T. Umeda, S. Yamasaki, J. Isoya, and K. Tanaka, Phys. Rev. B **59** 4849 (1999).

[181] P. A. Fedders, D. A. Drabold, P. Ordejon, G. Fabricius, D. Sanchez-Portal, E. Artcho, and J. M. Soler, Phys. Rev. B **60**, 10594 (1999).

[182] M. S. Hybertsen and S. G. Louie, Phy. Rev. B **34**, 5390 (1986).

[183] B. G. Pfrommer and S. G. Louie, Phy. Rev. B **58**, 12 680 (1998).

[184] J.-L. Li, G. -M. Rignanese, E. K. Chang, X. Blase, and S. G. Louie, Phys. Rev. B **66**, 035102 (2002).

[185] P. A. Fedders, D. A. Drabold and S. Nakhmanson, Phys. Rev. B **58**, 15 624 (1998).

[186] J. Bardeen, L. N. Cooper, J. R. Schrieffer, Phys. Rev. **108**, 1175 (1957).

[187] J. C. Phillips, Phys. Rev. Lett. **72** 3863 (1994); ibid. **59** 1856 (1987).

[188] J. G. Bednorz and K. A. Mueller, Z. Phys. B **64**, 189 (1986).

[189] D. A. Drabold and Jun Li, *Amorphous and Heterogeneous Silicon Based Films*, Materials Research Society Proceedings. Vol. **715**, 2002.

[190] P. W. Anderson, Phys. Rev. **109**, 1492 (1958).

[191] H. Overhof and P. Thomas, *Electronic Transport in Disordered Semiconductors* in: P. Boolchand (ed.), *Insulating and Semiconducting Glasses* (World Scientific, Singapore 2000), pp. 553-606.

[192] Jun Li and D. A. Drabold, Phys. Rev. B **68**, 033103 (2003).

[193] M. Cobb and D. A. Drabold, Phys. Rev. B 56, 3054 (1997).

[194] R. Atta-Fynn, P. Biswas, P. Ordejón and D. A. Drabold, Phys. Rev. B **69**, 085207 (2004).

[195] S. Aljishi, J. D. Cohen, S. Jin and L. Ley, Phys. Rev. Lett. **64**, 2811 (1990)

Appendix

Related Publications

1. "Systematic study of electron localization in an amorphous semiconductor,"

   R. Atta-Fynn, P. Biswas, & D. A. Drabold, Physical Review B **69**, 085207 (2004). [Chap. 6]

2. "Electron-phonon coupling is large for localized states,"

   R. Atta-Fynn, P. Biswas, & D. A. Drabold, Physical Review B **69**, 245204 (2004). [Chap. 7]

3. "Realistic modeling of amorphous silicon from reverse monte carlo approach,"

   P. Biswas, **R. Atta-Fynn**, & D. A. Drabold, Physical Review B **69**, 195207 (2004). [Chap. 4]

4. "Inclusion of experimental information in first principles modeling of materials,"

   P. Biswas, D. N. Tafen, **R. Atta-Fynn**, & D. A. Drabold, Journal of Physics Condensed Matter **16**, S5173 (2004).

5. "Recent developments in computer modeling of amorphous materials,"

   D. A. Drabold, P. Biswas, D. Tafen, & **R. Atta-Fynn**, In Non-Crystalline Materials for Opto-electronics, Edited by G. Lucovsky and M. Popescu, INOE, Bucharest pp 441-467 (2004).

6. "Constrained Monte Carlo Approach to Modeling Disordered Materials,"

   P. Biswas, **R. Atta-Fynn**, & D. A. Drabold, Microscopy and Microanalysis **10** (Suppl 2), 804 (2004).