2008-12-02

Simulation and Analysis of Cadmium Sulfide Nanoparticles

Chad Everett Junkermeier
Brigham Young University - Provo

SIMULATION AND ANALYSIS OF CADMIUM SULFIDE NANOPARTICLES

by

Chad Junkermeier

A dissertation submitted to the faculty of

Brigham Young University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Physics and Astronomy

Brigham Young University

December 2008

BRIGHAM YOUNG UNIVERSITY

GRADUATE COMMITTEE APPROVAL

of a dissertation submitted by

Chad Junkermeier

This dissertation has been read by each member of the following graduate committee and by majority vote has been found to be satisfactory.

Date	James P. Lewis, Chair
Date	Robert Davis
Date	Brett C. Hess
Date	Matthew C. Asplund
Date	Jean-Francois Van Huele

BRIGHAM YOUNG UNIVERSITY

As chair of the candidate's graduate committee, I have read the dissertation of Chad Junkermeier in its final form and have found that (1) its format, citations, and bibliographical style are consistent and acceptable and fulfill university and department style requirements; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the graduate committee and is ready for submission to the university library.

Date

James P. Lewis
Chair, Graduate Committee

Accepted for the Department

Ross L. Spencer, Chair
Department of Physics and Astronomy

Accepted for the College

Thomas W. Sederberg, Associate Dean
College of Physical and Mathematical Sciences

ABSTRACT

SIMULATION AND ANALYSIS OF

CADMIUM SULFIDE NANOPARTICLES

Chad Junkermeier

Department of Physics and Astronomy

Doctor of Philosophy

I used *ab initio* molecular dynamics calculations to model cadmium sulfide nanoparticles. The nanoparticles were originaly spherical, bulk-like zinc-blende structures. Constant temperature molecular dynamics calculations reveals that CdS nanoparticles that are about 2 nm in diameter and have unpassivated surfaces are in an amorphous structure with short range order. The nearest neighbor distance on the surface of the nanoparticles being near the wurtzite nearest neighbor distance. I wrote the program xyzSTATS and used its results in justifying the amorphous nanoparticles claim. I also estimated the band gap of the CdS nanoparticles with unpassivated dangling bonds.

ACKNOWLEDGMENTS

Mom and Dad, this is dedicated to you. Thank you for all of the love and support that you have always given me! I love you!

I would like to thank my advisor, Dr. James P. Lewis, for his direction. I would also like to thank the members of my research group, past and present, for their help and support: Dr. Hao Wang, Dr. J. Brandon Keith, Dr. Hong Wang, Dr. DeNyago Tafen, Dr. Khorgolkhuu Odbadrakh, Jacob W. Fennick, Jason Ard, Barry Haycock, Jinling Zhou, Richard Marsh, Brooke Adams, Larry Rush, Adam Dorsey, Craig Kammerman, Kylee Underwood, Alex Baker, Caleb Tchissi, Lamar Blake, Jessica Stalnaker, Ribeka Takahashi.

I would also like to thank the full-time office staff at BYU and at WVU: Nan Elen Ah Yu, Diann Sorenson, Siobhan Byrne, Sherry Puskar, and Sandy Johns. We all know who it is that really runs the department.

I would also like to thank Dr. George Cham for creating Piled Higher and Deeper. Your work has been a major factor in keeping me sane through the insanity of graduate school.

Finally, this work was only accomplished through the Lord's blessing.

Contents

Table of Contents vii

List of Figures ix

1 Introduction 1

2 Theory 5
- 2.1 Fireball . . . 5
  - 2.1.1 Harris-Foulkes Theory . . . 5
  - 2.1.2 DOGS . . . 8
  - 2.1.3 Pseudopotentials . . . 9
  - 2.1.4 Exchange-correlation (XC) theory and pregenerated basis set . 10
- 2.2 Band Structure of Bulk CdS . . . 11
- 2.3 Quantifying Localization . . . 13
- 2.4 Bulk Probability . . . 13
- 2.5 Modified Chi-squared Statistic . . . 14

3 *Ab Initio* tight-binding analysis of CdS nanocrystals 17
- 3.1 Introduction . . . 17
- 3.2 Computional Methods . . . 20
- 3.3 Surface relaxation . . . 23
- 3.4 Molecular Orbital States . . . 25
- 3.5 Summary . . . 34

4 Simiulating CdS nanoparticles at finite temperatures 37
- 4.1 Introduction . . . 37
- 4.2 Computional Plan . . . 38
- 4.3 Results . . . 41
- 4.4 Summary . . . 49

5 FireballUI: The Fireball User Interface 53
- 5.1 Introduction . . . 53
- 5.2 User Interface . . . 54
  - 5.2.1 Menu and Tree Directory . . . 56

5.2.2 Fireball Tab . . . 57
5.3 Framework integration . . . 63
5.4 Examples of how to use FireballUI . . . 65
5.4.1 Downloading files . . . 65
5.4.2 High-throughput computing . . . 66
5.4.3 *Ab Initio* Molecular Dynamics . . . 67
5.5 Conclusion . . . 69

6 xyzSTATS: A program for performing configurational analysis 73
6.1 Introduction . . . 73
6.2 Basic Statistics . . . 76
6.3 Input Files . . . 77
6.3.1 script.inxyz . . . 77
6.3.2 comparison.inxyz . . . 81
6.3.3 neighbors.inxyz . . . 81
6.3.4 Odoyle_rules.inxyz . . . 82
6.3.5 User Defined Subsystems . . . 83
6.4 Output Files . . . 84
6.4.1 atoms_with_m_nghbrs_extended.txt . . . 86
6.4.2 rad_xi_squared.txt . . . 87
6.4.3 NN_Ayres_timestep_SUBSYSTEM.txt and NNN_Ayres_timestep_SUBSYSTEM.txt . . . 88
6.5 Example . . . 90
6.6 Conclusion . . . 91

7 Conclusion and Future Directions 95
7.1 Conclusion . . . 95
7.2 Future Direction: Passivation with Hydrogen . . . 96
7.3 Future Direction: CdS-ZnS Core-Shell Quantum Dots . . . 99
7.4 Future Direction: Electron-Phonon Coupling in CdS nanoparticles . . . 106

Bibliography 111

A xyzSTATS Output File Types 119

Index 121

List of Figures

2.1 **a and b**. Band structure of the zinc-blende structure of CdS as produced by FIREBALL. **c**. Band structure produced by Díaz *et al.* using a semi-empirical method [1]. The light hole line will resolve into two lines upon resizing the figure. 12

3.1 **a**. Comparison of the Γ-point energy eigenvalues of bulk CdS with the Γ-point energy eigenvalues of CdS with vacancy defects, computed using the Harris-Foulkes functional. **b**. Comparison of the Γ-point energy eigenvalues of bulk CdS with the Γ-point energy eigenvalues of CdS with defects, computed using the DOGS functional. 19

3.2 This flowchart shows the steps used to obtain my results; rectangles represent specific calculations or simulations performed and the ovals represent results. 23

3.3 (Color online) Average radial movement of atoms after the surface reconstruction of each CdS nanoparticle. 24

3.4 (Color online) Surface realignment. The left (right) hand column shows each nanocrystal in its ZB (relaxed) configuration. Each row has a different size of nanocrystal. The top row is the CdS264 structure with more atoms in each structure going down each column. 26

3.5 (Color online) Determination of the gap energy for each crystal in both (a) the unrelaxed ZB structure and (b) the relaxed structure. The probability to be on bulk Cd atoms, dark gray (blue online) squares, and bulk S atoms, light gray (red online) circles, are shown for each single-particle state near the gap. The TVB and BCB are indicated by dashed lines. The left frames are the result of a single point calculation using the Harris-Foulkes functional when surface S atoms are removed and there are only Cd dangling bonds (used to define the TVB energy) and the right frames are the result when Cd surface atoms are removed, leaving the S dangling bonds (used to define the BCB energy). Notice in (a) that there is a slight shifting of the TVB and BCB energies such the energy gap is increased as the number of atoms is reduced. 28

3.6 Graphs showing (a) (Color online) the bulk probability and (b) the number of accessible atoms both before and after optimization. The dashed lines specify the TVB and BCB energies of the respective systems. 35

4.1 A flow chart specifying the MD simulations and analysis performed. Atomic structures are shown as parallelograms, MD simulations are rectangles, and analysis results are given as ellipses. 39
4.2 (Color online) Radial distribution of atoms. (a) The radial distribution of the bond-centered W259 and ZB264 bulk-like structures. (b) This is a representative time-step from CdS264, which shows how the atoms that were originally in the ZB264 positions are spread out during a MD simulation. 43
4.3 (Color online) Bar plots comparing the Chi-squared fit of the MD results against W and ZB. The rows are for the MD simulations. The columns specify the region of the nanoparticles that is being compared. 44
4.4 (Color online) Bar plots comparing the Chi-squared fit of the MD results against W and ZB. The rows are for the different regions of the nanoparticles that are being compared. 46
4.5 The nearest neighbor distance (a) and the next-nearest neighbor distances for Cd-Cd bonds (b) and S-S bond (c) for each MD simulation. The vertical dashed line represents the respective bond distance in ZB while the line with a long dash followed by two short dashes represents the bond distance in W. The bold black line within the boxes gives the median value, left and right edges of the boxes show the quartiles while the wiskers represents the full range of values. In (b) and (c) W and ZB next-nearest neighbor distances overlap. 47
4.6 Radial distribution function, g(R), for the nanoparticles. The last two columns on the right represent the g(R) for W and ZB structures. . . 50

5.1 (Color Online.) Example of how FireballUI appears before any parameters are set. 55
5.2 (Color Online.) Examples of the file formats needed to run fireball.x. The file *.bas specifies the cartesian coordinates of atoms in a system, where M is the number of atoms in the system. The file *.lvs gives the lattice vectors of a periodic system. The file *.kpts gives the kpoints used in recipricol space, where N is the number of kpoints used and the weights must sum to one. 60
5.3 (Color Online.) Example PBS job script used by FireballUI to submit jobs to remote clusters and supercomputers. The word "DIRECTORY" will be replaced, by the FireballUI, with the name of the bas file associated with this calculation when the job is submitted to the remote computer. 63

5.4 (Color Online.) FireballUI searching for the Fdata, Fireball executables, and input files. 67
5.5 (Color Online.) FireballUI after performing an energy calculation on the HPA, OVDAC, and TOPO molecules. The tree shows the folders created in the TOPO directory by FireballUI. 68
5.6 (Color Online.) FireballUI setting up a MD simulation on a CdS nanocrystal. 70
5.7 (Color Online.) Snapshot from the MD simulation of the nanocrystal heating from 4.17 K to 350 K over the course of a 10 ps time frame. . 70

6.1 (Color Online.) Example of the file script.inxyz. 78
6.2 (Color Online.) Example of comparison.inxyz. 81
6.3 (Color Online.) Example of neighbors.inxyz. 82
6.4 (Color Online.) Examples of Odoyle_rules.inxyz. (a) Example of the rule needed for keeping all atoms within, and on, a sphere of radius 5 Å. (b) Example of keeping all atoms within the unit cube. 83
6.5 (Color Online.) Example of a file that defines a set of two user defined subsystems. 84
6.6 (Color Online.) Example of logfile.txt 85
6.7 (Color Online.) Example of atoms_with_m_nghbrs_extended.txt This shows the results for a CdS cluster of 29 atoms. I have shown the results for the first time-step and the first few lines of the second time-step. 87
6.8 (Color online) Example of the file structure of rad_xi_squared.txt. . . 88
6.9 (Color Online.) Example of an Ayres plot. In this plot, I have plotted the radial distribution of atoms in a nanocrystal from center of the nanocrystal when in a bulk-like zinc-blende structure. 89
6.10 (Color online) Example of the file structure of rad_xi_squared.txt. . . 91
6.11 (Color online) Plot comparing the Chi-squared fit of the MD results against W and ZB. 92

7.1 Examples of how passivation of dangling bonds decreases the probability of occupied gap states. 100
7.2 Number of accessible atoms, W. **a.** zinc-blende structures. **b.** Relaxed structures. 102
7.3 (Color Online) The probability (**a**) and number of accessible atoms (**b**) associated with each state found in CdS522, CdS894, and CdS1418. While there is not technically an interface region as in the core-shell quantum dots, I present it here for comparison with Figs. 7.4 and 7.5. 103
7.4 Probability. **a.** zinc-blende structures. **b.** Relaxed structures. 105
7.5 Number of accessible atoms, W. **a.** zinc-blende structures. **b.** Relaxed structures. 107
7.6 Electron-phonon coupling of S-centered nanocrystals. 108

7.7 Electron-phonon coupling of Cd-centered nanocrystal. Note that in the current figure there is pictures at 400 K and not one at 450 K. . . 109
7.8 Electron-phonon coupling of bond-centered nanocrystals. Note that in the current figure there is pictures at 400 K and not one at 450 K. . . 110

Chapter 1

Introduction

Semiconductor nanoparticles are small clusters, typically with a lattice structure close to the bulk lattice structure, that are intermediate between bulk semiconductors and molecules with regard to their electronic structure. As the number of atoms in the particle increases, the discrete energies of the molecular orbitals merge toward a pseudo continuum of energy levels, converging to the solid state band-structure of the bulk material. The energy difference between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) increases with decreasing number of atoms. This increasing energy as the size of the nanoparticle decreases is known as the quantum size effect [2]. Nanocrystals function as artificial atoms, due to their discrete electronic spectra, which can be precisely tailored via the quantum size effect. The use of semiconducting nanoparticles is being intensely explored for a wide range of optical, electronic, and quantum technologies that require systems with discrete spectra [3]. Nanocrystals composed of type II-VI semiconductors are especially exciting because of their relatively large band gap, causing them to absorb or emit light in the visible spectrum. It is expected that semiconductor nanoparticles will be used in novel applications for energy and lighting, disease detec-

tion, and solid state quantum computing. As the nanoparticles shrink and become smaller than the the exciton Bohr radius the novel properties of the nanoparticles become more pronounced.

CdS is a component of a variety of nanoheterostructures, including CdS-HgS-CdS quantum-dot quantum-wells which are interesting for infrared applications and CdS-ZnS core-shell quantum dots which are important for blue-green applications [1,4,5]. Theoretical results show that strain due to lattice mismatch is important in nanoheterostructures [1]. However, the additional contribution by strain effects due to surface relaxation has not been determined for these CdS nanoheterostructures. Previously, Bryant and Jaskolski investigated gap states in large passivated and unpassivated CdS nanoparticles and CdS-ZnS core-shell quantum dots using empirical tight-binding models without strain to better understand how passivation and capping layers influence the contribution of gap states to the optical response [6]. To understand how surface reconstruction and the resulting internal lattice relaxation affect these results, Ch. 3 of this work considers fully relaxed small CdS nanoparticles. This includes the atomic rearrangement of nanoparticles including the important surface reconstructions and lattice relaxation effects. The primary goal is to determine how these effects influence the quantum size effect.

In addition to understanding electronic structure, the geometric structure of CdS nanoparticles is an important factor for engineering optoelectronic devices where CdS nanoparticles are integrated with electronic platforms. Macroscopic quantities of CdS are found to be in the wurtzite (W) crystal structure. However, in nanoparticle form it can be found in either the W structure or the zinc-blende (ZB) structure, and even in the rock salt structure under high-pressure. Experimental results indicate that the crystal structure of CdS nanoparticles is size dependent. However, for sizes smaller than 6 nm, there is considerable disagreement in the experimental literature [7–15].

Theoretical results are no less ambiguous [16–22]. Most of the theoretical basis for claiming that nanoparticles are W or ZB comes from finding the ground state configuration of nanoparticles. While this may be the first step in understanding their structure, using results of this type to claim that nanoparticles of a particular size are formed in a particular crystal structure is insufficient. This insufficiency is twofold: 1. The types of calculations performed cannot find the ground state, but, at best, a local minimum. Finding local minima is problematic because the energy difference between the ZB and W structures is miniscule and it is easy to imagine cases where the local minima found give the wrong indication as to the more stable configuration. 2. Sources of enthalpy (temperature, free energy, etc.) play an important role in determining the stability of a crystal structure, but are not generally well represented in energy calculations [7]. Ch. 4 of this work focuses on long period molecular dynamics simulations of CdS nanoparticles at finite temperature. Using this type of simulation I am able provide insight into the structure of the nanoparticles without being hampered by the problems just mentioned for energy difference calculations.

The graphical user interface to FIREBALL, called FIREBALLUI, has been a collaborative effort within Dr. Lewis's research group. The purpose of FIREBALLUI is to make using FIREBALL easier, enabling inexperienced users to more quickly learn how to use FIREBALL, and to enable experienced to more quickly and easily run calculations in FIREBALL. Particular emphisis has been placed on high throughput calculation, meaning being able quickly set up and run hundreds of separate calculations. Continuing improvements to FIREBALLUI will increase its functionality, resulting in being able to run more types of calculations within FIREBALL, as well as high through put data analysis. The user interface is discussed in Ch. 5.

In computational solid state physics molecular dynamics simulations are used to

model interactions between atoms and molecules. Often, analysis of the simulation, and thus any analysis tool, is based on finding some set of parameters that are dependent on studying the simulation as a whole (i.e. diffusion coefficient, phonon modes, etc.). We often end up finding a result that acts a little different then we expected, but because of the lack of fine grain analysis of the simulation we are unable to say for certain what is causing this. xyzSTATS was written to be able to focus on any subset of the configurational data and to produce results by time-step. A description of xyzSTATS's functionality and how to use it is given in Ch. 6.

Because a large proportion of the atoms in a nanoparticle belongs to the surface, many important physical and chemical properties may be explained in terms of the surface conformation. In an unpassivated nanoparticle, states appear in the band gap; these states are due to the dangling bonds of surface atoms where there is no hybridization satisfied either through passivating agents or through surface relaxation. Though we expect the transition probability into each surface state to be small, there are many and the sum affects the quantum efficiency of the nanoparticles. Continuing studies, on the effects of passivation by Hydrogen, and with a larger band gap semiconductor (i.e. ZnS) are discussed in Ch. 7.

In summary, extensive theoretical work has been performed by others on CdS, CdSe, and ZnSe nanoparticles; including semi-empirical tight-binding methods [6,23–30]. Density functional theory (DFT) methods have been used on CdS, CdSe, and ZnSe nanoparticles as well, but these have only been for small systems ($\approx$ 200 atoms or less) [16,17,19,31–37]. *Ab initio* calculations have been employed on only relatively small systems because the computational cost is much higher than with semi-empirical tight-binding methods. Because of the computational efficiencies that are built into the FIREBALL program I was able to perform calculations with thousands of times-steps on hundreds of atoms.

Chapter 2

Theory

This chapter deals with the theory behind the calculations that were performed for this work. None of the theory is original to this text, thus, only a brief overview is given. The general outline of Sec. 2.1.1, which describes the theoretical basis of the FIREBALL code, comes from [38]. Section 2.3 describes how to determine the number of accessible atoms that a state has; this was published in [39]. Section 2.5 gives the theory behind my use of the Chi-squared statistic of two sets of binned data with a different number of data points in each set; the text for this section comes from [40].

2.1 Fireball

2.1.1 Harris-Foulkes Theory

I used the *ab initio* tight-binding code FIREBALL to perform many of the molecular dynamics and electronic structure calculations contained in this work [38, 41–43]. FIREBALL consists of using density functional theory (DFT) where the charge density is expanded in terms of local orbitals derived from a non-local pseudopotential scheme.

FIREBALL uses the Harris-Foulkes functional: [44, 45]

$$E_{tot} = E_{BS} - E_{ee}\left[\rho(\mathbf{r})\right] + E_{xc}\left[\rho(\mathbf{r})\right] - \int \rho(\mathbf{r}) V_{xc}\left[\rho(\mathbf{r})\right] d\mathbf{r} + E_{ion-ion}. \tag{2.1}$$

Where E_{tot} is the total energy, ρ is the charge density, the band structure energy E_{BS} is given by $E_{BS} = 2\sum_{n\in\text{occupied}} \epsilon_n$ where ϵ_n is an eigenvalue of the one-electron Schrödinger equation,

$$\left(-\frac{1}{2}\nabla^2 + V\left[\rho\right]\right)\psi_n(\mathbf{r}) = \epsilon_n \psi_n, \tag{2.2}$$

V being the sum of the ionic potential $v_{ion}(\mathbf{r})$ (frequently represented by a pseudopotential), a Hartree potential, and an exchange correlation potential V_{xc},

$$V\left[\rho\right] = v_{ion}(\mathbf{r}) + \frac{1}{2}\int \frac{\rho(r)}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' + V_{xc}\left[\rho(\mathbf{r})\right]. \tag{2.3}$$

The second term of Eq. (2.1) is the average electron-electron energy

$$E_{ee}\left[\rho(\mathbf{r})\right] = -\frac{1}{2}\int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \tag{2.4}$$

and $E_{ion-ion}$ is the ion-ion interaction

$$E_{ion-ion} = \frac{1}{2}\sum_{i,j} \frac{Z_i Z_j}{|\mathbf{R}_i - \mathbf{R}_j|} \tag{2.5}$$

where Z_i is the nuclear or pseudopotential charge of atom i at position $\mathbf{R}$. With the given form of the total energy, the forces acting on an atom at position R_l are determined by taking the derivative of the total energy with respect to R_l. The band-structure force is evaluated using a variation of the Hellmann-Feynman theorem [46].

The main difference between the Harris-Foulkes functional and the Kohn-Sham functional is that the former is defined in terms of an input charge density, $\rho(\mathbf{r}) = \rho_{in}(\mathbf{r})$, and the latter is defined in terms of both an input and output charge density with the two functionals being equal when ρ is self-consistently determined [47]. The Harris-Foulkes functional gives a better approximation to the ground state energy

than the Kohn-Sham functional when the density ρ is not derived self-consistently. The input density is a sum of confined spherical atomic-like densities [41],

$$\rho_{in}(\mathbf{r}) = \sum_i n_i |\phi_i(\mathbf{r} - \mathbf{R}_i)|^2, \tag{2.6}$$

where the orbitals $\phi_i(\mathbf{r} - \mathbf{R}_i)$ are the basis functions used in solving the one-electron Schrödinger equation, and the n_i are the occupation numbers of the "atomic" orbitals. When using the Harris-Foulkes functional in FIREBALL, the n_i have integer values equal to the number of electrons, in the respective shells, of the neutral atom. The wave function of each local orbital is defined to vanish at a cutoff radius (similar to an "atom-in-a-box"), r_c ($\psi^{atomic}_{fireball}|_{r \geq r_c} = 0$) and have the effect of slightly raising the electronic energy levels (ϵ_s, ϵ_p, ϵ_d, ... atomic eigenvalues) due to confinement. The r_c's are chosen to preserve the chemical trends of the atoms, i.e., the excitation of the atoms must be done in a manner that preserves the relative ionization energies and relative atomic sizes. A basis for judiciously choosing these r_c's was discussed in [48]. Cutoff radii used for the sulfur atoms were $r_c^s = 4.2\, a_B$, $r_c^p = 4.7\, a_B$, where a_B is the Bohr radius. For the cadmium atoms, the radii were $r_c^s = 5.1\, a_B$, and $r_c^p = 5.0\, a_B$, $r_c^d = 4.5\, a_B$.

The "fireball" boundary condition yields two promising features: (1) the range of hopping matrix elements between orbitals on different atoms is limited; this leads to sparse matrices simplifying the calculation of the band structure energy, (2) the slight excitation of the atoms somewhat accounts for Fermi compression in solids which gives a better representation of solid-state charge densities [49]. These orbitals are only for valence electrons. Core electrons are modeled with pseudopotentials as described below.

2.1.2 DOGS

In Harris-Foulkes theory, the input density is not determined self-consistently. Instead, the occupation numbers n_i are assumed to be equal to the neutral atom values, $n_i = n_i^0$, producing second order errors in the density [44, 46]. The Harris-Foulkes approximation is shown to work quite well for a variety of systems, especially those that are strongly covalent [48]. Tests have shown that the Harris-Foulkes functional yields total energies similar to the LDA approximation but lie below them rather than above them as in a variational Kohn-Sham calculation [50–55]. Unfortunately, the use of the neutral atom densities means that it can only be applied to systems without a significant difference in the electronegativity of their constituents.

The DOGS functional is a self-consistent field (SCF) reformulation of the Harris-Foulkes functional. The effects of charge-transfer are included by extending the Harris-Foulkes functional to second order in $\delta\rho$ using a variation in density. The variation of ρ is determined by the variation in occupation numbers n_i as may be seen in Eq. (2.6). In terms of a change $\delta\rho(\mathbf{r})$ in the reference density $\rho_0(\mathbf{r})$ (neutral atom), the Hartree energy is,

$$U = \frac{e^2}{2}\int\int \frac{(\rho_0(\mathbf{r}) + \delta\rho(\mathbf{r}))(\rho_0(\mathbf{r}) + \delta\rho(\mathbf{r}))}{|\mathbf{r} - \mathbf{r}'|}\mathrm{d}^3r\mathrm{d}^3r'. \tag{2.7}$$

The zero and first order terms are already included in the Harris-Foulkes functional. As the density is defined in Eq. (2.6) the orbitals, ϕ_i are set. Thus, for there to be a variation in the density the occupation numbers must vary, δn_i. Therefore, the functional can be reworked in terms of the occupation numbers

$$E_{tot} = E_{tot}\left[n_i\right]. \tag{2.8}$$

Finding the ground state energy is then a matter of finding the proper value for

each n_i. This is done iteratively. Here the output occupation numbers are defined by

$$n_i^{out} = 2\sum_{\alpha} |\langle\psi_\alpha|\phi_i\rangle|^2, \tag{2.9}$$

where ψ_α are the eigenfunctions of the one partile Schrödinger equation, and ϕ_i are the molecular orbital solutions to Eq. (2.1).

2.1.3 Pseudopotentials

The complicated interactions between the nucleus, the core electrons, and the valence electrons produce wave functions that are not smooth near the nucleus. Pseudopotentials are used to replace the complicated interactions with a potential that produces smooth wave functions in such a way as to simplify the calculation of the electronic structure of a given system. As approximations, pseudopotentials, are not unique and some approximations will give better results for different types of systems than other approximations [56].

Pseudopotentials are derived from the solution of the Schrödinger equation for the all-electron eigenstates of the free atom. Relativistic and core-region effects are included by incorporating averaged spin-orbit coupling, mass-velocity and Darwin terms, the Hartree potential, the exchange-correlation potential, and the electrostatic potential of the nucleus. Various parameterizations of the local-density approximation (LDA) and of the generalized gradient approximation (GGA) are available for the exchange-correlation energy [57–64].

The pseudopotential and pseudoatomic wave functions may be generated in the Hamann form [65] or in the Troullier-Martins form [66] according to the method of Fuchs and Scheffler [67]. The semilocal representation of the pseudopotentials is calculated for $l_{\rm loc} = l_{\rm max} + 1$ and $\varepsilon_{l_{\rm loc}} = \varepsilon_{l_{\rm max}}$, where $l_{\rm max}$ is the momentum of the highest occupied orbital and $\varepsilon_{l_{\rm max}}$ is its energy. Choosing $l_{\rm loc} = l_{\rm max} + 1$ avoids

unphysical "ghost" states when the pseudopotential is transformed into the fully separable form of Kleinman and Bylander [68].

2.1.4 Exchange-correlation (XC) theory and pregenerated basis set

The approach for calculating each exchange-correlation interaction facilitates storing integrals in tables in the same manner as the electrostatic integrals. As the matrix elements H_{ij} and S_{ij} beyond some cutoff radius $(r_{ci} + r_{cj})$ become exactly zero, integrals are pre-calculated on a numerical grid and needed values are interpolated from them (using a cubic spline interpolation for the two-center interactions and a two-dimensional linear interpolator for the three center interactions). These integral tables depend only on the atom type, the cut off radii, and the type of DFT exchange-correlation functional used. The integral tables are pre-computed and stored before use in Fireball, enabling the user to compute them once and use the stored integral tables ever after.

Fireball offer three levels of theory for modeling XC interactions: (1) the Sankey-Niklewski method based on the "nearly uniform density approximation", (2) the Horsfield method [69], which uses a many-center expansion based on an expansion of the density a site at a time, and (3) the Multi-Center Weighted exchange-correlation Density Approximation (MCWEDA), which is also a many center expansion combining the best features of the previous two methods while maintaining computational efficiency [43]. Horsfield method works best for molecules and clusters of atoms. To work on solids, especially close packed solids, OLSXC (McWEDA) should be used. The Sankey-Niklewski method was used in earlier versions of Fireball, it has been superseded by the OLSXC method, but is kept to enable the comparison with previous

results.

2.2 Band Structure of Bulk CdS

As the first step in establishing the validity of the results in this dissertation, I present the band structure for bulk CdS computed using FIREBALL. First, I found the lattice constant a that minimized the total energy per atom by using the Harris-Foulkes functional to compute the total energy per atom for a spread of values around the experimental value of a. Searching in the range $a = [0.57, 0.59]$ nm, this procedure gave a value of $a = 0.584$ nm which is comparable to the experimental value of $a = 0.582$ nm [70]. From the computed lattice constant I calculated the band structure of the ZB structure; see Fig. 2.1 (a). I found a direct band gap energy of 1.96 eV at the Γ-point, which is similar to the value of 2.01 eV computed by Zunger and Freeman using a linear-combination-of-atomic-orbitals technique [71]. The experimental value of the band gap is 2.50 eV from reflectivity [72] or 2.4 eV from ellipsometry [73]. This discrepancy is expected for a DFT calculation. Using a planewave basis set DFT in the local density approximation would give band gap values of about half of the experimental value. However, using localized orbitals, which have slightly excited energy eigenvalues in an incomplete basis offsets the errors inherent to DFT and increases the band gap. From the band structure I calculated the hole and electron effective masses, giving a heavy hole effective mass of 0.76 m_e and an electron effective mass of 0.14 m_e; the heavy hole falls within the range of experimental values (0.53m_e to 0.8 m_e) while the electron is approximately 26% off of experiment (0.18m_e to 0.2 m_e) [74].

Díaz *et al.* calculated the band structure of CdS using a semi-empirical tight binding calculation [1]. Comparison of the band structure computed using the FIREBALL

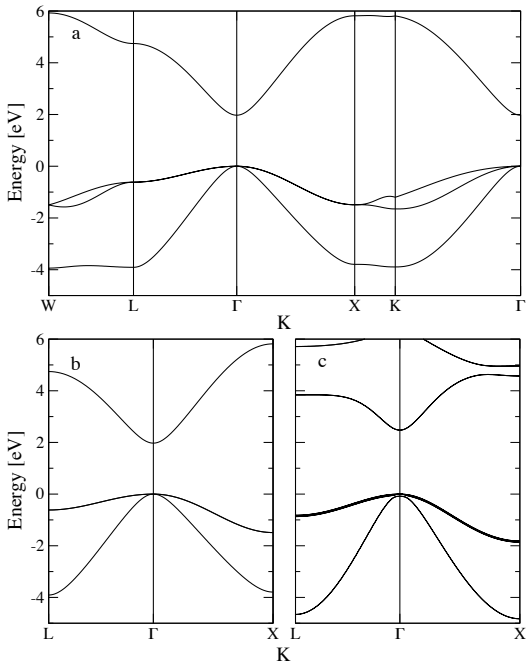


Figure 2.1 a and b. Band structure of the zinc-blende structure of CdS as produced by FIREBALL. **c**. Band structure produced by Díaz *et al.* using a semi-empirical method [1]. The light hole line will resolve into two lines upon resizing the figure.

method, Fig. 2.1 (b), with their band structure computed using the sp^3d^5 model, Fig. 2.1 (c) [75] shows that both have the same general features. The Díaz *et al.* band structure is fitted to experimental values so that it yields the correct band gap. It also takes into account electron spin, which produces a split off band (this can be observed by rescaling Fig. 2.1 (c)). The FIREBALL method is a ground state method, so the shapes of the heavy and light hole curves around the Γ point tend to be more accurate.

The results presented show that this model of CdS has validity and that nanoparticle results are meaningful.

2.3 Quantifying Localization

The number of accessible atoms W is used to quantify the degree of spatial localization [76]. The Lowdin charge, $p_i(\nu)$, is the probability that an electron in state ν can be found at atom i. Note that the sum over all atoms of $p_i(\nu)$ would equal one for each state; $\sum_i p_i(\nu) = 1$. Using $p_i(\nu)$, I can calculate the entropy-like quantity:

$$S(\nu) = -\sum_i p_i(\nu)\, ln\, p_i(\nu), \tag{2.10}$$

where the sum is defined over all atoms in the system. The number of accessible atoms $W(\nu)$ is thus defined as:

$$W(\nu) = e^{S(\nu)}. \tag{2.11}$$

This definition is independent of any definition of the nearest neighbor atoms. If the charge is spread uniformly over M atoms, then $W = M$. If the charge is localized to a single atom, then $W = 1$. I will use W to show that the dangling bond states are associated with only a few atoms.

2.4 Bulk Probability

In Ch. 3 I will discuss the "bulk probability" of a molecular orbital. It is used as a measure of how likely an electron in some state is to be found in the interior (bulk-like region) of a nanocrystal. I define a bulk atom to be any atom that has four nearest neighbors, since both Cd and S naturally have 4 neighbors in wurtzite and zinc-blende. This definition is also a handy way of specifying which atoms are on the surface, those atoms that have less than four nearest neighbors. The bulk probability BP is defined as

$$BP_\alpha = \sum_i P_{i,\alpha}, \tag{2.12}$$

where α specifies the molecular orbital, and i is a sum over all of the bulk atoms.

2.5 Modified Chi-squared Statistic

The Chi-squared statistic is a measure of the deviation of a sample from its expected value. If n samples are taken from a normal population then

$$\chi^2 = \sum_i^n \frac{(R_i - \mu)^2}{\sigma}, \tag{2.13}$$

where R_i is the result of a sample, μ is the mean value, and σ is the standard deviation.

The Chi-squared statistic can also be used to determine the goodness-of-fit for a set of binned data against a reference set of binned data. I used chi-squared to compare two sets of binned data, R and S (this can be thought of as a measure of how closely related two sets of binned data are), where the number of data elements in R is not equal to the number of data elements in S. The form of Chi-squared in this case is:

$$\chi^2 = \sum_i^n \frac{(\sqrt{S/R}\, R_i - \sqrt{R/S}\, S_i)^2}{(R_i + S_i)}, \tag{2.14}$$

where the number of bins is n, $R = \sum_i^n R_i$, and $S = \sum_i^n S_i$. When using the Chi-squared statistic below, I formed binned data out of the radial distance of the atoms from the center of the nanoparticle, this binned data will be the R_i and the binned data of either the W or ZB structure will be the S_i in Eq. (2.14). The best fit happens when $\chi^2 = 0$, and as χ^2 grows larger the fit becomes worse. In my analysis, I used χ^2 to measure the similarity of each time-step of the MD evolutions against the W or ZB structures.

The Chi-squared statistic for binned data was derived under the assumption that all of the bins have nonzero values. I will use it under conditions where for some set of bins $\{j|j \subset i\}$ the values of R_j and S_j are both zero, causing the denominator of Eq. (2.14) to be equal to zero. To get around this problem I define each term of

Chi-squared to be equal to zero at these points;

$$\chi_i^2 = \begin{cases} 0, & \text{if } R_i = 0 = S_i; \\ \frac{(\sqrt{S/R}\,R_i - \sqrt{R/S}\,S_i)^2}{(R_i+S_i)}, & \text{otherwise,} \end{cases} \tag{2.15}$$

where $\chi^2 = \sum_i^n \chi_i^2$. Thus, where as, the traditional Chi-squared statistic was defined to be used with a normal distribution that was assumed to be positive everywhere, I am allowing for distributions that are non-negative.

The Chi-squared statistic provides a means for determining if the data set can be related to the hypothesized sample. This determination is made by comparing the Chi-squared value with a pair critical values. If the Chi-squared value is between than the critical values then I accept the "null hypothesis;" in my case, that the MD simulation conforms to the reference system, W or ZB. If the Chi-squared value is greater than (or less than) both critical values then I reject the null hypothesis and accept the "alternate hypothesis" that the MD simulation does not conform to the reference system.

To find the critical value, first determine the number of bins used. Subtracting one from the number of bins used to obtain the degrees of freedom (a statistical term that is unrelated to the phase-space degrees of freedom). Once the degrees of freedom is determined, use look up tables from an introductory statistics book to obtain the critical value [77]. Tab. 2.1 contains the degrees of freedom and critical values that will be used later. The critical values in Tab. 2.1 correspond to being able to reject the null hypothesis with a 95 percent degree of certainty.

Thus, if I had a nanoparticle with 20 atoms, used 4 bins in my measure of Chi-squared, and compared it to some reference structure, I might have the values $R_1 = 5, R_2 = 5, R_3 = 5, R_4 = 5$ and $S_1 = 3, S_2 = 4, S_3 = 6, S_4 = 7$, which would result in a Chi-squared value of $\chi^2 = 1.009713$. I used four bins in my analysis, thus, I have three degrees of freedom in this example, and (from the look up table) a lower critical

value of 0.352 and a higher critical value of 7.81. Since χ^2 is between than the critical values I can say, with 95 percent surety that nanoparticle and the reference structure are the same.

Bins	DOF	$CV_{.95}$	$CV_{.05}$
4	3	0.352	7.81
22	21	11.59	32.67
23	22	12.34	33.92
24	23	13.09	35.17
25	24	13.45	36.41
26	25	14.61	37.65

Table 2.1 We give here a table containing the bins used in a binned data, the degrees of freedom "DOF" associated with that number of bins, and the corresponding critical values "CV" that are used with the Chi-squared statistic. The critical values given here represent a 5 percent error ("wrongly deciding significance" [77]).

Chapter 3

Ab Initio tight-binding analysis of CdS nanocrystals

3.1 Introduction

Nanoparticles made out of a substance that is crystalline in macroscopic quantities are often assumed to cleaved from the bulk structure when modeled. The band gap energy of the cleaved nanoparticle is often assumed to be solely dependent on its size, which is called the quantum size effect [2]. The cleaving of the substance produces many dangling bonds. Production of unpassivated dangling bonds populate the band gap region. Thus, the quantum size effect increases the size of the band gap while the dangling bonds decreases the size of the band gap. A simple experiment will show how the band gap is affected by dangling bonds.

Figure 3.1 (a) is a graph of the energy eigenvalues of bulk CdS in a periodic zinc-blende structure computed using the Harris-Foulkes functional. The red data points are for a CdS structure with 270 atoms in the supercell. The black values are the CdS supercell but with one (or two) Cd or S atom(s) removed from the ZB structure.

When one Cd atom was taken out of the structure it left the system with four S atoms that each had one dangling bond. I looked at two different configurations of when two atoms were taken out of the system. The first is when the atoms taken away were far apart from each other, this was an attempt to minimize the effect of the first missing atom on the atoms around the other missing atom. This also means that in this case there were a total of eight atoms that had one dangling bond. In the other case I picked an atom and then took away two of its neighbors, such that the picked atom had two dangling bonds, while the rest only had one dangling bond.

Furthermore, especially for CdS, many of the energy eigenvalues are degenerate, which is shown by the number of plus signs off to the right of the vertical line on which the system is placed. Much of the degeneracy is due to the point symmetries of the bulk lattice. Since the structural neighborhood, and thus the potentials that each atom feels, are the same, atoms of the same species have the same set of energy levels.

So what does this mean? I will use CdS as my reference. I find in 4S that there has been some symmetry breaking, but the states don‘t move much, and that importantly, it only slightly affects the band gap by moving the BCB state to a just barely higher energy. Similar changes are seen in 7S and 8S.

In 4Cd I find that symmetry breaking in the valence band is nearly the same as in 4S, but the broken symmetry in the conduction band is much different. The BCB state is pushed down into the band gap region, with the next six states, {BCB+a |a=1..6}, breaking their symmetries and spreading out. In 8Cd I find that the same thing happens to a greater degree. The appearance of a Cd atom with two dangling bonds pushes the BCB further into the gap region.

Figure 3.1 (b) is a graph of the gamma point energy eigenvalues of bulk CdS in a periodic zinc-blende structure computed using the DOGS functional; see Ch. 2

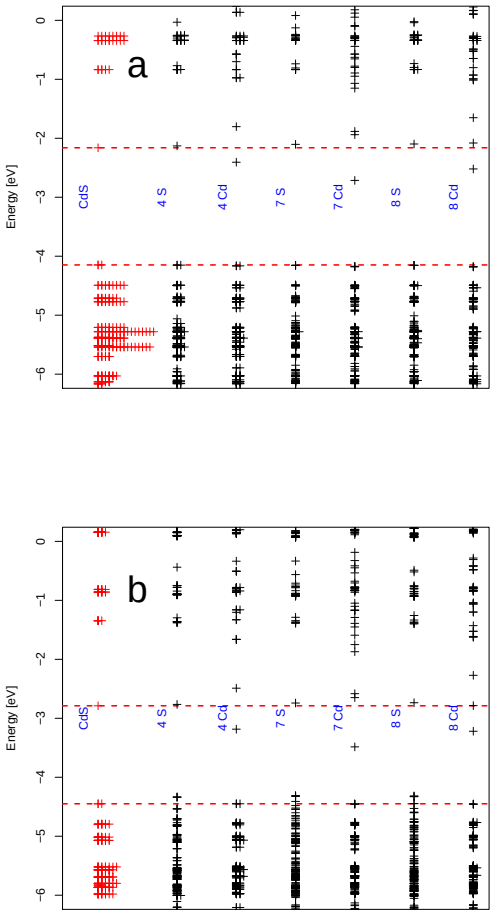


Figure 3.1 a. Comparison of the Γ-point energy eigenvalues of bulk CdS with the Γ-point energy eigenvalues of CdS with vacancy defects, computed using the Harris-Foulkes functional. **b**. Comparison of the Γ-point energy eigenvalues of bulk CdS with the Γ-point energy eigenvalues of CdS with defects, computed using the DOGS functional.

for a discussion of the DOGS functional. The formatting is the same as for the graph using the Harris-Foulkes functional. Most of the observations that were made concerning the Harris-Foulkes results also go for the DOGS results, except that when a Cd atoms is removed from the structure leaving S atoms with dangling bonds the TCB no changes a little. From this graph there appears to be very little difference between the two cases where two Cd atoms were taken out of the system, as compared to the two cases where two S atoms are taken out of the system.

Spherical Cd_nS_m nanocrystals were studied via an *ab initio* tight-binding analysis. Starting from the bulk zinc blende structure, these nanocrystals underwent relaxation as the geometries optimized to configurations that minimize internal forces. The resulting electronic-structure and molecular orbitals were analyzed and compared to the electronic-structure and molecular orbitals found in the bulk-like zinc blende structure. I conclude that the states found in the gap between the valence band and the conduction band are mainly on the surface of the nanocrystals and are due to the unpassivated bonds which remain after surface reconstruction. The text and figures of this chapter were published in Physical Review B [39].

3.2 Computional Methods

A semiconducting nanocrystal should have a "gap" in the energy of the internally confined states that is larger than the bulk band gap. These gaps are expected to vary in size with the number of atoms within the nanocrystal. In a perfect crystal the band gap would be defined as the energy difference between the valence band edge state and the conduction band edge state. In a nanocrystal with no dangling bonds, the HOMO state is the highest internally confined state and the LUMO is the next unoccupied state. Dangling bonds produce molecular orbital states that fall

around and within the band gap region, thus, the HOMO may be above the highest occupied internally confined valence band state. I will hereafter refer to the internally confined state at the top of the valence band (TVB) and the internally confined state at the bottom of the conduction band (BCB) to define the gap, instead of HOMO and LUMO, because there may be many occupied dangling bond states between the TVB and the BCB.

In the Harris-Foulkes functional electrons are not shared between atoms, therefore the molecular orbitals associated with the dangling bonds have energies near the energy gap. The transfer of electronic charge from one atom to another, in DOGS, causes the energies of the dangling bonds to shift further into the gap region. Thus, both Harris-Foulkes and DOGS obscured my finding the TVB and BCB because the dangling bonds reside around the band edges [33].

The purpose of this chapter is to understand how surface relaxation affects the band gap, to do so I had to first estimate the TVB and BCB states for each nanocrystal. To find the TVB and BCB states I took out the atoms, by species, that have dangling bonds and then performed a single point calculation using the Harris-Foulkes functional (i.e. I used the lattice structure and self-consistent potential previously found for, the full structure, remove the surface atoms of a particular species and then recalculated the states with those surface atoms removed, without iterating to self-consistency). Removing all of the sulfur atoms that have dangling bonds (leaving only cadmium dangling bonds) yielded an estimate of the TVB energy. Similarly, removing all of the surface cadmium atoms I obtained an estimate for the BCB energy. Removing the atoms by species is similar to adjusting the potentials associated with surface atoms as was done by Díaz *et al.* [1]. This produced crystal structures that were slightly smaller, with fewer atoms, and an estimated energy gap that is be slightly larger than the TVB-BCB difference of the nanocrystal. I used this procedure

on both the zinc-blende (ZB) and relaxed configurations of the nanocrystals.

Surface relaxation plays a big role in the number and type of dangling bonds. To achieve the proper surface relaxation I will first performed a quenching procedure where I used the Harris-Foulkes functional for a number of optimization time-steps followed by an optimization run using DOGS. The reason I used this two step process is that cleaving the perfect crystal produces an immediate metallization of the surface due to dangling bonds. The metallization of the surface becomes more pronounced as the charge transfer involved in the bonding process increases. A proper charge configuration can only be obtained via a self-consistent field (SCF) calculation; however, a SCF calculation on the ZB structure of the nanocrystals fails to converge. Therefore, I used a non-SCF optimization procedure, like the Harris-Foulkes functional within Fireball, to allow for surface relaxation and to remove this surface metallization. After the initial relaxation, I allowed further relaxation with SCF turned on.

To find the band gap of the unrelaxed ZB structures I removed the surface atoms, by species, and ran a single point calculation using the Harris-Foulkes functional; this procedure is similar to that proposed in Ref. [6]. We relax the nanocrystal with a 600 time-step (2.00 fs for each step) molecular dynamics simulation, using the Harris-Foulkes functional, for each nanocrystal. Starting from that relaxed configuration, I perform an additional 600 time-step SCF simulation; each time-step of which was allowed a maximum number of 20 SCF steps. In the cases of both the relaxation using the Harris-Foulkes functional and the relaxation using DOGS, the root-mean-squared (RMS) forces were less than 1.0 eV/nn. To find the band gap of the fully relaxed structures I removed the surface atoms by species, as described above, and ran the single point calculation using the Harris-Foulkes functional. A flow chart showing the procedures used in obtaining my results can be found in Fig. 3.2, and the purpose of each step will be given below.

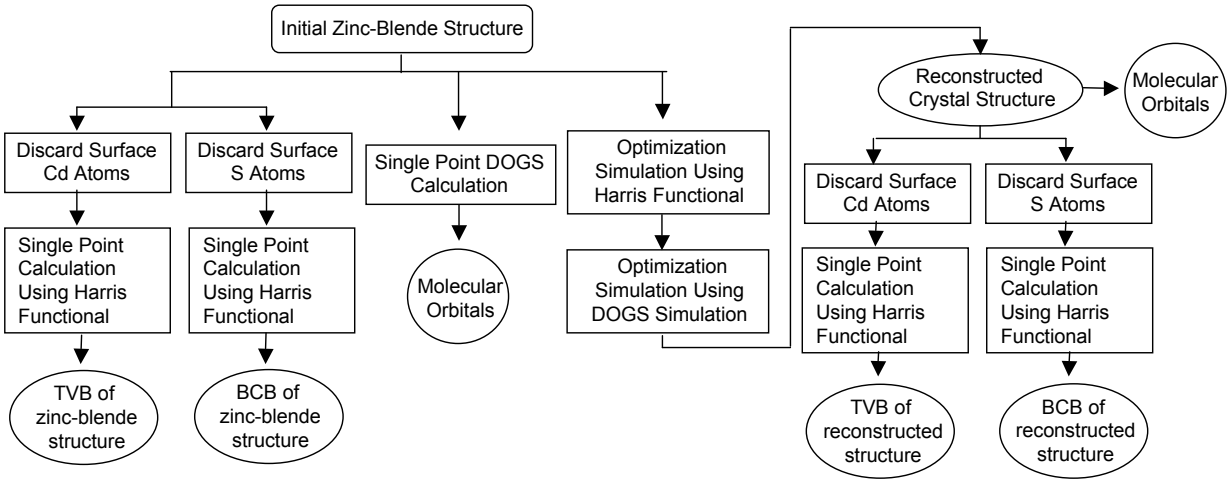


Figure 3.2 This flowchart shows the steps used to obtain my results; rectangles represent specific calculations or simulations performed and the ovals represent results.

3.3 Surface relaxation

For use in this work, six spherical CdS nanocrystals were considered, all initially in the ZB structure. These included a bond centered nanocrystal of radius $2a$, $Cd_{132}S_{132}$ (CdS264); and five atom centered nanocrystals: of radius 1.168 nm, $Cd_{140}S_{140}$ (CdS280); 1.168 nm, $Cd_{140}S_{141}$ (CdS281); 1.168 nm, $Cd_{152}S_{141}$ (CdS293); 1.2556 nm, $Cd_{152}S_{177}$ (CdS329); and 1.26728 nm, $Cd_{180}S_{177}$ (CdS357). Each of which is smaller than the Bohr radius of CdS, which is approximately 5 nm, this means that the novel quantum effects that are desirable in nanoparticles will be present in each of these models [14]. I will hereafter reference each nanocrystal by the designation given in parentheses. I chose to do the nanoparticles CdS264 and CdS281 because they have the same radius but have different numbers of atoms. The nanocrystal CdS280 was created by taking one of the S atoms off of the surface of CdS281 in order to have a second stoichiometric system. In the ZB structure a bulk atom will have four nearest neighbors; I will therefore define any atom in the unrelaxed structure with less than four near-

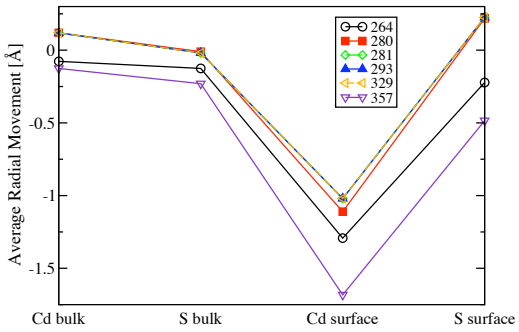


Figure 3.3 (Color online) Average radial movement of atoms after the surface reconstruction of each CdS nanoparticle.

est neighbors as a 'surface atom.' That designation is assumed to apply even after relaxation.

Fig. 3.3 shows the average radial displacement of the atoms in each of the nanocrystals after I performed the relaxation process. The interior atoms tended to move only slightly. However, on the surface, the Cd atoms moved inward on the order of 0.1 nm and the S atoms moved outward as they are displaced by the inward moving cadmium. In the cases of nanocrystal CdS264 and nanocrystal CdS357 I found that the surface S atoms moved inwards slightly, this is apparently because the Cd atoms moved in towards the center of the nanocrystal by such a large amount. In all cases the change in volume of the nanocrystal was minimal. Similar results have been found by others for CdS as well as other type II-VI systems [16, 19–21, 29, 33]. The relative displacements of the atoms were consistent with basic principles of chemical bonding. Because of the large electron affinity of the S atoms they form covalent bonds with the electrons in the filled valence shells of the Cd atoms, driving the Cd atoms on the surface to move deeper into the nanocrystal in an attempt to reclaim a filled orbital state, creating a S rich surface.

In Fig. 3.4, I present the initial positions of the atoms in a bulk-like ZB lattice

and their final positions after relaxation. The atoms near the center of the nanocrystals moved only slightly leaving them in a roughly ZB structure after the relaxation process. I also found that in the interior of the nanocrystals the relaxation process brought a Cd atom (one that had been on the surface in the ZB structure) into a position where its distance from an S atom, other than its ZB nearest neighbors, is about 10% longer than the nearest neighbor bonds in a ZB structure; thus becoming a fifth neighbor for that S atom. Finally, I do not observe the formation of S-S or Cd-Cd nearest neighbor bonds.

3.4 Molecular Orbital States

In Ref. [6], Bryant and Jaskolski used an empirical tight-binding (TB) method to study unrelaxed CdS ZB nanocrystals. In their work they considered four models for surface passivation: no passivation, only Cd dangling bonds passivated, only S dangling bonds passivated, and both Cd and S dangling bonds passivated. Passivation is modeled by shifting the energy of the dangling bonds, V_{db}, so that gap states are shifted well above (or below) the other bands, leaving only internally confined states near the band gap. By effectively passivating the dangling bonds in this manner they were able to calculate a band gap without the added complexity of adding a capping layer. By plotting the bulk probability for different levels of passivation, they showed how the band gap was influenced by the dangling bonds, as well as discuss where in the nanocrystal any gap states reside. I benchmarked my results to those of Bryant and Jaskolski by removing all surface atoms of the same species. Effectively, this removed all dangling bonds related to that species. From my results, I identified the TVB and the BCB, allowing me to determine how the surface atoms affect the band gap.

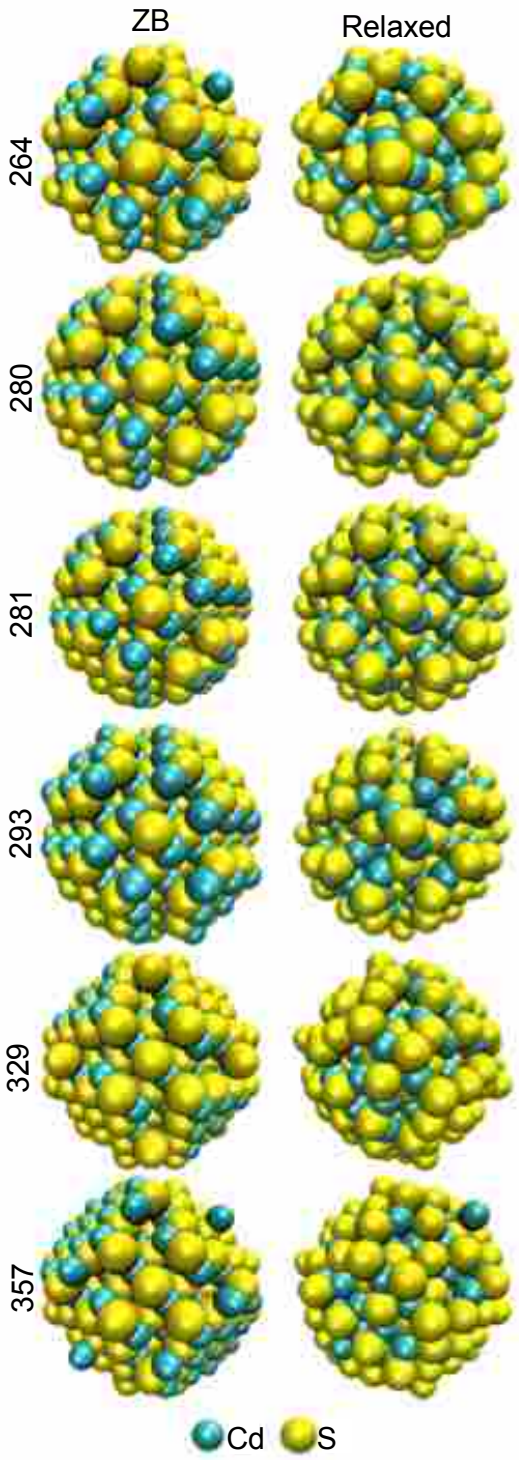


Figure 3.4 (Color online) Surface realignment. The left (right) hand column shows each nanocrystal in its ZB (relaxed) configuration. Each row has a different size of nanocrystal. The top row is the CdS264 structure with more atoms in each structure going down each column.

Before looking at the relaxed systems I will consider how removing the surface atoms by species affects the nanocrystals while in their zinc blende configuration. Doing so will allow me to test this technique on systems similar to those in Ref. [6] and later compare the results with those of the relaxed structures. My procedure differs slightly from that used by Bryant and Jaskolski. In their procedure, the surface atoms and the backbonds remain after the danling bonds are removed. In my approach, the surface becomes either Cd or S rich, with only Cd or S dangling bonds. The removal of the S derived states allows us to identify TVB, while removable of the Cd derived states, allows us to identify the BCB.

In Fig. 3.5 (a), I show the probabilities to be on bulk Cd and S atoms for each state near the gap. In the left hand panels I plot the bulk probability when the S atoms have been removed and in the right hand panels when all of the surface Cd atoms have been removed. I have results similar to those of Ref. [6]. From these results, I identified the TVB and the BCB. Tab. 3.1 presents the number of atoms remaining after the atoms were removed by species. I chose the TVB to be highest valence state with a large probability to be on bulk S atoms, just before the bulk S probability drops as the molecular orbitals become surface states. Similarly, the BCB state should be the lowest conduction-band state with a high probability to be on bulk Cd atoms. The dashed lines in the left and right hand panels pass through the TVB and BCB, respectively. In the left hand panels of Fig. 3.5 (a), states below the TVB look remarkably like valence states of [6]. In the right hand panels, the states to the right of the dashed line look like the conduction states of [6]. The TVB and BCB energies and band gaps for the ZB structures of the nanocrytal systems studied are given in Tab. 3.2.

In the ZB structure I found that the band gap generally decreases as the size of the particle increases. There is a minor exception to this when going from CdS280 to

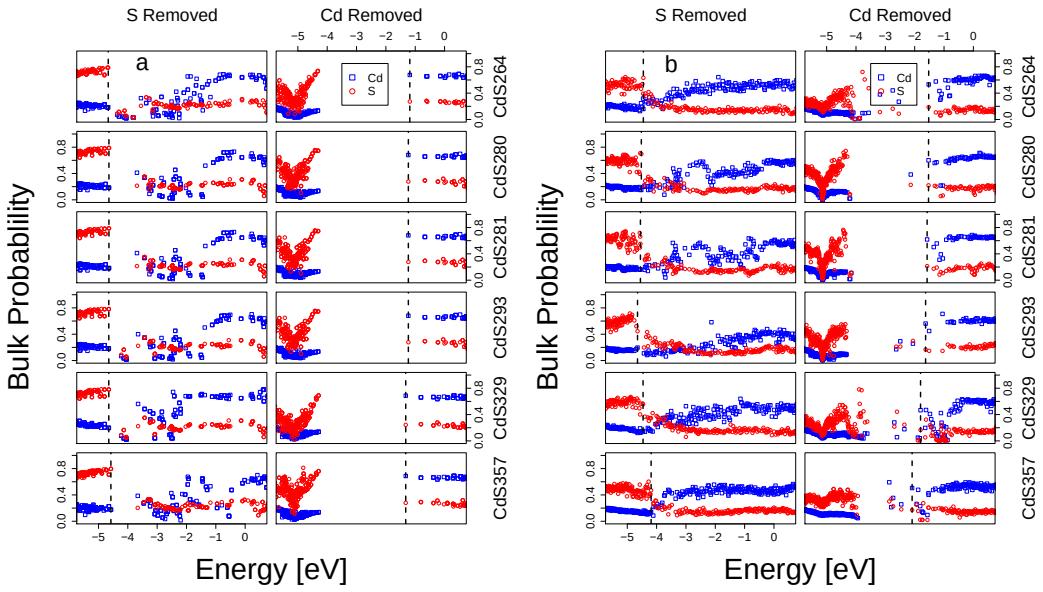


Figure 3.5 (Color online) Determination of the gap energy for each crystal in both (a) the unrelaxed ZB structure and (b) the relaxed structure. The probability to be on bulk Cd atoms, dark gray (blue online) squares, and bulk S atoms, light gray (red online) circles, are shown for each single-particle state near the gap. The TVB and BCB are indicated by dashed lines. The left frames are the result of a single point calculation using the Harris-Foulkes functional when surface S atoms are removed and there are only Cd dangling bonds (used to define the TVB energy) and the right frames are the result when Cd surface atoms are removed, leaving the S dangling bonds (used to define the BCB energy). Notice in (a) that there is a slight shifting of the TVB and BCB energies such the energy gap is increased as the number of atoms is reduced.

	Original		S Removed		Cd Removed	
System	# Cd	# S	# Cd	# S	# Cd	# S
CdS264	132	132	132	71	71	132
CdS280	140	140	140	83	80	140
CdS281	140	141	140	83	80	141
CdS293	152	141	152	83	80	141
CdS329	152	177	152	83	104	177
CdS357	180	177	180	111	104	177

Table 3.1 Number of Cd and S atoms in the resultant crystal when surface atoms are taken off the ZB structures.

CdS281 where the band gap increases by 0.04 meV, or about 0.001%, which lies within the tolerances of the present computational method and any variations expected for discrete atomistic effects. A more notable exception occurs when comparing the band gaps of CdS281 and CdS293. In this case there is a jump in the number of Cd atoms, and one would expect this to make the crystalline more metallic, and would expect this to decrease the band gap. This effect is the only instance I have seen in my analysis where the lack of stoichiometry has a noticeable effect.

Assigning the TVB and BCB states for the nanocrystals in the ZB structures is straight forward once the bulk probability is determined. The same is not true for the relaxed structures. For the relaxed structures there are more, rapid variations in the bulk probability, because some states that were degenerate in the ZB structures can become nondegenerate in the relaxed structures. In making my choice of where the TVB and BCB are in the relaxed structures I used the following guidelines: 1. The bulk probability of the Cd atoms and of the S atoms each had to follow the trend set by the unrelaxed ZB nanocrystals; for example when the surface Cd atoms

System	TVB [eV]	BCB [eV]	Gap [eV]
CdS264	-4.66644	-1.18442	3.44202
CdS280	-4.64307	-1.23413	3.40894
CdS281	-4.64307	-1.23409	3.40898
CdS293	-4.64744	-1.23409	3.41335
CdS329	-4.64744	-1.32342	3.32402
CdS357	-4.57584	-1.32342	3.25242

Table 3.2 Energy of the TVB and BCB and the band gap for nanocrystals in the ZB configuration.

were removed the value for Cd and S had to be around 0.6 and 0.3, respectively. 2. The energy of the TVB and BCB couldn't be radically different from the trends for the unrelaxed structures that are easier to judge. With these criteria in mind, I determined the TVB and BCB. The results are given in Tab. 3.3 and in Fig. 3.5 b with the number of atoms used in determining the TVB and BCB given in Tab. 3.4.

System	TVB [eV]	BCB [eV]	Gap [eV]
CdS264	-4.45219	-1.52261	2.92958
CdS280	-4.52048	-1.52794	2.99254
CdS281	-4.55150	-1.57639	2.97511
CdS293	-4.65063	-1.62789	3.02274
CdS329	-4.46160	-1.80015	2.66145
CdS357	-4.18726	-2.08743	2.09983

Table 3.3 Energy of the TVB and BCB and the band gap for nanocrystals in the relaxed configuration.

Explanation of the determination of the TVB and BCB in the relaxed structure

	Original		S Removed		Cd Removed	
System	# Cd	# S	# Cd	# S	# Cd	# S
CdS264	132	132	132	66	89	132
CdS280	140	140	140	81	90	140
CdS281	140	141	140	77	76	141
CdS293	152	141	152	77	76	141
CdS329	152	177	152	78	104	177
CdS357	180	177	180	118	142	177

Table 3.4 Number of Cd and S atoms in resultant crystal when surface atoms are taken off the relaxed structures.

cases may convince the reader of the validity of the energy values that I put forth. Since the cases where the Cd atoms are removed are simpler to justify I will concern myself with them first. Observing CdS264 and CdS281 in Fig. 3.5 (b) I see that states basically resemble the bulk probability of the bulk-like configurations, thus the lowest energy state above the gap should be the BCB. In CdS280 I would have a similar case if not for the one state at -2.13481 eV, this one state is not the BCB because it does not fit conditions 1 or 2 described in the last paragraph. Similarly, in CdS293 the three states between -3 eV and -2 eV do not fit either condition. The value determined in CdS329 was harder to come by. Here I have the bulk probability of the Cd and S atoms well separated as I found in the bulk-like case at higher energies and then as I go to slightly lower energies they both become much lower before rising again. The dip in values are due to surface states crowding into the conduction band as bonds are passivated. Again there are three states that are in the $(-3, -2)$ eV range, which I will ignore. There also happens to be two states with energies just lower than the state I have determined to be the BCB for CdS329, neither of these

are candidates because they are both surface states and thus fail condition 1. CdS357 is similar to CdS329 in that there are a few surface states that have moved into the conduction band, there are also three states in the $(-3, -2)$ eV range which fail for reasons mentioned above, the difference is that there is also a fourth state near -3 eV which has the correct form for the bulk probability but it fails condition 2.

Determining the TVB is harder due to noise from the unbounded Cd atoms throughout the whole band gap. For CdS264, CdS280, CdS281, CdS329 and CdS357 the determination of which state is the TVB is fairly easy because the last state that follows the valence band pattern set by the ZB systems also has a high S bulk probability, immediately after which the states have Cd and S bulk probability which are close together signifying a fundamental difference in the type of state. CdS293 is the only one which is questionable. For CdS293 I found the bulk probability of the three states with the highest energies before small gap in states move closer together mimicking the trend that I found in the slightly higher energy states after the gap. Ultimately, I decided that the last state is the TVB so as to most closely follow my second guideline.

While the fluctuation that I observed in the gap energy with change in crystal size are not fully understood, the literature has a number of examples of other groups seeing the same type of effect. Roy and Springborg have shown flucturations of the LUMO of up to about 0.5 eV, with flucturations of 0.1 to 0.2 eV for the HOMO in Indium Phosphide Clusters [20]. Their energy gap tops (bottoms) out at about 1 (0.3) eV for ZB and about 1.3 (0.6) eV for wurtzite. Similarly, for Cd_mSe_n results from Sarkar and Springborg, from Troparevsky et al., and from Yu et al. show that some nanocrystals of up to a few hundred atoms have gaps the are near the band gap energy of bulk CdSe [15, 19, 37]. Pal, Goswami, and Sarkar show that for Zn_mS_n the LUMO may fluctuate by almost 1 eV as the number of atoms increase [21].

Lippens and Lannoo in Fig 1 show a non-monotonic decrease in the band gap with an increasing number of atoms [74]. The band gap produced by semiempirical tight-binding calculations is generally close to the empirical band gap of a system because they change their parameters to fit it. This might explain why they get much smaller fluctuations in the enery of the LUMO state, with the energy gaps they find for CdS and ZnS staying near the experimental values obtained for the respective nanocrystals. Joswig, Springborg and Seifert have shown that the gap can fluctuate by as much as 2 eV for CdS [16]. Furthermore, as with my results, they end up with gap values that nears the bulk limit.

Joswig et al. state that they find that the BCB is localized between one and three surface atoms and thus the LUMO depends very sensitively on variations of the cluster surface [16]. They further argue that the experimental results of Lifshitz et al. support their claim [78, 79]. The fact that I found, in some cases, the energy gap between the TVB and BCB to be near my computed bulk limit for the band gap is related to the inherent difficulties of trying to determine a value that is obscured by a strong dependence on the coordination of the atoms.

Turning, now, to a comparison of the bulk-like ZB structures and the relaxed structures, I found that in each case there are fewer states within the energy gap when the crystal is in the relaxed configuration than when the same nanocrystal is in the unrelaxed ZB configuration; Tab. 3.5 gives the number of states found in the gap region for each crystallite structure before and after reconstruction.

In Fig. 3.6 (a) I present the bulk probability of the ZB structure, as found by the single-point SCF calculation, and of the relaxed structure, after the 600 time step DOGS calculation. In Fig. 3.6 (b) I give the number of atoms on which each state might be found as defined by the number of accessible atoms (Eq. (2)). For the ZB structure, most of the gap states were near the valence band edge. These gap states

Atoms	# ZB	ZB Gap [eV]	# Relax	Relax Gap [eV]
CdS264	123	3.44202	58	2.92958
CdS280	123	3.40894	77	2.99254
CdS281	122	3.40898	73	2.97511
CdS293	121	3.31335	54	3.02274
CdS329	74	3.32402	65	2.66145
CdS357	171	3.25242	44	2.18476

Table 3.5 Number of states in the energy gap for each system in the ZB structure and in the relaxed structure, along with the band gap energies of the structures.

are seen to be near the surface and most were spread over no more than 50 atoms. The gap states became increasingly bulk-like and increasingly delocalized towards the band edges. Once the crystallites are relaxed there are many fewer states in the gap and most are near the conduction-band edge. After relaxation the states near the valence band edge become even more localized and reside more fully on the surface. The gap states near the conduction band edge are delocalized and distributed nearly equally over surface and bulk atoms. The change in the number of states near the band edges along with the changes in the bulk probability and the characteristic number of accessible atoms for each state restores the semiconductor properties of the surface that was destroyed when the ZB was cleaved forming the spherical nanocrystal.

3.5 Summary

I found that the reconstruction process has broken the point symmetries of the crystal in its ZB structure, partially passivated the surface of the nanocrystal (decreasing the

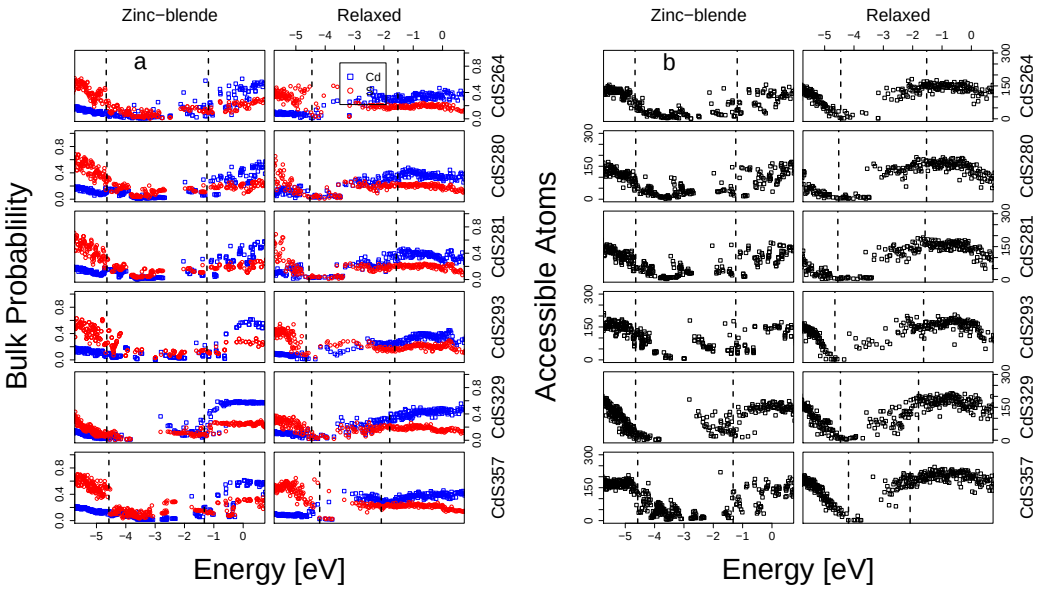


Figure 3.6 Graphs showing (a) (Color online) the bulk probability and (b) the number of accessible atoms both before and after optimization. The dashed lines specify the TVB and BCB energies of the respective systems.

number of gap states), and that the gap states are mostly surface states.

By removing surface atoms according to species I was able to give an estimation to where the top of the valence band and bottom of the conduction band should be. The estimation showed that band gap of the relaxed structures is smaller than the band gaps of the unrelaxed structures. This process also showed that the dangling bonds of the S atoms stayed mostly within the valence band, while the dangling bonds located on Cd atoms were located throughout the band gap.

Chapter 4

Simiulating CdS nanoparticles at finite temperatures

4.1 Introduction

Experimental and theoretical groups have assigned macroscopic properties to semiconducting nanoparticles, including crystal structure and thermodynamic properties [7–9, 11, 12, 14, 15, 80, 81]. While the validity of assigning macroscopic properties to nanoparticles is debatable, doing so acts as a starting point for further scientific inquiry. Macroscopic quantities of the semiconductor CdS are generally found in the wurtzite structure. Prior experimental and theoretical results confirm that the semiconductor CdS maintains a wurtzite structure for diameters greater than 6 nm. There is disagreement in the literature for sizes smaller than 6 nm. I used the DFT code Fireball and performed MD calculations on nanoparticles that are approximately 2 nm in diameter, considering different sized structures and different simulation temperatures. To predict the structure of the nanoparticles I looked at the radial distribution of the atoms about the center of the nanocrystal, the nearest

neighbor and next-nearest neighbor bond lengths, and the radial distribution function about individual atoms. Comparing the molecular dynamics simulations of relaxed nanocrystals against bulk-like wurtzite and zinc-blende nanocrystals. I predict that uncapped CdS nanoparticles, approximately 2 nm in diameter, are amorphous. This chapter has been submitted for publication to Physical Review B: Condensed Matter.

4.2 Computional Plan

Primarily, I am concerned with MD simulations at a desired sets of temperatures. A quick overview of the calculations performed is given here, along with a flow chart (Fig. 4.1) to illustrate the order of the calculations. Starting off with nanoparticles in bulk-like W and ZB structures, I optimized the nanoparticles. I then performed two sequential MD simulations on the optimized (relaxed) structures. The first to slowly increasing their temperatures to some desired value, the second fixed at the desired temperature. The fixed temperature evolutions are used to calculate characteristic properties of the nanoparticle systems.

For the starting configurations of the nanoparticles, I consider spherical crystalline structures in a zinc-blende configuration. The initial structures included a bond-centered nanoparticle of radius 1.164 nm, $Cd_{132}S_{132}$ (ZB264); an anion-centered nanoparticle of radius 1.164 nm, $Cd_{140}S_{141}$ (ZB281); and an anion-centered nanoparticle of radius 1.362 nm, $Cd_{180}S_{177}$ (ZB357) [39]. I applied a 600 time-step MD quenching procedure, at 0 K, in order to relax and partially passivate the surface. The relaxed versions of ZB264, ZB281, and ZB357 are repectively designated RZB264, RZB281, and RZB357. The total energy per atom is less after relaxation than before, showing that the relaxed configurations are energetically favorable than the bulk-like configurations.

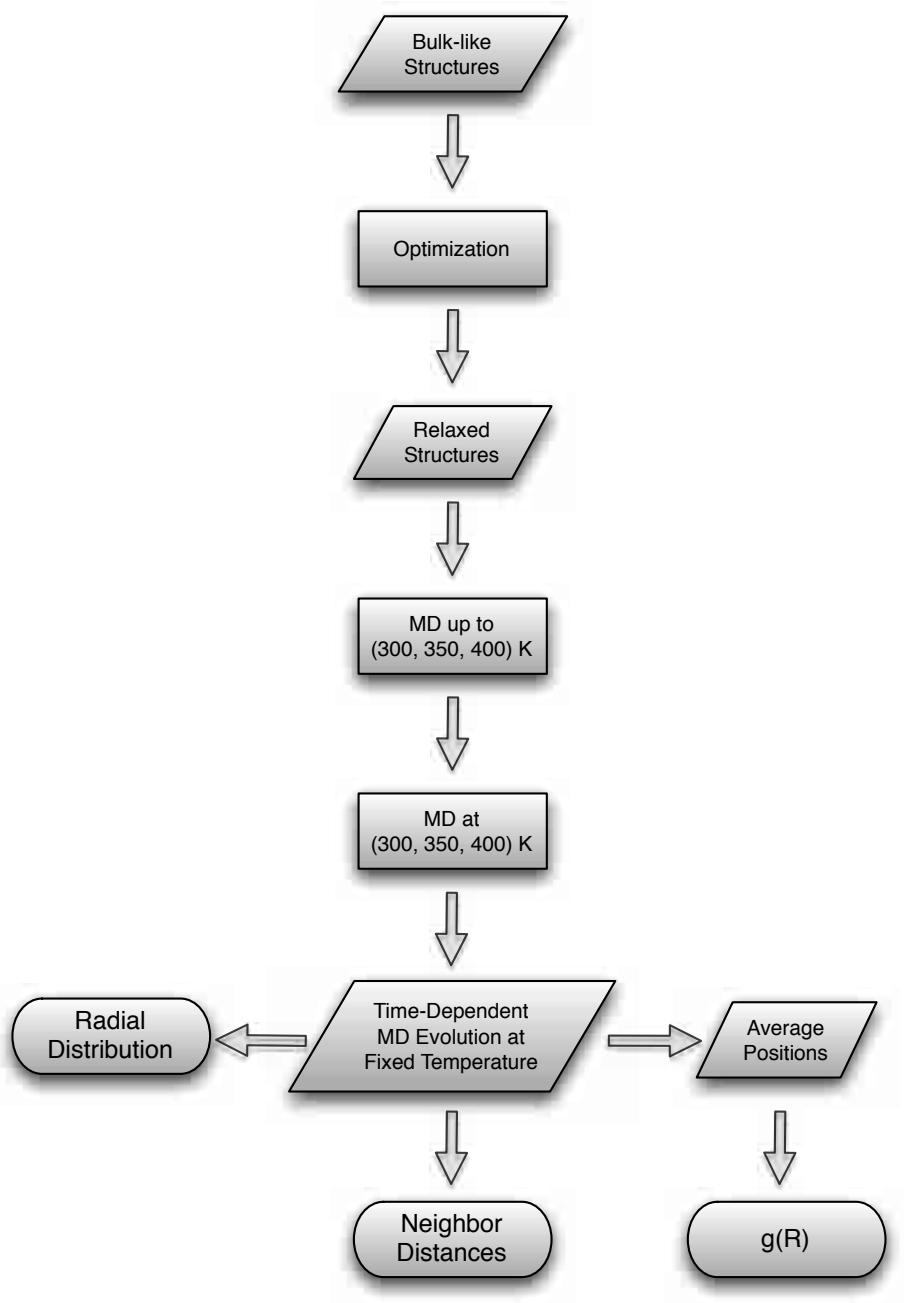


Figure 4.1 A flow chart specifying the MD simulations and analysis performed. Atomic structures are shown as parallelograms, MD simulations are rectangles, and analysis results are given as ellipses.

In a second set of simulations, I performed MD calculations to incrementally increase the temperature of the relaxed ZB nanoparticles up to 300 K. I also increased the temperature of RZB281 to 350 K and 400 K in separate calculations. I incrementally increased the temperature of the system with each time step, thus slowly changing the temperature from one time-step to the next (in the range of 0.025 K - 0.045 K per step). After each nanoparticle was raised to the desired temperature, each system underwent another MD calculation fixed at the desired temperature. I performed each calculation using a Nosé-Hoover thermostat, each simulation covering 10 ps with a 1 fs time-step. The output of each fixed temperature MD simulation is a MD evolution. When discussing the time-dependent MD evolution of a nanoparticle I will designate it by CdS264, CdS281, or CdS357. When discussing CdS281 I will also specify the temperature of the system concerned.

When performing the MD calculations I used a number-volume-temperature (NVT) ensemble with an explicit velocity Verlet integrator in which a factorization of the Liouville operator is used to propagate the ionic positions. A Nosé-Hoover chain thermostat was used to control temperature [82–85]. The Hamiltonian is replaced with a non-conservative system that adds a series of damped oscillator terms to the potential energy. This conservative system cools, or heats, the system if its instantaneous kinetic energy is higher, or lower, than k_BT, where T is the desired temperature. This integrator is reversible in time, allows long time steps, and creates ergodic motion for evaluation of thermal averages. In the FIREBALL code the default settings have four thermostats in the Nosé-Hoover chain with a characteristic frequency of each damped oscillator is set to $1.2\frac{rad}{fs}$.

We also looked at spherical wurtzite structures that were similar to the ZB structures: a bond-centered nanoparticle of radius 1.164 nm, $Cd_{124}S_{135}$ (W259); an anion-centered nanoparticle of radius 1.164 nm, $Cd_{138}S_{138}$ (W276); and an anion-centered

nanoparticle of radius 1.362 nm, $Cd_{178}S_{159}$ (W337). I ran the same MD quenching procedure as was performed on the ZB nanoparticles resulting in relaxed structures. These relaxed structures where then subjected to the incremental increase of temperature and the fixed temperature MD simulations. For brevity I will not focus on these MD simulations in this paper, but the results were simular to what I discuss below.

4.3 Results

In analyzing my results, I examined four regions of the nanoparticle, based on radial distance from the center of the nanoparticle. These regions were $r \leq 0.5$ nm, $0.5 < r \leq 0.9$ nm, $0.9 < r \leq 1.3$ nm (the region initially including the surface), and $r \leq 1.3$ nm. I did this based on the expectation that the atoms at the center of the nanoparticle would behave differently than the atoms on the surface. The region nearest the origin is slightly larger in radius than the next two regions so that the central region has enough atoms for the results to have statistical meaning. I use these regions throughout this chapter.

The first type of analysis which I consider is the radial distribution of the nanoparticle's atoms from the center of the nanoparticle. This is different from the radial distribution function, $g(R)$, which will be discussed later. The distribution data was condensed into binned data binned data for easier data analysis.

In Fig. 4.2 (a) I show the binned data of the bulk-like W and ZB structures; where I used 100 bins over $0 < r \leq 1.3$ nm. Notice that the distribution of atoms, in each region, of the W and ZB structures have a number of similarities. The similarities will play a part later as I look at the radial distribution of atoms.

In Fig. 4.2 (b) I compare the radial distribution of the atoms in the bulk-like ZB

structure with the radial distribution of the same atoms during a sample time-step from a MD simulation. Even though an atom starts in a particular region, it doesn't necessarily stay in that region. Since the atoms may not stay in the region where they started, I must allow for the possibility that atoms exchanged positions, leaving the structure of the nanoparticle unchanged, as I analyze the MD simulations. To do this I determined which atoms are in each region during each time-step of a MD simulation.

I now turn to analyzing the MD simulations. I computed the radial distribution of the atoms in the nanoparticles for each time-step in the MD simulations. Here I used 25 bins over $0 < r \leq 1.3$ nm. The 25 bins that I use here have more atoms per bin, which is desirable for the Chi-squared statistic that I use, whereas the 100 bins that I use for Fig. 4.2 shows where atoms are located. The binned data of each time-step was then compared with the binned data of both the W and ZB structures using the Chi-squared statistic. In Fig. 4.3 I show a represetative data set, CdS281 at 300 K, the critical value (given in Tab. 2.1) is given by the dashed line.

On the left-hand-side of Fig. 4.3 I have a plot of Chi-squared as a function of time, the box plots on the right-hand-side summarizes the data on the left. The thick center line of the box plot gives the median value, the top and bottom edges of the box give the quartiles, and the whiskers show the rest of the data. Each row of Fig. 4.3 gives the Chi-squared data for a different region of the nanoparticle. In Fig. 4.4, I show the summarized data for each of the systems that I have studied.

In the plots of the Chi-squared data as a function of time, the trends of the W and ZB data often have similar features because the radial distribution of the W and ZB phases have similar features. When the bulk-like W and ZB nanoparticles are centered at the midpoint of a Cd-S bond, the radial distributions for W and TB are the same (as seen in Fig. 4.2 (a)) for the region $r < 0.5$ nm, causing the Chi-squared

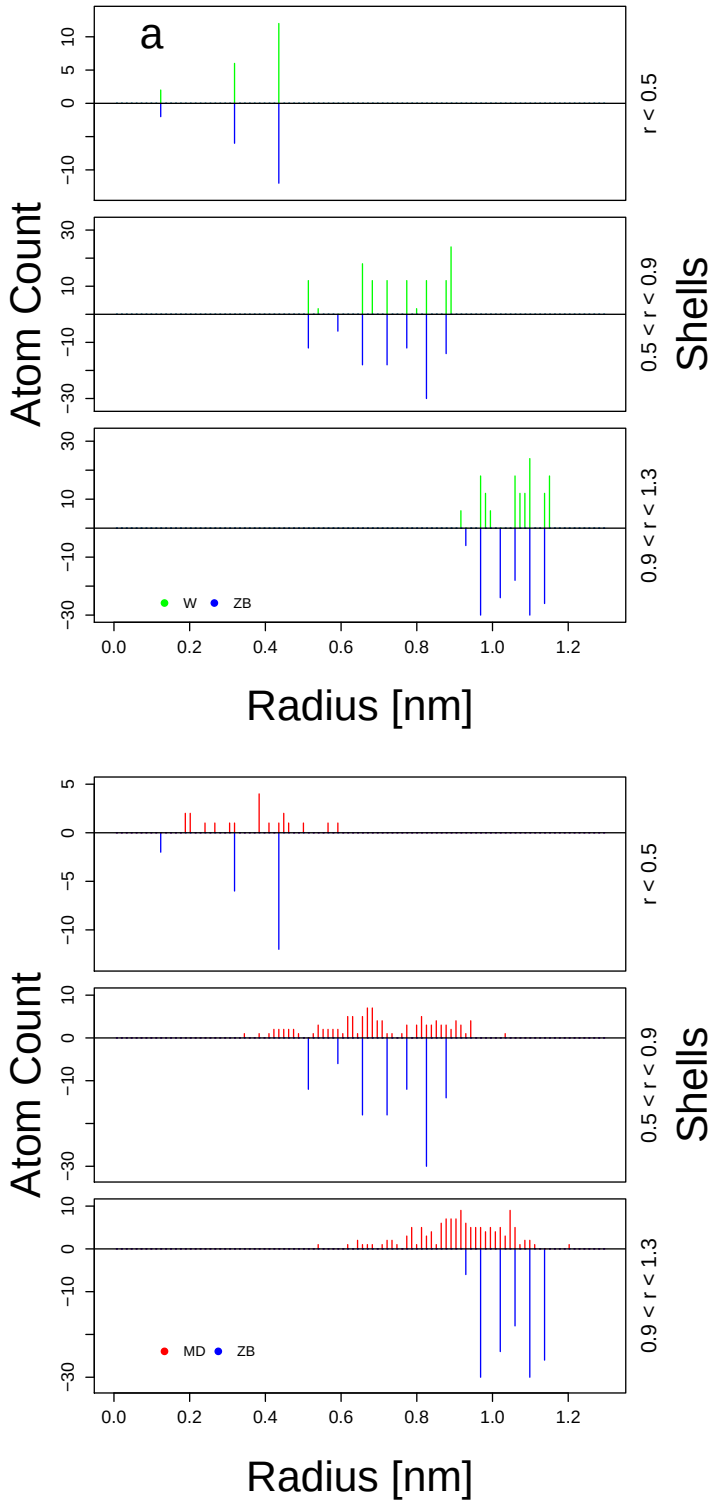


Figure 4.2 (Color online) Radial distribution of atoms. (a) The radial distribution of the bond-centered W259 and ZB264 bulk-like structures. (b) This is a representative time-step from CdS264, which shows how the atoms that were originally in the ZB264 positions are spread out during a MD simulation.

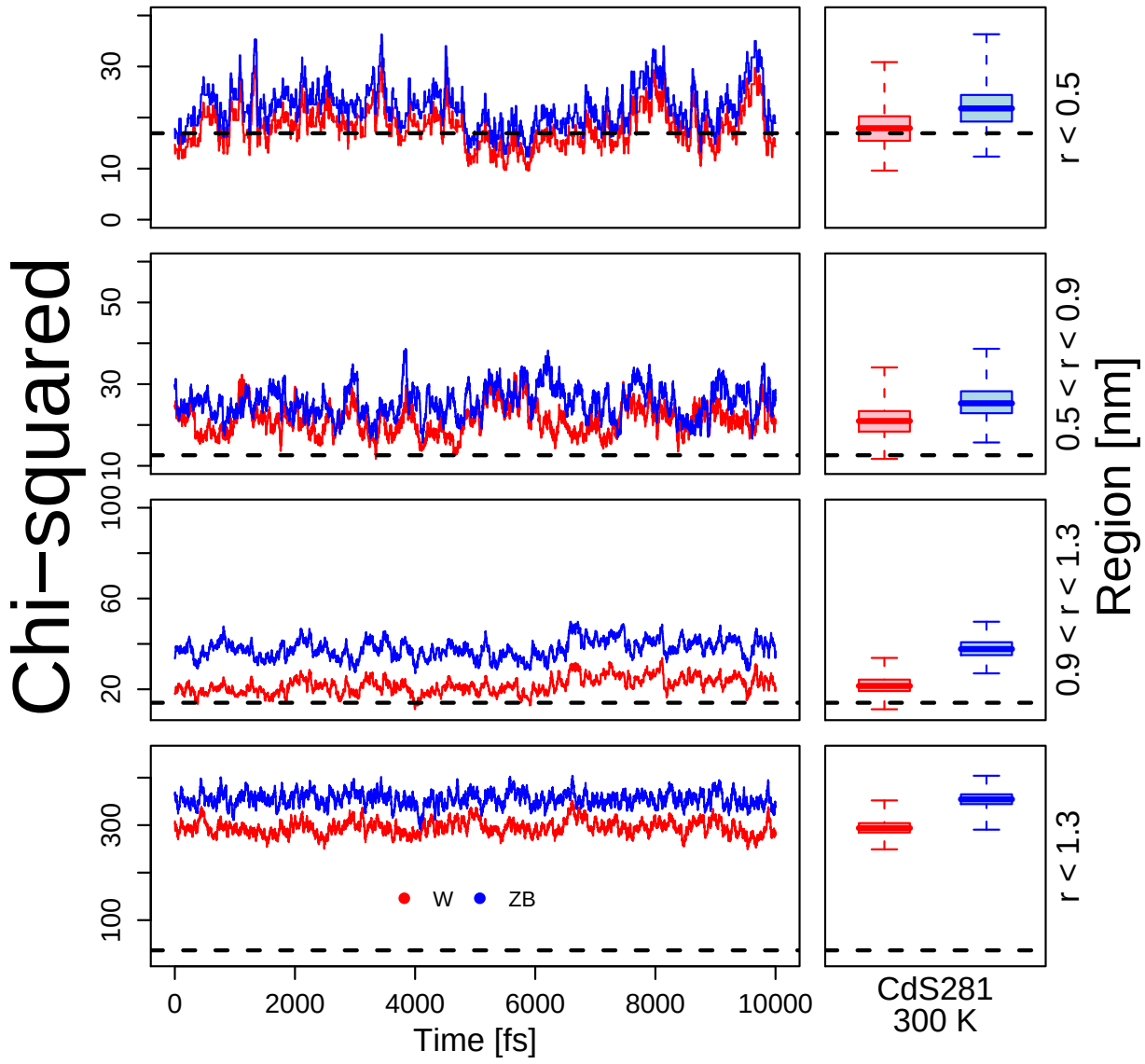


Figure 4.3 (Color online) Bar plots comparing the Chi-squared fit of the MD results against W and ZB. The rows are for the MD simulations. The columns specify the region of the nanoparticles that is being compared.

fit of W and ZB to be the same.

Figure 4.4 shows that the average Chi-squared is above the critical value in all cases. As stated earlier, if the Chi-squared value is greater than the critical value then the MD simulation does not conform to the reference system. Therefore, according to the radial distribution, the MD simulations do not correspond to either the W or ZB structure.

We computed the nearest neighbor distances by computing the distance between an atom and all of the other atoms in the nanoparticle for each time step. An atom was determined to be a nearest neighbor if it was of the other species and if it was closer than 0.27 nm from the first atom (for a maximum distance that is about 20 percent longer than the bulk nearest neighbor distance of either W or ZB).

Fig. 4.5 (a) presents a box plot representation of the nearest neighbor distances measured over the course of each MD simulation (the size of each sample set used to produce the box plot, for each region, is from 100's thousands to millions of data points). Again, the thick center line of the box plot gives the median value, the left and right edges of the box give the quartiles, and the whiskers show the rest of the data. I broke the measurements up by region. I find that the median nearest neighbor distance near the nanoparticle surface is close to the W nearest neighbor distance, while the interior regions do not resemble either W or ZB. Because it is possible to directly measure the nearest neighbor distance of the surface atoms of nanoparticles, this may be the reason why some groups have found that CdS nanoparticles of this size are in a W structure.

Fig. 4.5 (b) and (c) give the box plot representation of the Cd-Cd and S-S next-nearest neighbor distances, respectively. My analysis shows that the next-nearest neighbor distances in the MD evolutions are shorter than the corresponding nearest neighbor distances for bulk W or ZB CdS. The average Cd-Cd distance is shorter by

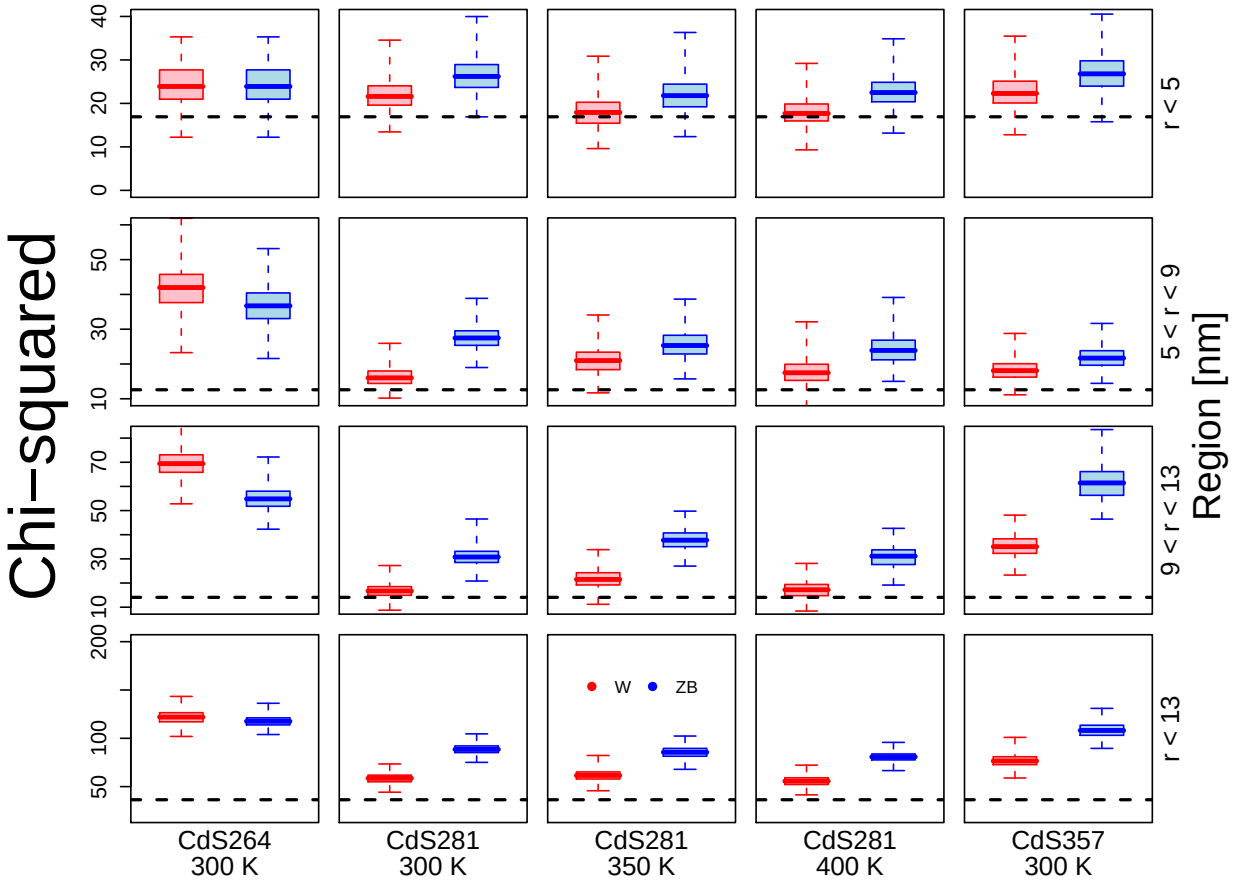


Figure 4.4 (Color online) Bar plots comparing the Chi-squared fit of the MD results against W and ZB. The rows are for the different regions of the nanoparticles that are being compared.

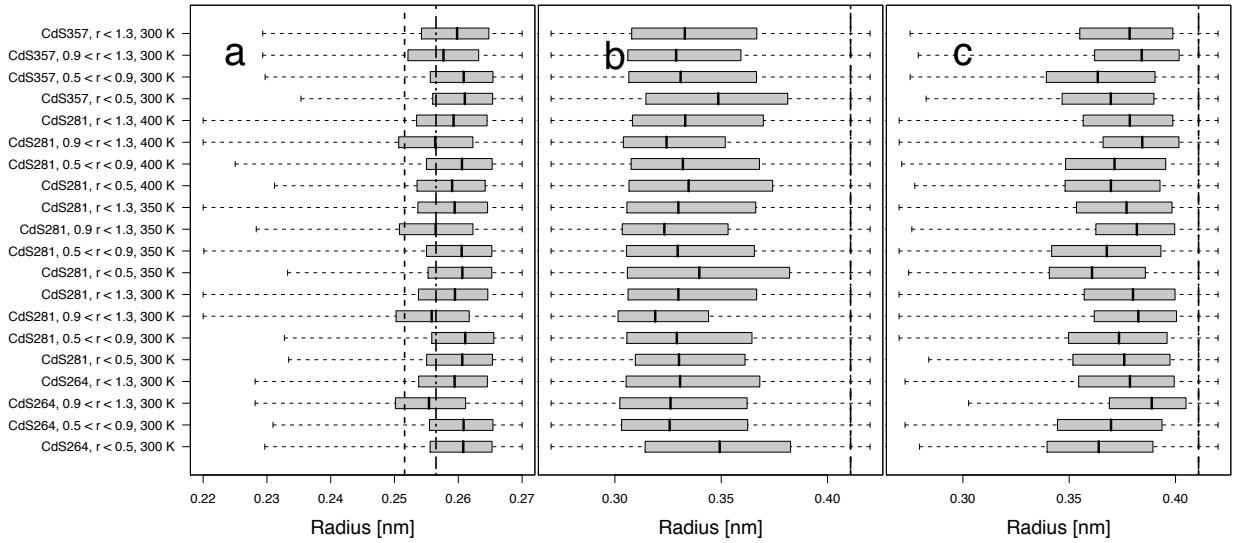


Figure 4.5 The nearest neighbor distance (a) and the next-nearest neighbor distances for Cd-Cd bonds (b) and S-S bond (c) for each MD simulation. The vertical dashed line represents the respective bond distance in ZB while the line with a long dash followed by two short dashes represents the bond distance in W. The bold black line within the boxes gives the median value, left and right edges of the boxes show the quartiles while the wiskers represents the full range of values. In (b) and (c) W and ZB next-nearest neighbor distances overlap.

about 15% to 21%. While the average S-S distance is shorter by about 5% to 12%.

The results that I have presented here appear to correspond to the results shown in Fig. 3 of Wen and Melnik [22]. In their figure, they presented a binned data of the bond lengths between 0.22 nm and 0.34 nm. In most of the cases that they presented, there is a peak around the nearest neighbor distance and another smaller peak around where I find the Cd-Cd next-nearest neighbor distance.

Earlier I discussed the radial distribution of atoms with the origin at the center of a nanoparticle. I now consider the radial distribution function, $g(R)$, which is the radial distribution of atoms around a particular atom. The radial distribution function is often used to determine if a system is crystalline or amorphous. Crystal structures have long range order. If a system is crystalline then there will be several sharp peaks representing the distances to the nearest-neighbors and the next-nearest-neighbors with additional peaks for extended-neighbors at higher radii. Amorphous structures have an absence of long range order. When the g(R) of an amorphous structure is computed, there may be any number of broadened shell-like distributions, but there are not regions where there are an absence of neighbors [86].

To compute g(R), I first computed the average position of each atom of a system for each of the MD evolutions. The g(R) was then computed on each system (CdS264, CdS281, CdS357) using the average position of the atoms from the respective evolution. If I had not computed the average positions, but instead used the values obtained by plotting the g(R) of all the time-steps in one graph, I would expect to see a more diffuse distribution of atoms then what I see for bulk-like W and ZB structures. Whereas, using the average positions, if I find a diffuse g(R), characteristic of an amorphous structure, then this is because of the structure of the nanoparticle.

To show that this line of reasoning is justified, I optimized a supercell of bulk zinc-blende CdS, linearly increased the temperature of the system using a MD simulation

and then ran a 10,000 time-step MD simulation at 300 K. I then computed the average position of the atoms in the supercell, and computed the radial distribution function. The resultant graph of the radial distribution function had sharp peaks similar to what is found in the far right column of Fig. 4.6. Thus, this method of computing the g(R) gives the expected crystal structure result.

As shown in Fig. 4.6, I find that in the $r \leq 0.5$ nm region most of the distribution is close to the nearest neighbor peak, with no gap between the nearest neighbors and next-nearest neighbors peaks. The $0.5 < r \leq 0.9$ nm region is similar to the $r \leq 0.5$ nm region in that there is no gap but there are also more neighbors on further away from the reference atom, this is because there are more atoms in this region. The $0.9 < r \leq 1.3$ nm region has a broad distribution in the outer regions but also shows sharp next-nearest neighbor peaks in the cases of CdS264 and CdS281 and 350K. This may indicate that there is short range order on the surface of the nanoparticles, which is also consistent with what I discussed about the nearest neighbor distance as found in Fig. 4.5, but is more likely due to the fewer atoms in the distribution about an atom and thus less broadening. As I see, that for the overall nanoparticle $r \leq 1.3$ nm, the radial distribution function appears to indicate that nanoparticle is in an amorphous structure.

4.4 Summary

In summary, I used *ab initio* MD simulations to model CdS nanoparticles of slightly over 2 nm in diameter. Using the radial distribution of the atoms in the nanoparticles I employed a Chi-squared statistic to show that the nanoparticles could not be in a W or ZB structure. The NN bond lengths do not resemble either W or ZB. Computing the radial distribution function, $g(R)$, I found that the nanoparticle systems are

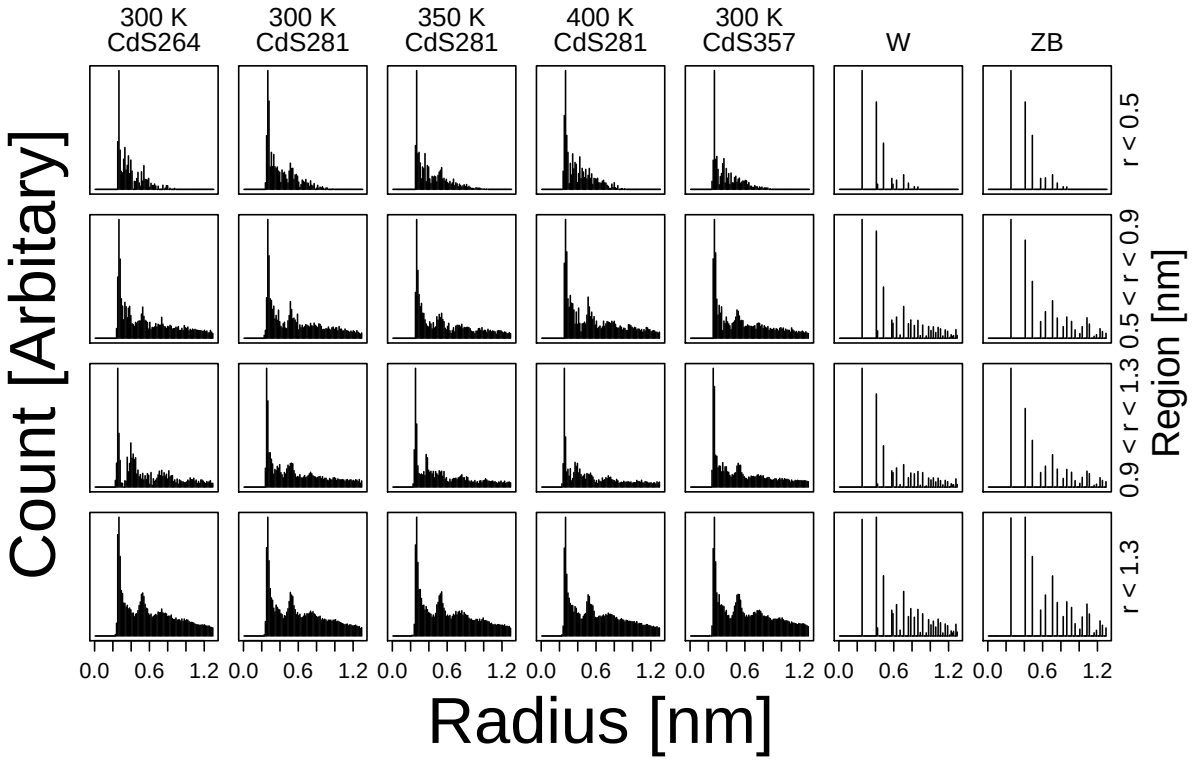


Figure 4.6 Radial distribution function, g(R), for the nanoparticles. The last two columns on the right represent the g(R) for W and ZB structures.

amorphous. Therefore, for the size of nanoparticles that I considered, I found that these nanoparticles are *not* nanocrystals but are amorphous nanoparticles that appear to have short range order on the surfaces. My results agree with the findings of Vorokh and Rempel [87]. I expect that this amorphous nature will disappear for larger nanoparticles with diameters near the 6 nm, where all of the experimental results I know of agree that the nanoparticles are in the W structure.

Chapter 5

FireballUI: The Fireball User Interface

5.1 Introduction

Quantum mechanical simulations are essential for understanding the structure, properties, and dynamics of molecules and materials. Currently, first-principles electronic structure methods, especially those based on density functional theory (DFT), are limited to systems of around 100 atoms on most workstations. To accurately calculate the electronic structure for systems of several hundred atoms, as is common in nanostructured materials, molecular devices, or large biomolecules, one is forced to use semi-empirical methods. These methods, although fast, have poor transferability when applied to new systems for which they have not been parametrized.

Tremendous progress has been made in developing rapid, parameter-free, approximate electronic structure methods completely from first principles [42, 43, 69, 88–90]. Especially useful are the so-called *ab initio* tight-binding (AITB) methods which can accurately model systems of over 1000 atoms. AITB methods typically employ ac-

curate reference electron densities, finite-range atomic orbitals, and a set of look-up tables for Hamiltonian integrals to reduce overall CPU time. Thus, they are similar to empirical tight-binding approaches but can be applied to any system and typically with better accuracy.

An early endeavor in AITB methods was the code of Sankey and Niklewski [46] which has undergone a series of modifications [41–43, 48] until its current manifestation as FIREBALL, a preeminent AITB code. The purpose of this paper is to discuss the capabilities of a FIREBALL interface, called FIREBALLUI, and the advantages it provides. The advantages that I enumerate in this paper can be easily extended to other UIs developed for other electronic-structure methods. FIREBALLUI is built so that it can be deployed via Java Web Start from a web page (http://fireball.phys.wvu.edu/fireball) or from within the CSE-Online environment [91]. In Sec. 5.2 I walk through how to execute its algorithms in FIREBALLUI. Sec. 5.3 discusses its inclusion in CSE-Online, and Sec. 5.4 illustrates using FIREBALLUI with a series of examples. Sec. 5.5 ends the paper with a brief discussion and conclusion.

5.2 User Interface

FIREBALLUI is divided into two graphical areas; see Fig. 5.1. The area on the left is a tree structure which displays workspace information. The area on the right contains tabs for the various input parameters and functionality of FIREBALL. Currently, there is only one tab, **FIREBALL**; tabs to create basis sets, data files, and perform structural and other simulation analysis will appear in later versions.

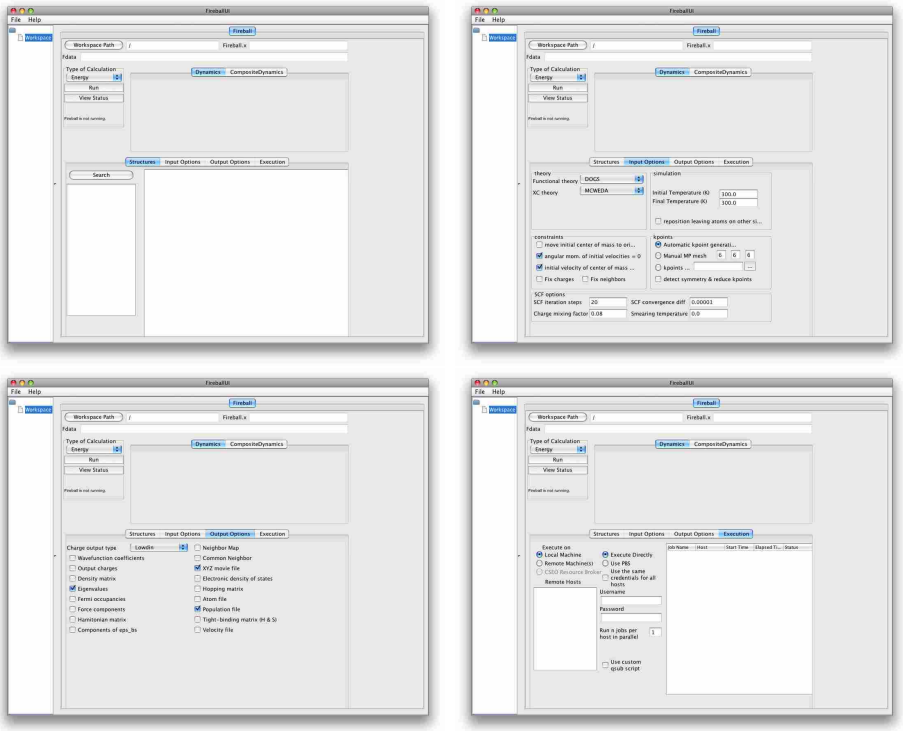


Figure 5.1 (Color Online.) Example of how FireballUI appears before any parameters are set.

5.2.1 Menu and Tree Directory

The **File** menu contains **Download Files**, **New Workspace**, **Open Configuration**, **Save Configuration**, and **Exit**. The **Download Files** menu option opens up a new window containing all of the available executables (binary versions for different computer architectures), Fdata, and psuedopotential files. A username and password is needed to download any file listed, the files that can be downloaded depend on the username and password; this protects commercial users so that any Fdata created for them can only be downloaded by them. Usernames and passwords can be obtained by requesting them through the Lewis Group website (http://fireball.phys.wvu.edu).

FIREBALLUI allows users to arrange the settings for many calculations simultaneously. Each group of settings is saved into a workspace, and users can switch between workspaces by selecting a workspace in the directory tree. To open a new workspace, click **New Workspace** in the **File** menu. Complex configurations may take some time to generate. FIREBALLUI allows configurations to be saved to disk and subsequently restored, allowing changes to be made incrementally. To save a configuration, click **Save Configuration**. All of the workspaces will be saved to disk exactly as they appear in FIREBALLUI. A saved configuration can be opened by use of the **Open Configuration**. FIREBALLUI should then appear exactly as it was when the user saved the configuration.

The **Help** menu contains **Manual** and **About**. **Manual** opens the FIREBALL wiki in the user's default web browser. Clicking on **About** opens another window which gives important compatibility information between FIREBALLUI and the Fortran version of FIREBALL. The directory tree, at the far left, lists all of the workspaces that the user has open. The user can name each workspace by clicking on the workspace in the tree and then pressing the **F2** key. This brings up a small window in which to give the workspace a name. The user can expand a workspace to

see its various files and subfolders. If the contents of the workspace have changed, pressing **F5** refreshes the tree. Double clicking a text file within the tree will open it in a new window, where it can be viewed or edited. If the file is in xyz format, the user is prompted to open it as text or to visualize it in JMol (www.jmol.org).

5.2.2 Fireball Tab

At the top of the panel are three text fields for specifying the path to a workspace, the path to a FIREBALL executable, and the path to a Fdata directory. FIREBALL's input and output files will be placed in the workspace, which can be specified by clicking on **Workspace Path**. The **Fireball.x** field specifies which FIREBALL executable that will be used.

The **Fdata** field contains the path to the pregenerated basis set to be used. Fdata datafiles are named for the species they contain, what type of pseudopotential and exchange-correlation theory they use, what orbitals are in the basis, and what cutoffs were imposed on the local orbitals. I give two examples of how Fdata datafiles are named. The first is

Fdata3t_Os3.6_p4.1_Tis6.3_p6.8_d5.7.

The number "3" indicates that the datafiles were computed using LDA psuedopotentials with the Ceperly-Perdew-Zunger exchange-correlation. The underscores separate the atom and orbital types. The nomenclature "Os3.6_p4.1" means this Fdata contains oxygen with s and p orbitals. The s orbital is cutoff at $r_c^s = 3.6\ a_B$, and the p orbital at $r_c^p = 4.1\ a_B$. Similarly,"Tis6.3_p6.8_d5.7" shows titanium with the s orbital cutoff at $r_c^s = 6.3\ a_B$, the p orbital cutoff at $r_c^p = 6.8\ a_B$, and the d orbital cutoff at $r_c^d = 5.7\ a_B$. Another Fdata example is

Fdata9_Hs4.2_9DMOL_Cs4.4_p4.8_Ns4.0_p4.5_Os3.6_p4.1_Ps4.9_p5.4.

The number "9" indicates that the datafiles were computed with GGA psuedopotentials with Becke(B88) exchange and Lee-Yang-Parr (LYP) correlation (BLYP); "9DMOL" means the atoms listed afterward use the same XC (BLYP) but have a double numerical basis [92] instead of a single numerical basis set. The nomenclature "Hs4.2" is for hydrogen with an s orbital cutoff of $r_c^s = 4.2\ a_B$. The string "Cs4.4_p4.8" shows carbon with cutoffs of $r_c^s = 4.4\ a_B$ and $r_c^p = 4.8\ a_B$, and "Ns4.0_p4.5_Os3.6_p4.1_Ps4.9_p5.4" nitrogen with cutoffs of $r_c^s = 4.0\ a_B$ and $r_c^p = 4.5\ a_B$, oxygen with cutoffs of $r_c^s = 3.6\ a_B$ and $r_c^p = 4.1\ a_B$, and phosphorous with cutoffs of $r_c^s = 4.9\ a_B$ and $r_c^p = 5.4\ a_B$. Fdata datafiles are created by the CREATE and BEGIN modules which are part of FIREBALL. The functionality to run these modules will be included in the next version of FIREBALLUI.

Immediately below the **Workspace Path** button and the **Fdata** field is the **Type of Calculation** panel. The top element of this panel is a drop-down menu which allows users to perform a variety of calculations such as energy calculations (**Energy** in the drop down menu), optimizations (**Optimization**), and molecular dynamics (**MD**). Additional types of calculations will be available in the next version of FIREBALLUI. Also located on this panel are the **Run** and **View Status** buttons, and text stating if FIREBALL is running.

The **Dynamics** panel changes depending on the type of calculation chosen. If an **Energy** calculation is chosen then the dynamics panel remains blank. If **Optimization** is chosen, a new subpanel appears where the user selects **Slow quench** [93], **Fast quench** [93], **Conjugate gradient** [93–95], or **Specified quench** [93]. Choosing one of these produces another options panel on the right containing the maximum number of time steps, the timestep size (**delta t**), and force/energy tolerances. If the conjugate gradient method is chosen, further options are given such as the maximum atomic displacement (**Max displacement**), the maximum number of iterations used

in the conjugate gradient approximation (**Max internal interations**), or the **Refine with quenching** toggle. The **Refine with quenching** option performs quenching when the conjugate gradient places the system closer to the local minimum.

The MD suboptions appearing in the dynamics panel include an integrator choice (**predictor-corrector** vs. **velocity-verlet**), ensemble choice (**NVE** vs. **NVT**), thermostat choice, length of the MD run, etc. There are two choices for the thermostat, Nosé-Hoover Chain dynamics (**NHC**) [82, 83] and a simple isokinetic velocity/force rescaling. Below the **Type of Calculation** and **Dynamics** panel there are other panels titled **Structures**, **Input Options**, **Output Options**, and **Execution**. I discuss each of these areas in more detail.

The **Structures** panel, Fig. 5.1, is used for locating atomic coordinate files (ending ".bas") and lattice vector files (ending in ".lvs"). As explained in depth in the user's manual [93], a .bas file has the number of atoms in the system on the first line and each atom listed on a separate line as atomic number and cartesian coordinates (in Å), see Fig. 5.2. A lvs file contains the three lattice vectors (in Å), each on a separate line, see Fig. 5.2. The **Search** button opens a popup window which allows the user to search for all FIREBALL executables (fireball.x), Fdata datafiles, .bas files, and .lvs files. The directory to be searched can be changed by typing in the textbox or double clicking the textbox, which will open a file chooser. All subdirectories of this directory will also be searched. At any time during the search, users can specify the executable and Fdata by selecting the appropriate items and clicking the **Add** button. FIREBALLUI can also locate xyz files and automatically convert them to .bas format. Users may also search for custom file patterns. For example, to search for all files with the extension "kpts", enter the regular expression ".*kpts" into the rightmost textbox and click **Add Search Pattern** button. The regular expression is compared against the entire path of the file, so most expressions should start with

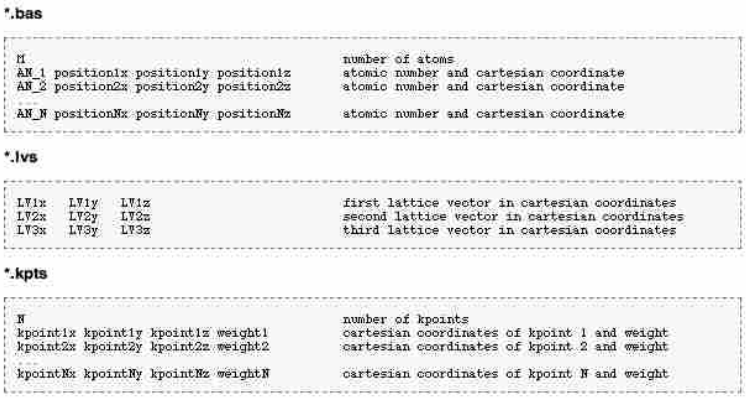


Figure 5.2 (Color Online.) Examples of the file formats needed to run fireball.x. The file *.bas specifies the cartesian coordinates of atoms in a system, where M is the number of atoms in the system. The file *.lvs gives the lattice vectors of a periodic system. The file *.kpts gives the kpoints used in recipricol space, where N is the number of kpoints used and the weights must sum to one.

".*" to ignore path.

Upon selecting a .bas file in the window below the search button the window to the right will display a drop down menu to specify if the system is **Periodic** or non-periodic (**molecule**). If the system is chosen to be periodic then a panel appears where either the lattice parameters or lattice vectors may be input. Lattice parameters will be automatically converted to lattice vectors. The lattice vectors can also be input by selecting a .bas file and then using the search function to find its corresponding lvs file.

The **Input Options** tab, Fig. 5.1, is for specifying options common to all run types. In **theory** the user can choose the type of density functional (**Harris**, **DOGS**, **extended Hubbard**) and the type of exchange correlation treatment (**MCWEDA**, **Horsfield**, **SN**).

Under **simulation** the user specifies simulation parameters such as the beginning and ending temperatures. There is also a checkbox, **reposition leaving atoms on**

other side, used to allow atoms in a periodic system to leave the unit cell from one side and return on the other side of the cell. Under **constraints** are a variety of geometrical and energetic constraints. There are checkboxes to move the system so that its center of mass is at the origin (**move initial center of mass to origin**), to fix the center of mass so that there is no translation (**initial velocity of center of mass = 0**), etc. Each of these will be applied to only the first time step of a calculation. The **Fix charges** textbox prevents wave function optimization, which is useful in band structure calculations where multiple kpoints must be evaluated at the same charge density.

For periodic systems FIREBALL uses Bloch's theorem to calculate electronic structure. It can integrate the Brillouin zone using an automatically generated Monkhorst-Pack grid [96] (**Automatic kpoint generation**) or a mesh specified by the user (**Manual MP mesh**). **Fireball** can also take custom kpoints (**kpoints file**) placed into a plain text file in the format, where the first line specifies the total number of kpoints and each of the following lines specifies a kpoint in cartesian space followed by a weighting factor for that kpoint. The weighting factors must add to one. An example of how the file should appear is found in Fig. 5.2. FIREBALL can also automatically detect the system symmetry and appropriately reduce the number of kpoints (**detect symmetry & reduce kpoints**). Various self-consistency options may also be specified in the SCF options section, such as maximum number of iterations (**SCF iteration steps**), the charge mixing factor between new and old electronic structures (**Charge mixing factor**) using the Anderson algorithm for mixing, SCF convergence tolerance (**SCF convergence diff**), and smearing temperature (**Smearing temperature**) [97, 98]. The default values will work in most situations.

The user can choose to write out a large number of potentially useful results by

marking the checkboxes under the **Output Options** tab, Fig. 5.1. A few useful quantities that can be output are atomic charges (**Löwdin** or **Mulliken**), energy eigenvalues of the molecular orbitals (**Eigenvalues**), xyz trajectories (**XYZ movie file**), or orbital populations (**Population file**).

The **Execution** tab, Fig. 5.1, is used for specifying the the computer(s) where FIREBALL will be ran. The default setting is to run on the local machine. If the user desires to run one or more **Remote Machines** then they must double click on the **Remote Hosts** field and type the computer name(s), or IP address(es), into the pop-up window. When executing remotely, workspace directory, executable (fireball.x), and Fdata datafiles are in reference to the remote machine. A unique username and password may be entered for each remote machine. If all the usernames and passwords are identical, users should select the **Use the same credentials for all hosts** checkbox. **Use PBS** allows for using clusters and supercomputers that have the Portable Batch System (PBS) job scheduling system. A simple job script is used by default, but a computer/calculation specific script can be used by selecting the **Use custom qsub script** checkbox. The default script will simply run FIREBALL on a single processor from each job's directory. If a custom script is used the user must write the keyword "DIRECTORY" at the end of the path where the job should be written on the remote machine. The name of each structure will replace "DIRECTORY"; see Fig. 5.3. When using non-PBS machines, the user has the option to tell the GUI how many jobs it should run on a computer at the same time by use of the **Run n jobs per host in parallel** text box. Once all jobs have been set up and the **Run** button has been clicked, the user can watch the progress of the jobs being run locally or submitted to other machines in the field on the right. If using PBS, there is currently no way to determine status, so FIREBALLUI simply displays "done" as soon as jobs are submitted to the queue.

```
#!/bin/bash
#PBS -q long
#PBS -l nodes=2:ppn=4
#PBS -m ae
#PBS -M fjkl@mail.wvu.edu
#PBS -N jobDIRECTORY

source ~/.bashrc


cd /home/fjkl/DIRECTORY

/usr/local/common/mpich/bin/mpirun \
       -machinefile $PBS_NODEFILE \
       -np 8\
        /fireball.x
```

Figure 5.3 (Color Online.) Example PBS job script used by FireballUI to submit jobs to remote clusters and supercomputers. The word "DIRECTORY" will be replaced, by the FireballUI, with the name of the bas file associated with this calculation when the job is submitted to the remote computer.

5.3 Framework integration

As mentioned previously, FireballUI can be run either in stand-alone mode or within a larger framework. Various frameworks currently exist where AITB capabilities would be useful, such as DANSE and CSE-Online. DANSE is a distrbuted computing effort to complement neutron scattering experiments at the Spallation Neutron Source (neutrons.ornl.gov). DANSE uses primarily the Python language to implement the algorithms of neutron scattering data reduction. It also includes bindings to standard simulation codes to create scattering kernels for Monte Carlo neutron scatting simulation of these materials. FireballUI has been used within DANSE in two ways. Initially it's backend interface to Fireball was wrapped with Python bindings using the Babel technology (https://computation.llnl.gov/casc/components/babel.html). A scientific interface definition language (SIDL) is used by Babel to pass scientific-oriented data objects such as large-scale arrays between disparate languages quickly and easily. The Babel tool suite consists of a SIDL parser, a code generator, a small

run-time support library, and the Alexandria component repository. Thus DANSE's Python framework interacts with the FireballUI Java application via Babel's SIDL bindings.

However, recently DANSE has moved to a more web-services based interface in which JavaScript, Java applets, and Java web start play a significant role. This allows a user to launch FireballUI directly from within the DANSE portal. Fireball's output is then parsed by other DANSE components and used in scattering kernel construction. For example, electronic structure energies can be applied to lattice dynamics models of quasiharmonic solids from which inelastic neutron scattering intensities are calculated. Likewise *ab initio* MD trajectories may be used to calculate quasielastic/inelastic neutron scattering intensities of arbitrary solid and liquid materials. These scattering intensities can then be utilized during Monte Carlo neutron instrument simulation to generate instrument-specific scattering spectra. The value of the Fireball method to neutron scattering is that spectra of new types of materials, especially nanomatrials involving 1000's of atoms, may be simulated with *ab initio* accuracy. In general, FireballUI integration in DANSE depends on these accompanying scattering spectra components which, in their totality, are due for completion by 2011 (http://danse.us).

CSE-Online is a cyberenvironment which allows access of various applications, tools, data, and computing resources from one desktop [91]. It is platform independent and requires no additional software beyond the Java Runtime Environment. CSE-Online is extremely versatile and has many application areas such as catalysis, quantum and molecular simulation, bioinformatics, and crystallography. Beyond regular CSE-Online applications, third party developers can use the application programming interface to host their applications on their own servers and integrate them into the framework. This streamlines user effort by automating file conversion and

transfer and interprocess communication. All applications are automatically updated with bug fixes and improvements; software is upgraded to new versions as soon as they are deployed. To access FireballUI within CSE-Online, simply check it out of the catalog and click on its desktop icon. Workflow examples that are enabled by CSE-Online include (1) searching for crystal structures in public databases and exporting them to Fireball, (2) performing *ab initio* calculations in Fireball and exporting to kinetic rate constant calculators, (3) exporting Fireball energies to forcefield fitting programs and performing classical MD, etc.

5.4 Examples of how to use FireballUI

5.4.1 Downloading files

To initially use FireballUI one must download the underlying Fortran binary. Select **Download Files** in the **File** menu. After a username and password are provided, a list of fireball executables and Fdata appear. Supposing I are going to do the examples in this paper, I select

Fdata3_Ss4.2p4.7_Cdd5.1s4.5_Zns5.0_d4.5_Ses4.4_p4.9.7z,

Fdata9_Hs3.8_9DMOL_Cs4.0_p4.5_Ns3.7_p4.1_Os3.4_p3.8_Ps4.6_p5.0.7z,

Fireball-bin.7z,

and press download. Once a directory is chosen ("fireballDemo" in this case), downloading begins. All files are compressed with 7zip to save space. The files are automatically extracted when downloading is complete, and the original archives are deleted. Most files are quite large and should be downloaded over a high-bandwidth connection.

5.4.2 High-throughput computing

As a first example of the power of FireballUI for efficiently performing electronic structure calculations, I demonstrate its high-throughput feature. This eases the workload when a user needs to perform many calculations which have the same configuration, but differ in the atomic structure. To demonstarte this I perform an **Energy** calculation to obtain the molecular orbital energies of three molecules; they are trioctylphosphine oxide (TOPO), Hexylphosphinic acid (HPA), and octadecyl-p-vinylbenzyldimethylammonium chloride (OVDAC). The atomic coordinates are given on the release page http://fireball.phys.wvu.edu/fireball under the heading "Test data". Clicking on "High-throughput Fireball" under this heading, saving to "fireballDemo", and unziping gives the "highThroughput" subdirectory with the three bas files "HPA.bas", "OVDAC.bas", and "TOPO.bas". Clicking on **Workspace Path**, I select "highThroughput" as my workspace and FireballUI discovers the three bas files and loads them into the **Structures** tab.

To specify the Fireball executable, the Fdata files, and the bas files, I click **Search**, set the path to "fireballDemo" or any directory above it, and click the inner **Search** button. It finds the binary "fireball_mac_intel.x" (the version of fireball.x for Intel processor based Apple Mac's running OS X), the two Fdata I downloaded (under the **Fdata** tab), and the three previous bas files. I select fireball_mac_intel.x, Fdata9_Hs3.8_9DMOL_Cs4.0_p4.5_Ns3.7_p4.1_Os3.4_p3.8_Ps4.6_p5.0.7z, the three previous bas files, and click **Add** for each (see Fig. 5.4). After closing the **Search** window I note my choices are placed in their respective text boxes, with all bas files listed in the textbox under the **Search** button. Clicking on each of these bas files allows us to specify their system (**Molecule** or **Periodic**) and even to visualize them in JMol.

We accept defaults under **Functional Theory**and **kpoints** but change **XC Theory** to **Horsfield** since this particular basis set was created with just Horsefield XC

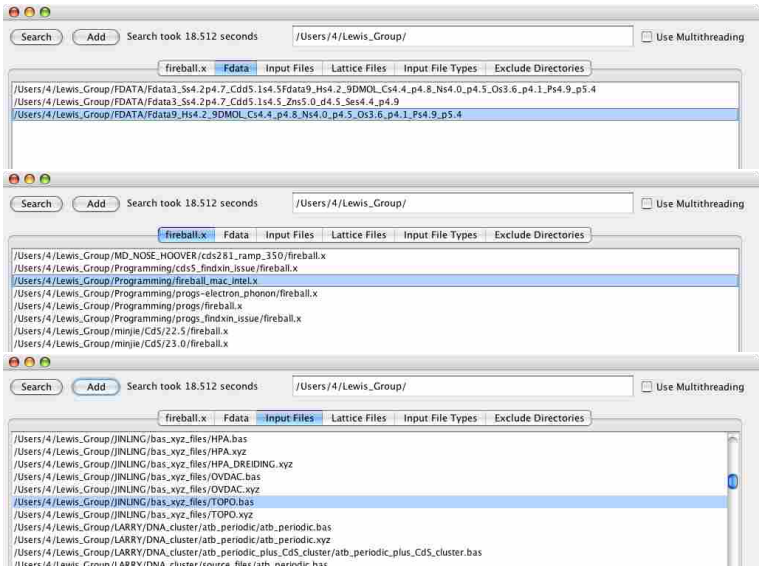


Figure 5.4 (Color Online.) FIREBALLUI searching for the Fdata, FIREBALL executables, and input files.

theory. I accept all other defaults in this and the **Output Options** tab. Since this is a simple calculation I accept **Local Machine** on the **Execute** tab and click **Run**. In this case the self-consistent electronic structures of all three molecules are calculated sequentially and three new folders are created in the current workspace; each folder is named after the bas file that is placed in it. In these folders are placed input and output files for fireball.x as shown in Fig.5.5. Output files are viewable by double clicking them. For example, one may double click "f.log" in each of these three new directories and see detailed information about the electronic structure convergence, including the final energies. The eigen energies are located in the "eigen.dat" files.

5.4.3 *Ab Initio* Molecular Dynamics

Now I demonstrate a molecular dynamics calculation using a multi-100-atom CdS nanocrystal. Before starting this simulation one may either save the previous set-

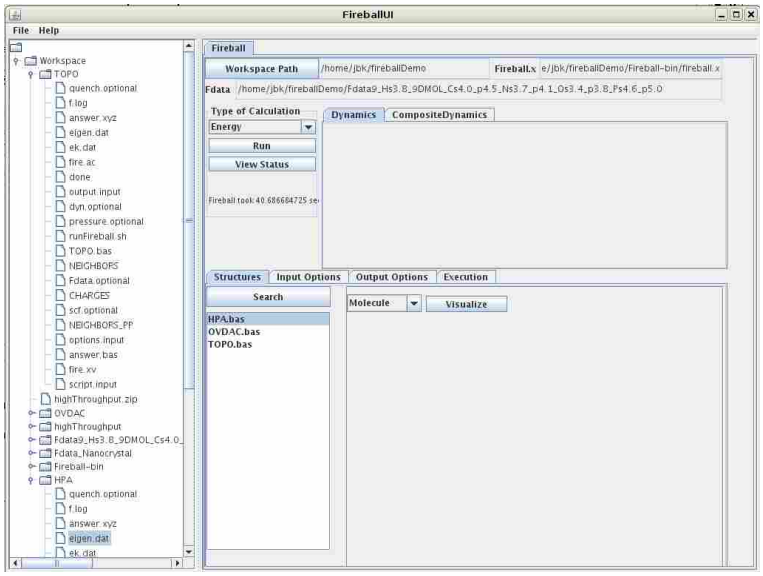


Figure 5.5 (Color Online.) FIREBALLUI after performing an energy calculation on the HPA, OVDAC, and TOPO molecules. The tree shows the folders created in the TOPO directory by FIREBALLUI.

tings for my high-throughput electronic structure example and restart FIREBALLUI or simply open a new workspace. To obtain the atomic coordinates I go to Fireball website discussed above, click on "CdS Molecular Dynamics", and unzip its contents into "fireballDemo". This creates a new sudirectory, "CdS", containing atomic coordinates ("answer.opt.bas") of the optimized spherical cluster of zinc-blende-type CdS ($Cd_{140}S_{141}$). I set **Workspace Path** as before if it is not already set and use **Search** to load the Fdata containing S and Cd that I previously downloaded, the executable, and the new input structure. I note this set of Fdata contains $3d^{10}4s^24p^0$ orbitals for Cd and $3s^23p^4$ for S with Hamann type pseudopotentials and Horsfield XC theory. From its name I see cutoff radii for the sulfur and cadmium atoms are $r_c^s = 4.2\ a_B$, $r_c^p = 4.7\ a_B$ and $r_c^d = 5.1\ a_B$, $r_c^s = 4.5\ a_B$, respectively.

In the **Type of Calculation** panel I choose **MD** and **NVT NHC velocity verlet**. I would normally select 10,000 **Time steps** for good statistics but for the

purposes of this demo I input 10. I select a **delta t** of 1.00 fs. In **Input Options** I set **Functional Theory** to **Harris** and **XC theory** to **Horsfield**. I set **Initial Temperature** to 4.170736 K and **Final Temperature** to 350.0 K. All other quantities and settings are kept at the default values.

Since quantum MD is typically cpu-intensive, I would typically submit this type of job to a remote compute cluster as FireballUI is designed to do. This would be done by going to the **Execution** tab and selecting **Remote Machines(s)**, inputting the server's name into **Remote Hosts**, and setting up the job submission requirements. In such a scenario the executable fireball.x and Fdata would need to be previously installed on the remote machine, but all input files are transferred to the remote computer via sftp. Once the job is finished, the results can be transferred back to a local workstation for analysis. For this demonstration however I submit the job on my local machine by accepting defaults in the **Execution** tab and clicking **Run**. In about an hour on a typical processor the job finishes. Fig. 5.6 provides a screen shot of FireballUI with settings used to perform the MD simulation; while Fig. 5.7 shows a model of the resultant nanocrystal.

5.5 Conclusion

In summary, the use two functionalities–high-throughput electronic structure and quantum md–of Fireball have been demonstrated through the use of FireballUI. Other examples might have shown how to perform an optimization or a calculation on a periodic structure. It is anticipated the graphical user interface to Fireball will put the power of fast *ab initio* methods more within reach of potential users (chemists, materials scientists, physicists, biologists). The Java interface also allows easy incorporation of quantum MD by software architects working in large appli-

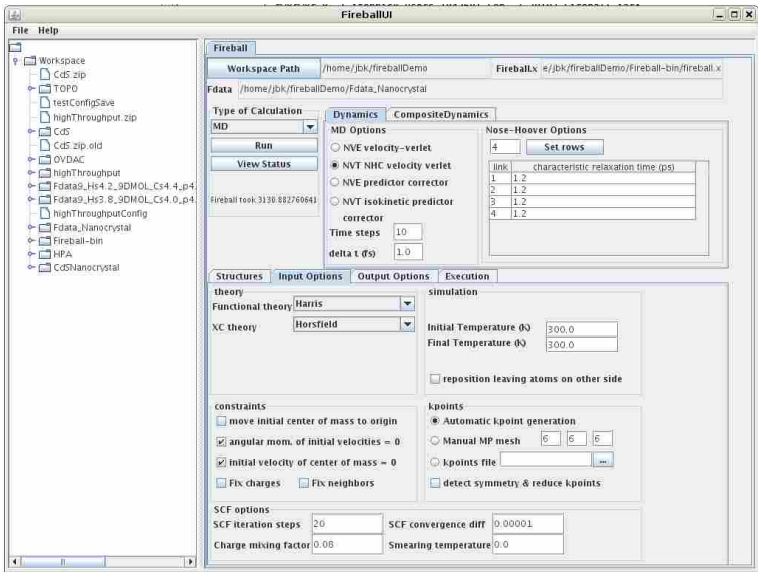


Figure 5.6 (Color Online.) FIREBALLUI setting up a MD simulation on a CdS nanocrystal.

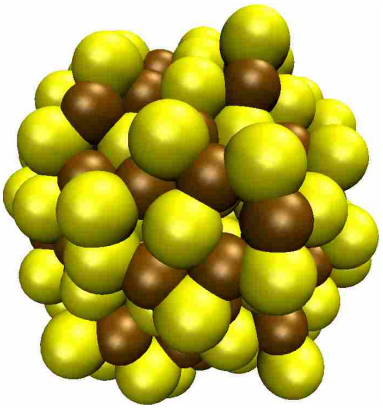


Figure 5.7 (Color Online.) Snapshot from the MD simulation of the nanocrystal heating from 4.17 K to 350 K over the course of a 10 ps time frame.

cation frameworks. Development of FireballUI is an ongoing work. Many new improvements are expected in the next release, such as basis set creation, dynamical matrix calculation, easy-to-use band structure calculation, automated unit cell volume optimization, umbrella sampling, transition point barrier crossing, thermodynamic integration, and many others. A number of data analysis tools will also be incorporated into future versions. If users are interested in implementing further capabilities, collaboration is welcome. A new wiki has been developed for this purpose at http://fireball.phys.wvu.edu/wiki/index.php/Main_Page.

Chapter 6

xyzSTATS: A program for performing configurational analysis

6.1 Introduction

Molecular dynamics (MD) simulations have been used to determine how atoms are bonded together into molecules and how systems of millions of atoms react to some stimulus as a whole. Analysis of these results, especially cases where there are many atoms, has been relegated to individual codes that are designed for a very specific purpose and that are not widely disseminated. The lack of dissemination of these codes forces each research group to develop their own tools causing a lack in scientific productivity. My goal was to write a code that could be used by a wide audience, is extensible, and can take the large amounts of data produced in a molecular dynamics simulation to distill out many types of useful information.

In designing this program a few objectives were held paramount: 1. the user is given control over the atoms and time-steps that they want to analyze. 2. I contained all analysis subroutines within modules that can be easily swapped in or

out in order to add or correct functionality without affecting the rest of the program. 3. Look forward to announced properties of Fortran 2003 and Fortran 2008 that are not yet implemented in most Fortran compilers but will make xyzSTATS better when implemented.

Because a molecular dynamics simulation can contain many atoms and run for many time-steps, the first thing I needed to do was fight information overload. To do this I allow the user to break the system up into more manageable pieces. These pieces can be defined via species-type, position in cartesian space, or via position in the xyz file. The user can also choose the time-steps that they want to analyze. In making these choices the user also limits the amount of memory that is used, because xyzSTATS only stores the atoms that fits the criteria given.

Since Fortran 95 is the most widely available version of Fortran today, I used it, but I would really like to use some of the announced features of Fortran 2003 and Fortran 2008, namely Object Oriented Programming and the data-parallel formalism found in Co-array Fortran. As such, I tried to write the code so that as the these features become widely available in Fortran compilers their use will be easy to include. Thus, to mimic objects, I contained the subroutines needed for a particular type of analysis within a single module. The main routine then calls one subroutine which acts as the class routine for that module.

Currently, the results that can be produced are:

- **Average position.** This outputs an xyz file that has the average position of each atom, by subsystem, over the course of the molecular dynamics simulation.
- **Movie files of subsystems.** This outputs the molecular dynamics evolution of each subsystem into a xyz movie file.
- **Sets of individual time-step files.** Often, I will compute the molecular

dynamics simulation in one code, or under one set of interactions, and then want to compute the electronic configuration, for a set of time-steps, with a different code or with a better set of interactions. This will extract the relevant time steps to be ran again.

- **Center-of-mass.** This computes the center-of-mass for each subsystem. For any pair of subsystems, this calculates the distance between their calculated centers-of-mass for each time-step.

- **Radial distance from a point.** This creates binned datas of the radial distance of the atoms from some fixed point in space. If reference xyz files are given the program will also quantify the quality of fit of the molecular dynamics simulation against the reference atomic positions as a function of time-step.

- **Radial distribution function, g(R).** This is a new feature which I have just added. I have not yet determined how I will analyze the raw data.

- **Neighbor-neighbor and next-nearest-neighbor distances.** Besides giving raw data, I have instituted subroutines that output the average bond length by bond type, average bond length by atom, average bond length by subsystem, average bond length by time-step. I also implemented subroutines to track the number of neighbors on a particular atom and the number of atoms that have a specific number neighbors, both as a function of time-step.

- **Angles between sets of neighboring atoms.** Here again, I produce the raw angle data, as well as, the average angle by subsystem, average angle by atom, average angle by subsystem, average angle by time-step. I also plan on implementing a routine that tracks the angles between four atoms to be used in psi vs. phi plots that are popular in computational chemistry.

- **Diffusion.** This computes the diffusion average displacement from one time-step to the next. Currently, the user will have to plot the results and fit a trend-line to determine the diffusion coefficient. I could have xyzSTATS automatically fit a trend-line to the data, but multiple cases exist where a single trend-line would be insufficient.

Sec. 6.2 will be concerned with some of the basic statistical concepts that I use (and in one case have changed). In Sec. 6.3 I will discuss the structure of the input files used by xyzSTATS. I will discuss a few of the more interesting output files that are produced by xyzSTATS in Sec. 6.4. A listing of all output files is given in the Appendix. I will give an example of running xyzSTATS in Sec. refsec:xyzexample, followed by a few concluding remarks.

6.2 Basic Statistics

Definitions of common descriptive statistical analysis techniques can be found in any textbook, and I have used most without any change; these techniques include mean, median, standard deviation, variance, coefficient of variation, and Pearsonian skew [99].

Because using binned data (i.e. in a binned data) is the best way to represent some types of data, particular effort was made in making sure the analysis of the binned data could be quantitatively compared with reference systems. Of particular use when comparing two sets of binned data is the Chi-squared statistic [77, 100].

6.3 Input Files

To run xyzSTATS there are three files that the user will need, not including the xyz file of the molecular dynamics evolution or any xyz files of the reference systems: script.inxyz, comparison.inxyz, and neighbors.inxyz. Three optional files may be used: Odoyle_rules.inxyz which creates a subsystem based on a set of rules on the positions of the atoms, a file containing lattice information, and a file containing information on user defined subsystems. I will discuss each of these files below.

6.3.1 script.inxyz

The file script.inxyz (see Fig. 6.1) allows the user to choose the types of analysis to perform. The first five lines specify parameters that are used by all of the analysis subroutines. The first line specifies the name of the xyz file containing the molecular dynamics evolution. The next two lines specify the number of times-steps in the molecular dynamics evolution and the length of each time-step. As I have stated, one of my major goals was to allow the user to break up the molecular dynamics evolution into smaller, more manageable chunks, the fourth line of script.inxyz allows the user to specify the time-steps on which to start and stop the analysis. The fifth line specifies if the system is non-periodic (set this line to "0"), or periodic and the user has a lattice vectors file "1" or a file with lattice-constants "2".

The rest of the lines in script.inxyz are used to control the various types of analysis being performed. The basic format for each type of analysis starts off with a toggle for turning the analysis on or off. The following line allows the user to specify the subsystems that will be considered in the analysis. This is followed by lines for specifying any other parameters needed to do the particular analysis.

While having the user specify the subsystems that will be studied for each type of

```
answer.xyz                      MD file name
10000                               Number of time-steps in MD file.
1.00                            Length of each time-step (in fs)
1,10000                             Beginning and Ending time steps
0                               cluster(0), lat vectors(1), lat constants(2)
0                               Single point toggle
0                               whole system(0),species(1), user define(2)
0                               Produce an single point file every nth steps
0                               Average Position Toggle
0                               whole system(0),species(1), user define(2)
0                               Center of mass toggle
0                               whole system(0),species(1), user define(2)
1                               radial toggle
0                               whole system(0),species(1), user define(2)
0.0,0.0,0.0                     Position to measure from
1                               Radial compare
25                              Number of bins in Radial histogram.
0.00,13.00                      Histogram cutoff radius.
0                               GofR toggle
0                               whole system(0),species(1), user define(2)
1                               GofR compare
100                             Number of bins in GofR histogram.
0.00,13.00                      Histogram cutoff radius.
0                               Neighbors toggle
0                               whole system(0),species(1), user define(2)
2.2,2.7                         min,max NN cutoff
100                              Number of bins in NN histogram.
2.2,4.2                         min,max NNN cutoff
400                              Number of bins in NNN histogram.
1                               neigh_hist_toggle
0                               Diffusion Toggle
0                               whole system(0),species(1), user define(2)
0                               produce zero_one data from comparison files
zeroone.txt                         optionally read if user define(1)
```

Figure 6.1 (Color Online.) Example of the file script.inxyz.

analysis might seem superfluous, it gives the user the ability to tailor the calculations performed, allowing them to cut down on the number of times they need to run xyzSTATS. Several ways exist in which the user can control which set of subsystems they analyze. The set of subsystems can change from one analysis type to the next. The most obvious subset to analyze is the complete set, this choice is made by setting the appropriate line to zero. xyzSTATS can break a system into a set of subsystems based on the atomic species of the molecular dynamics simulation. To do this, the user sets the appropriate line equal to one. The may also specify that certain atoms make up a subsystem through setting the appropriate line to the value 2 and by supplying a file that specifies which atoms are in the subsystem; this file is specified on line 34 of script.inxyz. The format of this file will be discussed in more detail later. The user is further able to specify a subsystem by specifying a region of cartesian space in the optional file Odoyle_rules.inxyz.

The first type of analysis is more clerical than scientific, the user might want to take a subset of the molecular dynamics evolution time-steps and create a set of files of the atomic positions at individual time-steps. To do this, I turn on the toggle on

line six by changing the zero to a one. On line seven the user specifies the subsystem, and on line eight the number of time-steps skipped between each single point file. The user is also able to use the same subroutine to create a single xyz file that contains a molecular dynamics evolution but with only the specified subsystem and the specified time-steps. This is done be setting line six to the value 2.

The second type of analysis is also clerical. It will write out a set of files containing the average positions of the atoms in the original molecular dynamics evolution. The number of files written out is dependent on the number of subsystems chosen using line ten. No other parameters need to be specified for this type of analysis.

The third type of analysis listed in script.inxyz focuses on calculating the center-of-mass of a set of subsystems as a function of time-step. If more than one subsystem is specified the subroutine also calculates the distance between each pair of center-of-mass as a function of time-step. No other parameters need to be specified for this type of analysis.

The fourth type of analysis is the radial distribution of atoms from a point in cartesian space as a function of time. This program is an outgrowth of analysis tools that were originally used for working with spherical nanoparticles and quantum dots, thus the limitation to distribution from a point. In future versions of this code this routine will be generalized to look at distributions from a line as well as a plane. The first two lines are normal, turning on the analysis and specifying subsystems, but there are four lines of parameters. The first parameter (on line 15 of script.inxyz) is a set three of real numbers that specify the point in cartesian space that the distribution is measured from. The next line is a toggle that turns on or off the comparison of the molecular dynamics evolution with the reference systems given in comparison.inxyz (more on the reference systems later). In this module, I bin the radial data to form binned datas. The number of bins to use in this analysis is specified on line 17 of

script.inxyz. The user specifies the region of space over which the bins should be made by giving a minimum radius and maximum radius on line 18.

On lines 19 through 23 I work with the g(R) analysis. On line 21 I turn on or off the comparison of the molecular dynamics evolution with the reference systems. Line 22 sets the number of bins used in the analysis. And line 23 sets the region of space over which the bins should be made by giving a minimum radius and maximum radius.

Lines 24 through 30 are concerned with the analysis of neighbor distances and with the angles between atoms. By setting line 24 to the value of 1, 2, or 3 the user turns on the neighbor analysis, angle analysis, or both neighbor and angle analyses. Lines 26 and 27 specify the number of bins and the domain of values in which a nearest neighbor should be found. Lines 28 and 29 define the number of bins and the domain of values in which a next-nearest neighbor should be found. Line 30 turns on or off the Ayres plot function for nearest neighbors. Lines 31 and 32 are for computing the diffusion of the system.

By setting line 33 to the value "1", the user will cause the first time-step of the molecular dynamics evolution to be compared with the reference systems specified in comparison.inxyz. This setting create two files: "MDvscomparisonatoms01.txt" and "MDvsNOTINcomparisonatoms01.txt." The format for these two files follow the description of the user defined susbsystem files that are discussed in Sec. 6.3.5. If there are atoms that are in the same positions in the molecular dynamics evolution and in a reference system, then it's position in the MD file will be noted (as well as it's position in the comparison file) in MDvscomparisonatoms01.txt. All atoms in the MD file that are not found to be in the same position as in the comparison file wil be denoted in MDvsNOTINcomparisonatoms01.txt.

```
Place the names (and paths if needed) of all of the files
you would like to compare the MD results to
************************************************************
2                        number of files to read in
************************************************************
W.xyz
ZR.xyz
```

Figure 6.2 (Color Online.) Example of comparison.inxyz.

6.3.2 comparison.inxyz

As one of the goals of this project was to simplify the process of comparing a molecular dynamics evolution against some reference system, or systems, I need a way to specify what those reference systems are. This is accomplished by using comparison.inxyz, (see Fig. 6.2 for an example of the format). The first three and fifth lines of comparison.inxyz are explanations of the purpose of that file and also serve to delineate the placement of the data. On the fourth line I specify the number of files I will be comparing against the molecular dynamics evolution. On the sixth lines and following, I state the names of the files. Each file must be in the folder in which xyzSTATS is being ran.

6.3.3 neighbors.inxyz

The file neighbors.inxyz is used to specify the number of neighbors and the species of the neighbors that each atom species may have. An example of the format for neighbors.inxyz is given in Fig. 6.3. The first line in neighbors.inxyz gives the number of species combinations of nearest neighbors. The next line or lines, depending on the value given on the first line, give rules concerning the maximum number of nearest neighbors an atom of a particular species may have. The format for the lines giving the neighbor rules is

species$_1$, species$_2$, number-of-nearest neighbors, interchange toggle

Figure 6.3 (Color Online.) Example of neighbors.inxyz.

where **species**$_1$ and **species**$_2$ are a set of atomic species, which may be the same species, **number-of-nearest neighbors** is the maximum number of nearest neighbors of type **species**$_2$ that **species**$_1$ may have, and setting *interchangetoggle* to a value of one indicates that *species*$_2$ may also have **number-of-nearest neighbors** of type **species**$_1$. Setting **interchange toggle** to the value of two indicates that **species**$_2$ may not have **number-of-nearest neighbors** of type **species**$_1$. For example, if the user is working with CdS in a wurtzite or zinc-blende configuration, there will only be one pair of atom types Cd and S. Whereas, if the user was working on TiO_2 in an anatase structure, the user would need two have two different rules for the number of nearest neighbors since the Ti would have six O neighbors while the O have four Ti neighbors.

6.3.4 Odoyle_rules.inxyz

In some simulations the user might want to break up a system into subsystems based on position of the atoms in space. The creation of subsystems based on the atomic positions is determined by setting rules to define which atoms are to be included. Any number of rules can be applied. Each rule describes a simple geometric shape. The rules are supplied using an optional file called Odoyle_rules.inxyz; see Fig. 6.4. It is optional in that xyzSTATS is designed to read it and use it if the file exists in the folder in which xyzSTATS is being ran.

Figure 6.4 (Color Online.) Examples of Odoyle_rules.inxyz. (a) Example of the rule needed for keeping all atoms within, and on, a sphere of radius 5 Å. (b) Example of keeping all atoms within the unit cube.

As stated above each rule produces a simple geometric shape. All of the rules have the following form

$$ax^i + by^j + cz^k \spadesuit d^l \tag{6.1}$$

where (a,b,c,d,i,j,k,l) are real numbers that are set by the user, and ♠ can be any of the following equality symbols (<=, <, =, >=, >). The first line in the file contains the number of rules. The rest of the lines list the rules. Fig. 6.2 (a) gives the general structure of a file, (b) the example of the rule needed for keeping all atoms within, and on, a sphere of radius 5 Å, (c) and the example of keeping all atoms within the unit cube.

6.3.5 User Defined Subsystems

As was stated in Sec. 6.3.1 the user gives the name of the file containing the user defined subsystems on line 34 of script.input, see Fig. 6.1. Whatever the name of the file is, the first line of it should contain the number of subsystems that will be defined. Line 2 is not read by xyzSTATS and may be left blank or may contain human readable information. The next line will contain the number of atoms in the first subsystem

Figure 6.5 (Color Online.) Example of a file that defines a set of two user defined subsystems.

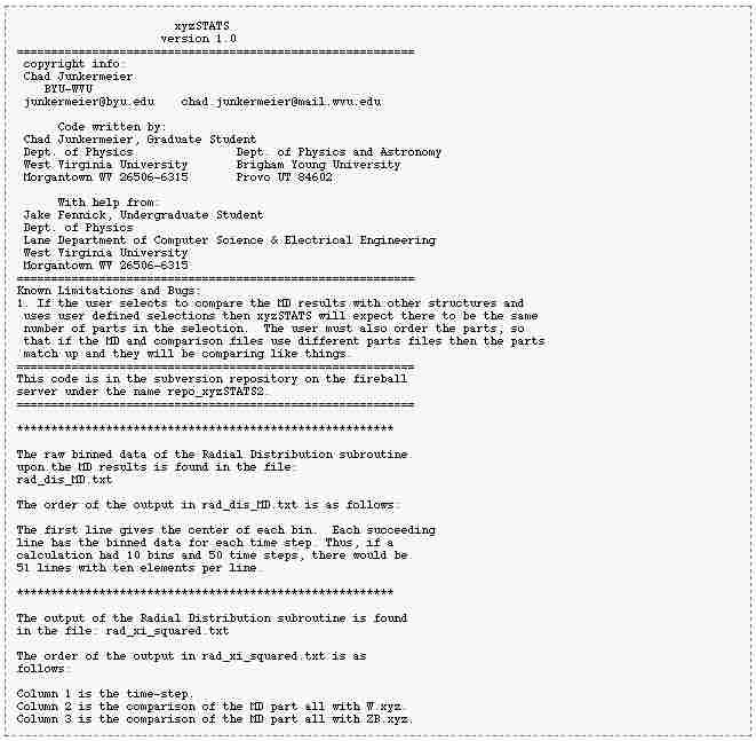
defined along with a name for the subsystem. Each of the immediately following lines containing the file position of one of the atoms in the subsystem.

If the file contains a second subsystem the user will need to skip a line, and then give the number of atoms and subsystem name for the next subsystem. The line containing the number of atoms and subsystem name will be followed by lines specifying the file position of each atom in that subsystem. Each successive subsystem will be defined in the same way. Fig. 6.5 contains a simple example of this type of file.

6.4 Output Files

There are over twenty different output files that may be produced using xyzSTATS. Discussions of each output file type is outside the scope of this work. Discussion of the format for most of the output files is written during each run of xyzSTATS into a file called logfile.txt. In logfile.txt, I first write out information concerning the version of xyzSTATS that was used to create the output files. This information is followed by author and copyright information, known bugs, an information regarding where the user can obtain the software. See Fig. 6.6 for an example of a logfile.txt.

While I are not going to give a full review of all of the output files, there are



```
                    xyzSTATS
                   version 1.0
============================================================
 copyright info
 Chad Junkermeier
    BYU-WVU
 junkermeier@byu.edu    chad.junkermeier@mail.wvu.edu

      Code written by:
 Chad Junkermeier, Graduate Student
 Dept. of Physics                Dept. of Physics and Astronomy
 West Virginia University        Brigham Young University
 Morgantown WV 26506-6315        Provo UT 84602

      With help from:
 Jake Fennick, Undergraduate Student
 Dept. of Physics
 Lane Department of Computer Science & Electrical Engineering
 West Virginia University
 Morgantown WV 26506-6315
============================================================
Known Limitations and Bugs:
1. If the user selects to compare the MD results with other structures and
 uses user defined selections then xyzSTATS will expect there to be the same
 number of parts in the selection.  The user must also order the parts, so
 that if the MD and comparison files use different parts files then the parts
 match up and they will be comparing like things.
============================================================
This code is in the subversion repository on the fireball
server under the name repo_xyzSTATS2.
============================================================

**********************************************************

The raw binned data of the Radial Distribution subroutine
upon the MD results is found in the file
rad_dis_MD.txt

The order of the output in rad_dis_MD.txt is as follows:

The first line gives the center of each bin.  Each succeeding
line has the binned data for each time step.  Thus, if a
calculation had 10 bins and 50 time steps, there would be
51 lines with ten elements per line.

**********************************************************

The output of the Radial Distribution subroutine is found
in the file  rad_xi_squared.txt

The order of the output in rad_xi_squared.txt is as
follows:

Column 1 is the time-step.
Column 2 is the comparison of the MD part all with W.xyz.
Column 3 is the comparison of the MD part all with ZR.xyz.
```

Figure 6.6 (Color Online.) Example of logfile.txt

a few that I would like to make mention of: atoms_with_m_nghbrs_extended.txt, rad_xi_squared.txt, and a type of Ayres plot. A listing of all of the other file types is given in Appendix.

Some of the output files contain only the information of one of the subsystems. When a file contains information on only one subsystem the filename reflects this by having the name of the subsystem as part of the name of the file. I have signified this in my description of the output files by using the flag "SUBSYSTEM." One such file that I describe later is called gr_raw_SUBSYSTEM.xyz.txt

6.4.1 atoms_with_m_nghbrs_extended.txt

The file structure for atoms_with_m_nghbrs_extended.txt gives the time-step, followed by the number of atoms with zero neighbors. If there are any atoms with zero neighbors, they are written out on the lines immediately following. On these lines, the species of the atom is written out followed by the position of the atom in the xyz file. After all of the atoms with zero neighbors are written out, the following line specifies the time-step and the number of atoms with 1 neighbors. That line is followed by a listing of all of the atoms with 1 neighbor. The specifying of the number of atoms with m neighbors and a listing of those atoms continues through the maximum number of neighbors specified in neighbors.inxyz. Once this is done for the first time-step it goes on the the next time-step and repeats the process until all time-steps have been analyzed. An example of the file appears in Fig. 6.7. Since I discussed the atoms_with_m_nghbrs_extended.txt file I should also note atoms_with_m_nghbrs.txt which doesn't have as much detailed informtion in it but is more streamlined and ready for analysis using something like R [101].

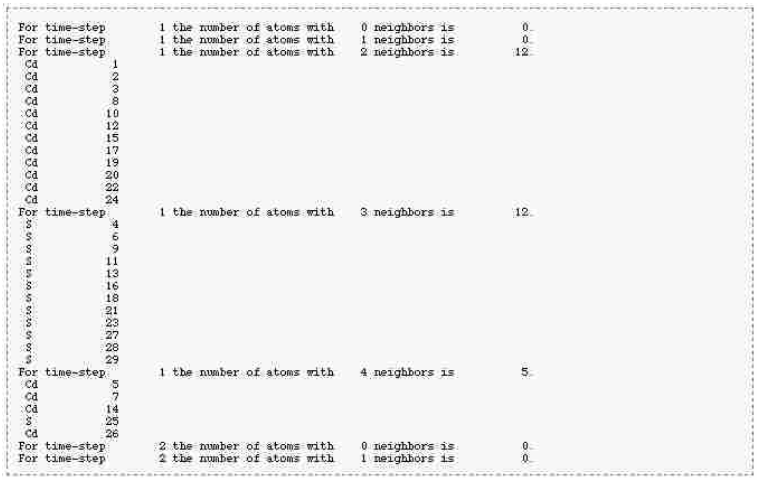


Figure 6.7 (Color Online.) Example of atoms_with_m_nghbrs_extended.txt This shows the results for a CdS cluster of 29 atoms. I have shown the results for the first time-step and the first few lines of the second time-step.

6.4.2 rad_xi_squared.txt

As was stated before, one of the aims of the xyzSTATS project was to be able compare and contrast molecular dynamics evolutions with any given set of reference systems. So far, one of the few things I have written into xyzSTATS, in an effort to achieve this goal, is a Chi-squared analysis of the radial distribution of molecular dynamics evolution data as compared to the radial distribution of a set of reference structures. The output of this analysis is placed into a file called rad_xi_squared_SUBSYSTEM.txt, where SUBSYSTEM is the name of some subsystem (i.e. all, species, etc.). Each line contains the results for a single time-step. The first element in a line is the time-step. This is followed by a set of real numbers each corresponding to the Chi-squared comparison of a particular reference system against a time-step of the molecular dynamics evolution. We use logfile.txt to indicate which real number is associated with which Chi-squared comparison. An example of a few lines from a rad_xi_squared_SUBSYSTEM.txt is given in Fig. 6.8.

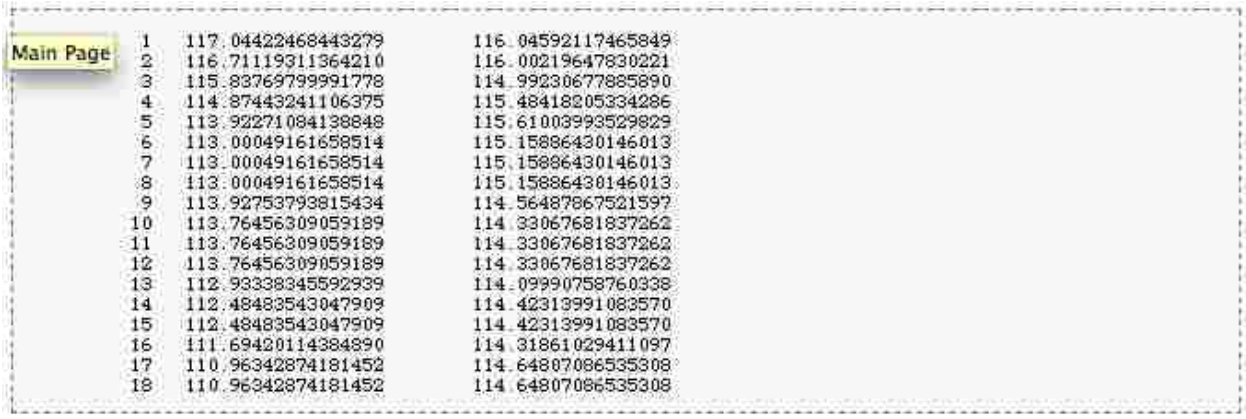


Figure 6.8 (Color online) Example of the file structure of rad_xi_squared.txt.

6.4.3 NN_Ayres_timestep_SUBSYSTEM.txt and NNN_Ayres_timestep_SUBSYSTEM.txt

Edward Tufte, an eminent author and speaker on the conveyance of information through graphs and figures, gives an example of a binned data designed by Colonel Leonard P. Ayres that uses symbols to a new symbol to signify each data point found in a binned data column (this is different from a stem-and-leaf style binned data). The symbol that he uses is the value of the data. An example of this type of plot is given in Fig. 6.9. Tufte argues that this is an effective graph because it has a higher information density then a normal binned data graph [102]. Since I have found no other name for this type of graph I will call this type of binned data an "Ayres plot." While publishing this type of graph is only practical for small data sets, it can be very useful in determining what is causing a particular peak in a binned data, thus the analysis types that use binned datas can produce one of these graphs. The user can choose to obtain this type of graph for the average nearest neighbor and average next-nearest neighbor distance as a function of time-step by setting line 30 of script.inxyz to one. A simple graph of this type can be found in Fig. 6.9. Fig. 6.9 (a)

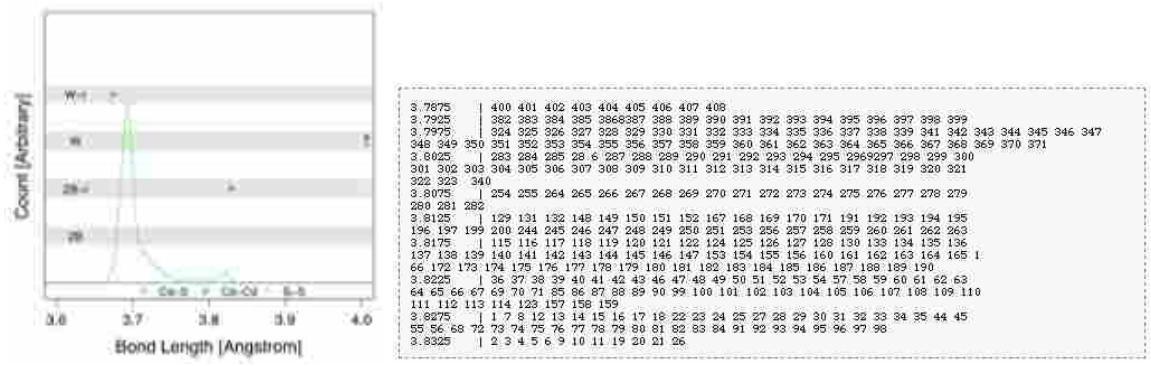


Figure 6.9 (Color Online.) Example of an Ayres plot. In this plot, I have plotted the radial distribution of atoms in a nanocrystal from center of the nanocrystal when in a bulk-like zinc-blende structure.

is a binned data showing average next-nearest neighbor distances between every set of two S atoms over the course of a 10,000 time-step molecular dynamics simulation. The simulation is of a CdS nanoparticle that started off in a relaxed structure (relaxed from the bulk-like zinc-blende structure). Inspection of this graph shows that there are two peaks present, one being much the larger. The question becomes, what is the cause of the two peaks? Knowing which time-steps make up each peak may be helpful in understanding the result. In Fig. 6.9 (b) I show a portion of he Ayres graph that was made of the S-S average next-nearest neighbor distances and is given in NNN_Ayres_timestep_S.txt. I only give the portion of the Ayres graph that focuses on the smaller of the two peaks shown in (a). I find that the smaller peak is wholey made up of the average next-nearest neighbor distances from the first few hundred time-steps.

6.5 Example

I will now provide an example of how a user might use xyzSTATS to obtain useful information concerning a system of interest. My test system is a CdS nanoparticle that has undergone a molecular dynamics simulation of $10,000$ time-steps at constant temperature. I will compare the radial position of the molecular dynamics evolution's atoms at each time-step with a couple of reference systems: nanocrystals in bulk-like wurtzite and zinc-blende structures.

First, I place the file containing the molecular dynamics evolution, called answer.xyz, as well as the two reference files, W.xyz and ZB.xyz, in a folder. Then, within the same folder I also place script.inxyz, comparison.inxyz, and neighbors.inxyz, each file having the settings found in Figs. 6.1, 6.2, and 6.3, respectively. I now run xyzSTATS from within the folder.

Since I was looking at the radial distribution of the atoms in the molecular dynamics evolution and compared them against the reference wurtzite and zinc-blende systems, and turned off all other types of analysis xyzSTATS produced only three output files: logfile.txt, rad_dis_MD.txt, and rad_xi_squared.txt.

The log file (see Fig. 6.6) specifies that I have performed the radial distribution analysis on the molecular dynamics evolution and that I compared the radial distribution of the molecular dynamics evolution against the radial distribution of the reference systems. It also gives us the format of the data in rad_dis_MD.txt and rad_xi_squared.txt (see Figs. 6.8 and 6.10).

Using a graphing program I produced a graph (see Fig. 6.11) of the data from rad_xi_squared.txt. The dashed line near the bottom of the graph is the critical value of my system. The critical value was found by taking the number of bins used, 25, and subtracting by 1 to obtain the number of degrees of freedom, 24. Once I have

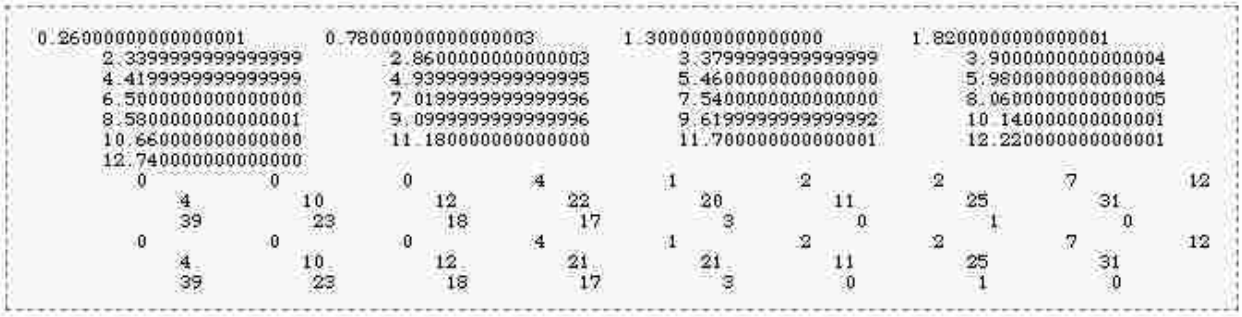


Figure 6.10 (Color online) Example of the file structure of rad_xi_squared.txt.

the number of degrees of freedom, I then look up the critical value in an introductory statistics book to find that the value is 36.41. In reviewing my plots of the Chi-squared values of the molecular dynamics evolution, I find that it appears to be neither wurtzite or zinc-blende in nature [40].

6.6 Conclusion

Development of this project is still ongoing. I would like to find more ways to automate the comparison of the molecular dynamics simulations against reference data. I would also like to be able to take the data that I produce in separate modules and produce correlation tables to identify variables that might be highly linearly correlated.

Our research group controls the development of the *ab initio* tight-binding molecular dynamics code called FIREBALL as well as the graphical user interface front end to FIREBALL called FireballUI [38, 41–43]. Ongoing development of FireballUI includes many types of analysis. In order to perform this analysis, functionality has been added to it in order to read all of the input files and most output files used/produced by FIREBALL. xyzSTATS is slated for inclusion in future versions of FireballUI as an

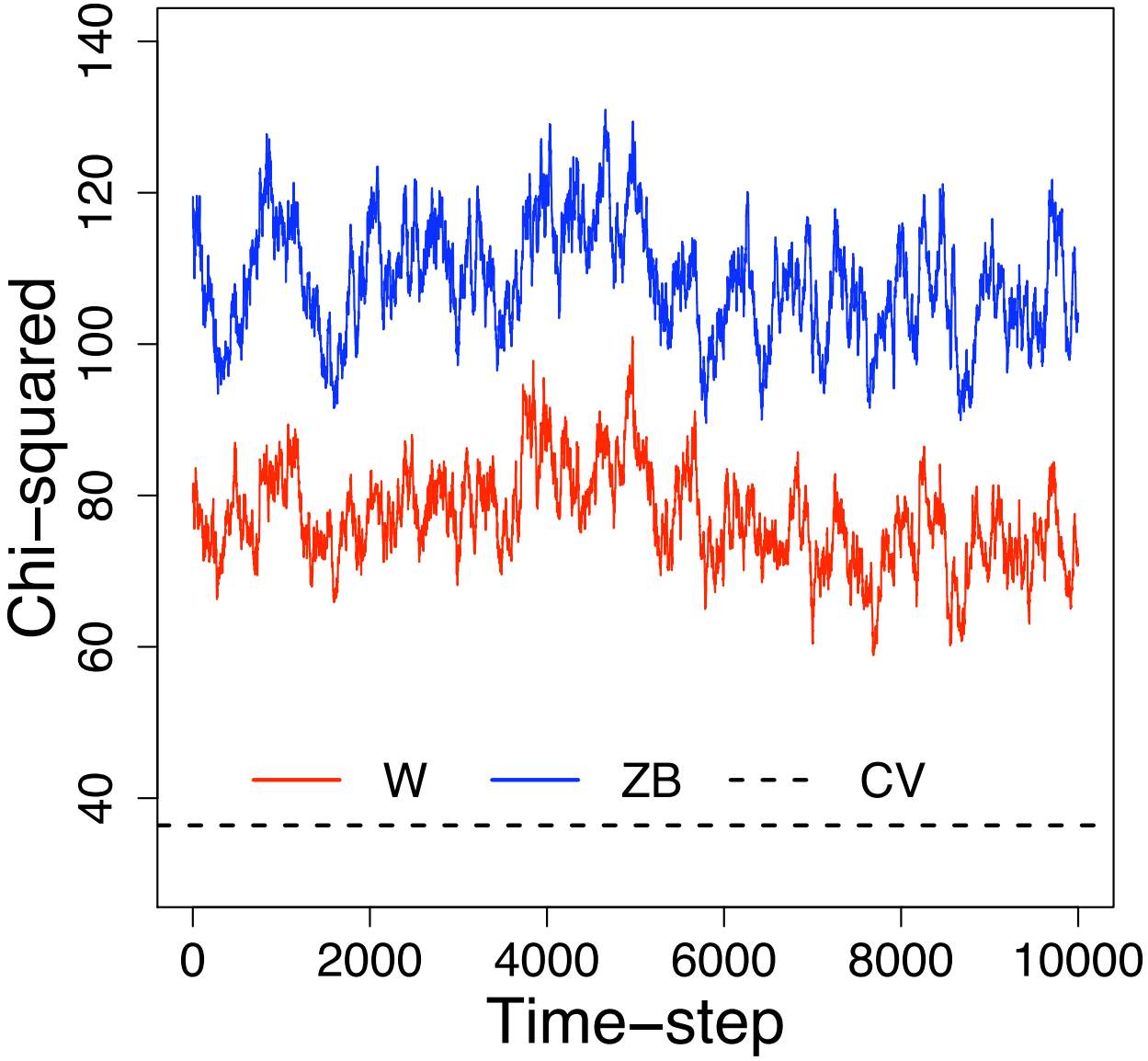


Figure 6.11 (Color online) Plot comparing the Chi-squared fit of the MD results against W and ZB.

analysis tool. When included, xyzSTATS will obtain parameters for neighbor types and nearest neighbor and next-nearest neighbor ranges directly from FIREBALL input files.

Chapter 7

Conclusion and Future Directions

7.1 Conclusion

When I started the research that culminated in this dissertation, to my knowledge, there had been no published results of *ab initio* calculations of semiconducting nanoparticles with over 200 atoms [16,17,19,20]. Further, even at this time there are only a few publications, that I know of, containing results from *ab initio* calculations on over two hundred atoms [18,21,22].

The calculations in this dissertation were performed using FIREBALL enabling *ab initio* simulations of CdS nanoparticles of up to 357 atoms. With this ability I was able to show that partial passivation due to relaxation of the surface recovers some of the bulk-like semiconductor properties even though each system is in nanoparticle form this includes the following (not all of which are necessarily desirable). The band gap of the relaxed structures are smaller than the band gaps of the coresponding bulk-like structures bringing them more in line with the bulk band gap. The number of dangling bonds are decreased and with that the number of states in the gap are decreased.

In each of the references given so far in this chapter the MD calculation(s) were used to optimize nanoparticle structures. This means that the forces were (or energy was) minimized to give a ground state structure at some desired temperature (often zero Kelvin). The goal of this type of calculation is to find a structure where the atoms are not moving and any analysis is performed on only the last time-step; this type of simultation was performed for Ch. 3. In addition I performed molecular dynamics simulations where the goal of the calculations were not to find a ground state structure, but to model how atoms move when in thermodynamic equilibrium. The calculations I performed allowed for small perturbations to the atomic energies that mimic thermal flucturation. To my knowledge this type of simulation on a nanoparticle has only been reported one other time, and it was on a thirteen atom Germanium cluster [103]. Using xyzSTATS to analyze long period molecular dynamics simulations at constant temperature, I showed that CdS nanoparticles, at least for those with no passivating layer, are amorphous in nature.

I participated in the design and testing of the FireballUI. With it, future users of FIREBALL will be able to more efficiently create Fdata, run simulations and analyze the results. As part of the analysis, I created the program xyzSTATS. With xyzSTATS the user is able to determine a number of properties, statistical and otherwise. With the functionality planned for FireballUI and xyzSTATS high throughput analysis will enable scientists to quickly and accurately produce significant results.

7.2 Future Direction: Passivation with Hydrogen

In Ch. 3 I discussed nanoparticles without a passivation layer. In real world systems each of these particles would have a layer of passivating atoms or molecules. In each system I studied I failed to included a layer of passivating because they added an extra

level of complexity. I thought it best to first understand the simplest systems I could design and then incrementally add levels of complexity to build on my understanding. The next level of complexity that I need to understand is attaching simple passivating agents to the surface of a nanoparticle. The simplest choice is adding H atoms. Hydrogen attaches to the S atoms passivating their dangling bonds.

Yu *et al.* claim that structural relaxation cannot passivate enough of the dangling bonds to clear the band gap of dangling bond states [37]. The data presented in Ch. 3 supports this claim. In the same work, they state that the gap states are largely due to surface atoms that have two, or more, dangling bonds. This statement is surprising, in that the majority of surface atoms have only one dangling bond.

I started with the nanoparticle configuration RCdS264 from Ch. 3. For lack of a better naming convention I named each atom in RCdS264 by its position in the configuration file; so the first atom in the file is named RCdS264_1 and the last atom is named RCdS264_264.

Using the program xyzSTATS, I find that RCdS264 has 147 atoms with one dangling bond, 31 atoms with two dangling bonds, and three atoms with three dangling bonds. To determining how the (partial) passivation an atom with either one, two, or three dangling bonds I created three separate systems. The first system had a H atom attached to an atom with one dangling bond. The second system had a H atom attached to an atom with two dangling bonds. The third had H attached to one with three dangling bonds.

In actuality, I had to, in each case, attach one H atom to each of two atoms, because FIREBALL will not find the correct results for a system with an odd number of electrons. The atoms that the H atoms were attached to were pick so that they were on opposite sides of the nanoparticle, this was done to minimize the affect of one passivated atom on the other. Thus, the three systems became:

1. H attached to atoms with one dangling bond. H atoms were attached to RCdS264_6 and RCdS264_259, each of which had one dangling bond. I will call this system 1HRCdS264.

2. H atoms were attached to RCdS264_87 and RCdS264_176, each of which had two dangling bonds. This system is called 2HRCdS264.

3. H atoms were attached to RCdS264_36 and RCdS264_176. Atom RCdS264_36 has three dangling bonds, while RCdS264_176 still only has two. Even though there were two other atoms that had three dangling bonds, I decided that they were each too close to RCdS264_36 and opted to instead reuse atom RCdS264_176 because it was on the opposite side of the nanoparticle. This system is called 3HRCdS264.

To determine the effect of passivating dangling bonds I plotted the Löwdin population of each state on the atoms that are being passivated. By comparing the Löwdin population on each atom before and after passivation I can determine if any states in the band gap region are due to dangling bonds. The other method is to look at how the passivation of a single dangling bond affects the electronic state of the whole nanoparticle. To do this, I plot the bulk probability before and after passivation. The state(s) that moves is affected.

The Löwdin population as a function of energy, on atoms RCdS264_6 and RCdS264_259, is found in Fig. 7.1 (a). While both RCdS264_6 and RCdS264_259 have one dangling bond, neither of them have a state in the band gap region. I do find that both atoms have states that are of stronger localization (stronger than the surrounding states) just below the band gap. After passivation with H, the localization moves to lower energies.

Both RCdS264_87 and RCdS264_176 have highly localized states near the valence

band edge, as shown in Fig. 7.1 (b). Attaching the H atoms again moves the localization to lower energies, but with the states being far less localized on that surface atom. This confirms the notion of Yu *et al.* that atoms with one dangling bond do not play a significant role in closing the band gap of CdS nanoparticles. That atoms with two or more dangling bonds affect the band gap of CdS nanoparticles is further justified in Fig. 7.1 (c). I see that RCdS264_36 has three states of high localization within the band gap before passivation. Adding on H atom, should passivate one of the dangling bonds. After passivation, RCdS264_36 is left with two states in the band gap region that are highly localized on that atom.

With the preliminary results given here, I have shown that as dangling bonds are passivated, the band gap is opened up. This should in turn increase the quantum yield and my ability to tune a nanoparticle to absorb/emit light at a desired wavelength.

7.3 Future Direction: CdS-ZnS Core-Shell Quantum Dots

In Ch. 3, I studied the effects of the unpassivated surface of CdS nanoparticles on the band gap of the nanoparticles [39]. I defined every atom that had four nearest neighbors to be a bulk atom. In analyzing the nanoparticles I used the bulk probability, which is the probability that an electron in a molecular orbital would be found on an atom that I considered to be a bulk atom. Further, I used the number of accessible atoms to determine the localization of each molecular orbital. The smallest of these systems, a stoichiometric system which was centered on the midpoint of a nearest neighbor bond, is the structure that I use for the core of the core-shell quantum dots that I study here. I give representations of the bulk probability and number of accessible atoms that I obtained in that study, in Fig. 7.2. I determined the TVB

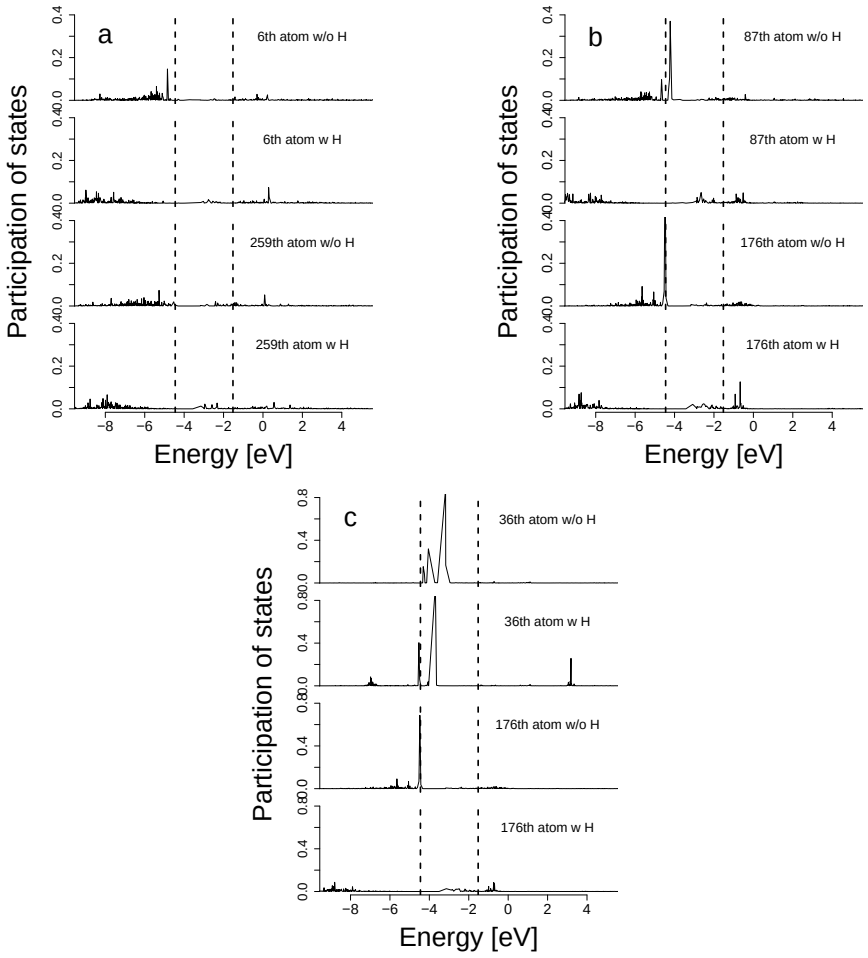


Figure 7.1 Examples of how passivation of dangling bonds decreases the probability of occupied gap states.

and BCB of the CdS core. For the bulk-like zinc-blende structure, the TVB to be at -4.45219 eV and the BCB to be at -1.52261 eV resulting in a band gap of 2.92958 eV. I will return to this result shortly.

In Fig. 7.3 I present the Löwdin probability of each state, by region, as well as the number of accessible atoms. I find that the gap states become more localized and are centered on the surface of the nanoparticles, following the trend discussed in Ref. [39]. More interesting is what appears to be the disappearance of any semblance of a band gap in the shell of CdS1418. The disappearance being caused by dangling bonds on the surface of the nanoparticle.

In Refs. [6] and [39] the measure of the bulk probability, which is the measure of the probability of an electron in a molecular orbital being found somewhere other than the surface of the nanoparticle, was shown to have a particular behavior in the absence of dangling bonds. This behavior had a large probability on the S atoms in the valence band, and a much lower probability on the Cd atoms in the valence band, giving the shape of a two pronged fork with the points of the tines towards the band gap. The reversed shape was found in the conduction band, with the higher probability being on the Cd atoms.

In the case of the shell probability of CdS1418 in Fig. 7.3, I find two very well defined pronged structures. But, whereas in the smaller nanoparticles I found that there was a space between the valence and conduction bands, there is none here. The lack of a gap leads us to believe that the as the number of dangling surface bonds increases, even if the number of dangling bonds per surface area remains nearly constant, the surface becomes more metallic. As I are only concerned with the ZB structures at this point, some of this may become inconsequential as the structures are relaxed, causing a partial passivation of the surface.

Since the gap states are located on the surface, the core of the structures end up

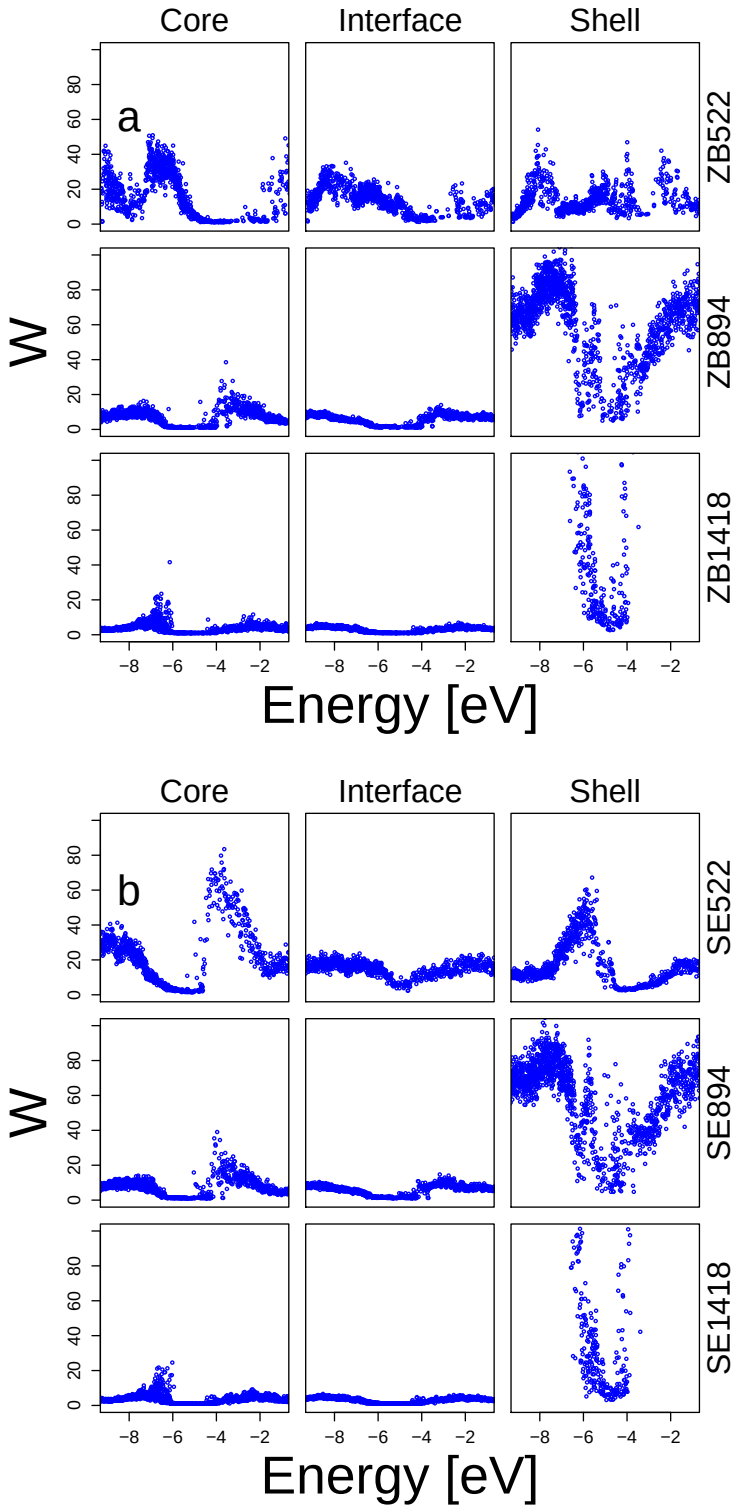


Figure 7.2 Number of accessible atoms, W. **a.** zinc-blende structures. **b.** Relaxed structures.

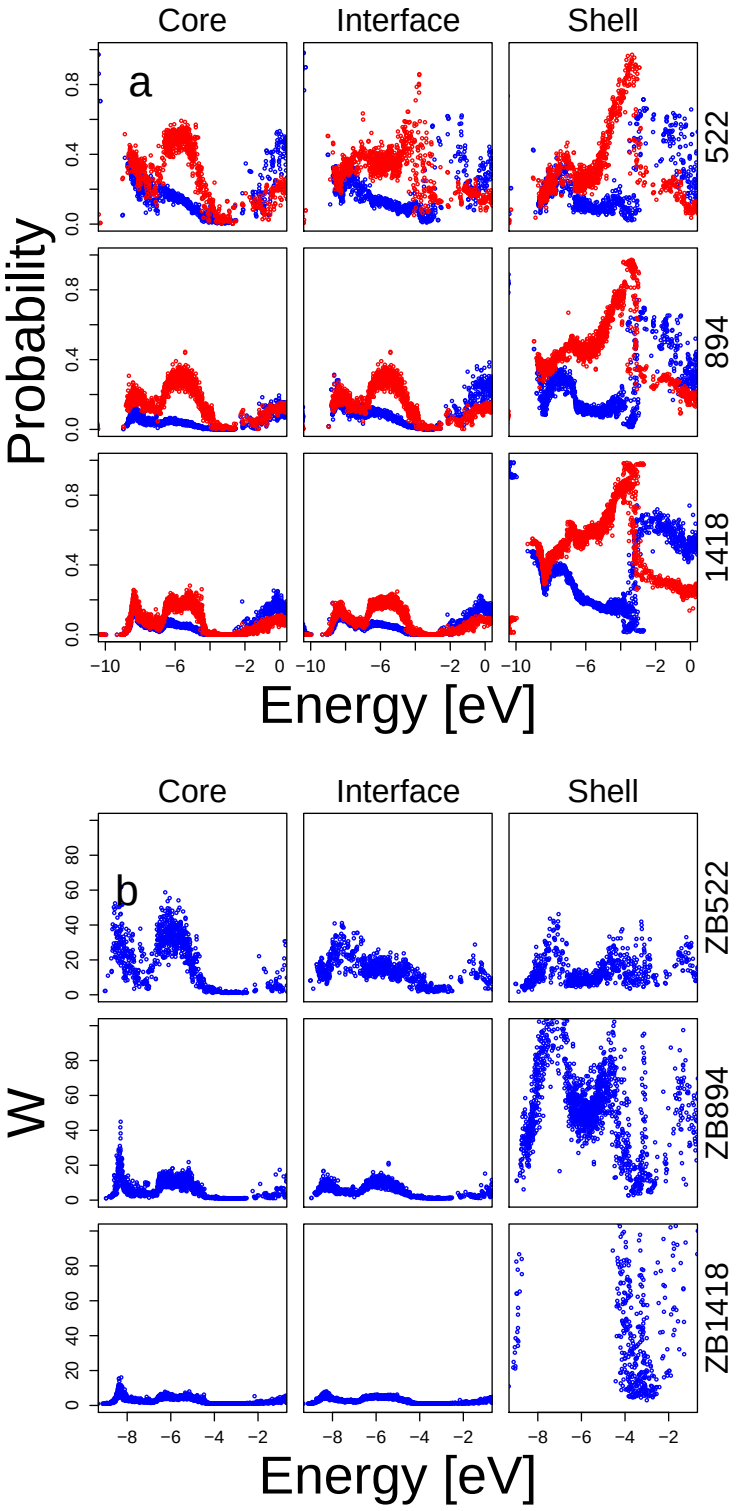


Figure 7.3 (Color Online) The probability (**a**) and number of accessible atoms (**b**) associated with each state found in CdS522, CdS894, and CdS1418. While there is not technically an interface region as in the core-shell quantum dots, I present it here for comparison with Figs. 7.4 and 7.5.

having band gaps. This would be good news if the shell structures weren't of the same material. Since the shells are made of the same material I cannot treat the cores as entities separate from the shells, and thus the whole system must have the loss of the band gap associated with the dangling bonds.

Turning now to CdS-ZnS core shell structures, which are the same sizes as were just discussed in the CdS nanoparticles. They have some important differences from the CdS nanoparticles. Fig. 7.4 (a) shows that the band gap of ZB522 is in the same place as in CdS522, but ZB894 and ZB1418 have energy gaps that are at lower energy values than the corresponding CdS nanoparticles. The decrease in energy of the top of the valence band and bottom of the conduction band indicates that these structures, even with the large lattice mismatch, are more energetically stable. Another change that I find when switching from the CdS nanoparticles to the core-shell quantum dots is that on the surface of ZB1418, there appears to be a slight diffusion of states in the -5 eV to -4 eV energy range. This diffusion is reminiscent of the scattering of states through out the band gap.

Fig. 7.4 (b) is concerned with the core-shell structures where the interface between the core and shell has been relaxed using the semi-empirical of Ref. [1]. The relaxation does not seem to play a large role in the electronic structure of SE894 and SE1418, when compared to ZB894 and ZB1418, respectively. In SE522 see a much bigger difference. The band gap region of SE522 is lower in energy, so that it is in the energy range of the band gaps of SE894 and SE1418.

Comparison of the number of accessible atoms in the cores of each quantum dot with the values found in Ch. 3 for ZB264, shows that the number of accessible atoms found in each core is much lower than in ZB264; see Fig. 7.5. The number of accessible atoms in the gap region being near zero within each core (and on the interfaces). In the shell region of each quantum dot there are still apparent gap regions, but the

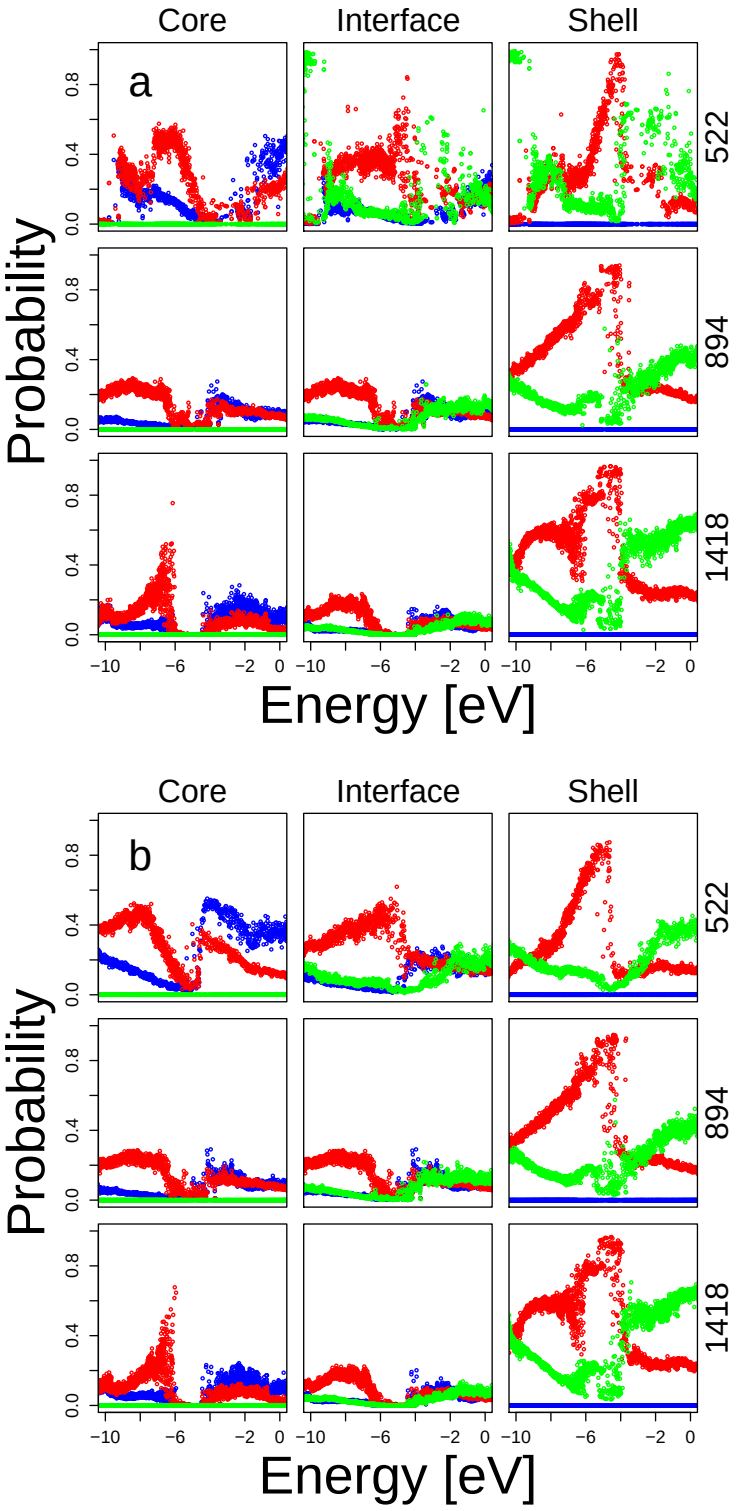


Figure 7.4 Probability. **a.** zinc-blende structures. **b.** Relaxed structures.

many gap states may have as many accessible atoms as do the valence and conduction band regions. I expect that some of these states that are in what I are calling the band gap region are actually not, dangling bond states but valence and conduction band states that have been pushed into the gap region as the number of dangling bonds increased. I saw a similar movement of conduction and valence states for the nanoparticles in Ch. 3.

Further work on this will include determining the TVB and BCB as was done in Ch. 3, studying the degree of charge transfer in the atoms on the core-shell interface, and using the Löwdin population to study if the interface atoms are covalently or ionically bonded.

7.4 Future Direction: Electron-Phonon Coupling in CdS nanoparticles

I ran FIREBALL to compute the Electron-Phonon Coupling (EPC) for nanoparticles with 29 atoms to 200 atoms. Currently, the code fails to consistantly produce results for systems with more than 200 atoms. Analyzing the data did not show any consistent trend in the results. I believe that this is due to the dangling bonds that resided on the surfaces of the nanoparticles, see Fig. 7.6 through 7.8. I believe that if I were to go back and passivate the dangling bonds and then run these simulations again, that a trend would emerge. It is also possible that the sizes of the nanoparticles used here were too small, and that there were not enough atoms that have bulk-like characteristics in order to results that were obviously meaningful.

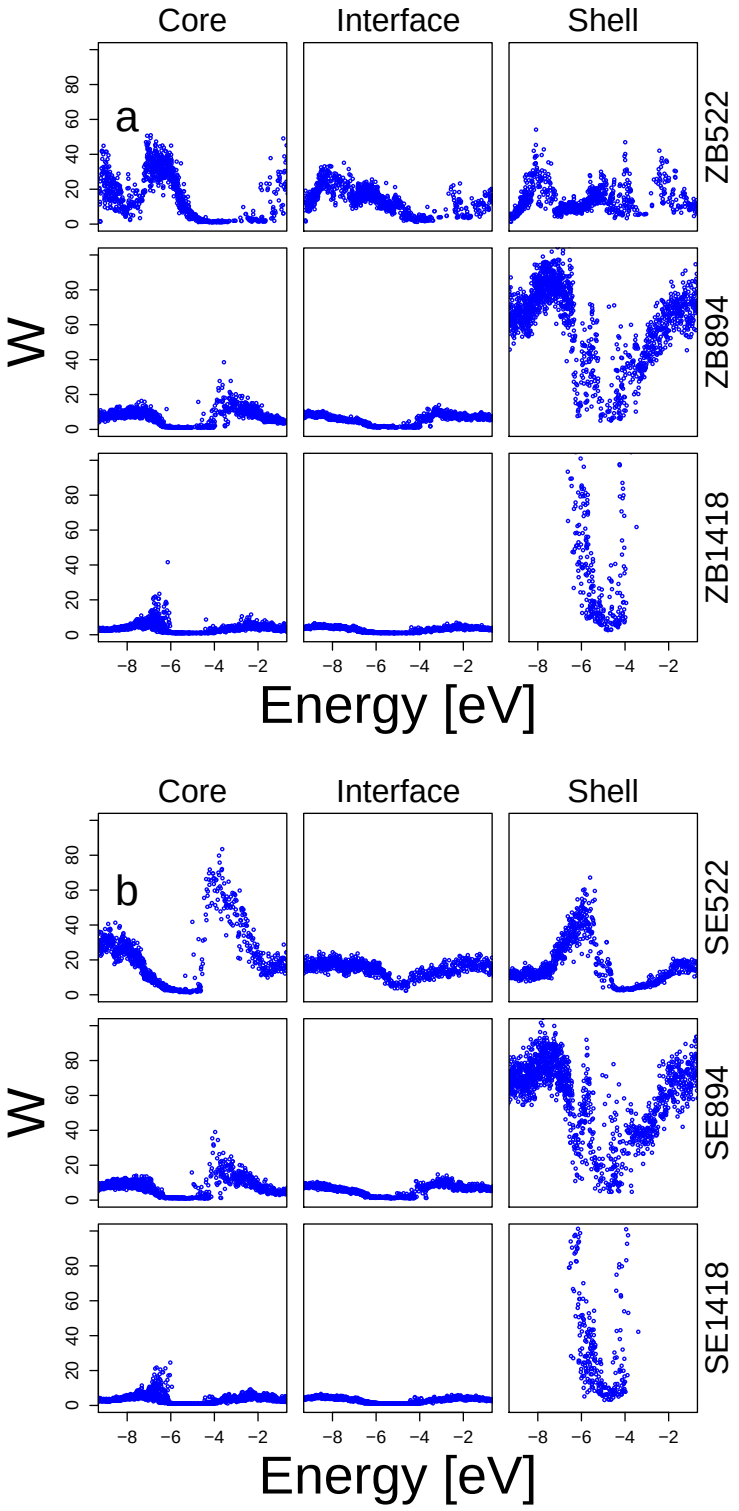


Figure 7.5 Number of accessible atoms, W. **a.** zinc-blende structures. **b.** Relaxed structures.

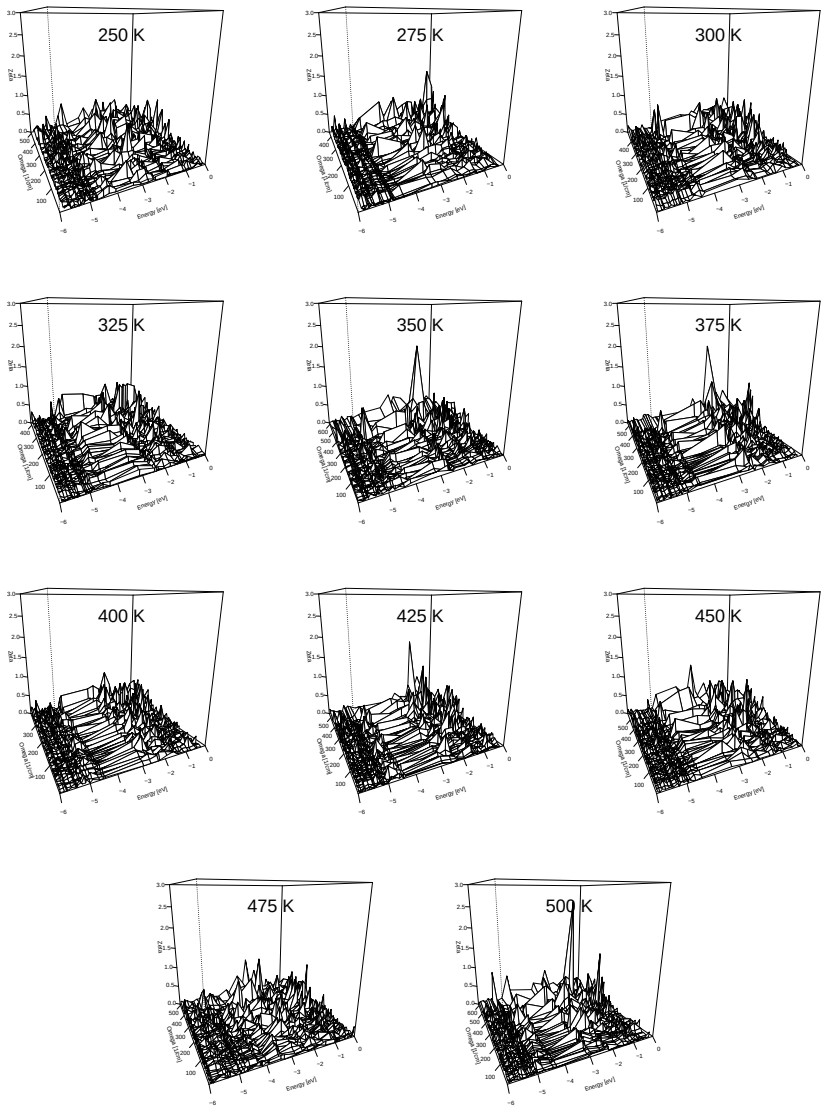


Figure 7.6 Electron-phonon coupling of S-centered nanocrystals.

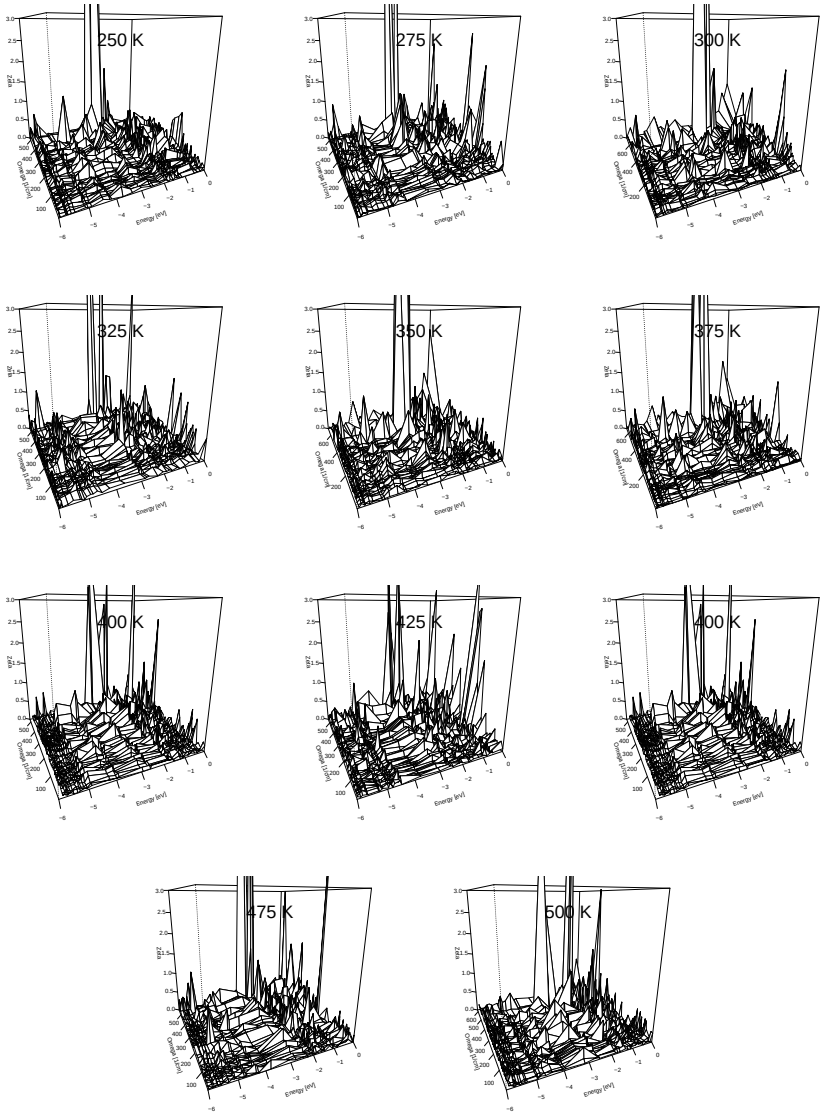


Figure 7.7 Electron-phonon coupling of Cd-centered nanocrystal. Note that in the current figure there is pictures at 400 K and not one at 450 K.

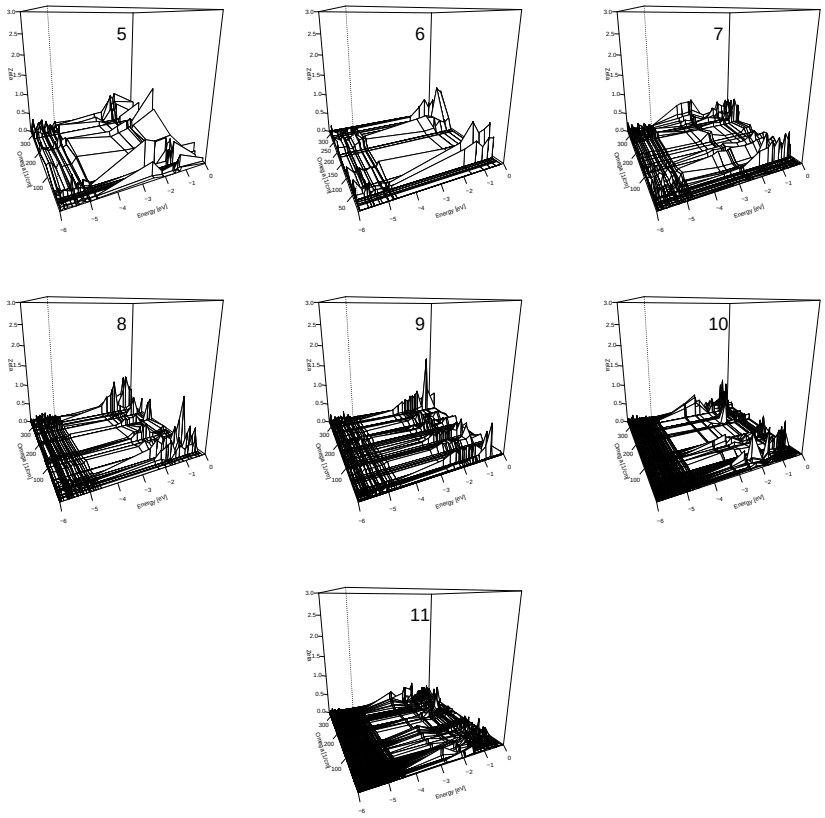


Figure 7.8 Electron-phonon coupling of bond-centered nanocrystals. Note that in the current figure there is pictures at 400 K and not one at 450 K.

Bibliography

[1] Díaz, J. G., Zielinski, M., Jaskolski, W., and Bryant, G. W., Phys. Rev. B **74** (2006) 205309:1.

[2] Bawendi, M. G., Steigerwald, M. L., and Brus, L. E., Ann. Rev. Phys. Chem. **41** (1990) 477.

[3] Ouellette, J., Ind. Phys. **9** (2003) 14.

[4] Bryant, G. W. and Jaskolski, W., Phys. Rev. B **67** (2003) 205320.

[5] Little, R. B., El-Sayed, M. A., Bryant, G. W., and Burke, S. E., J. Chem. Phys. **114** (2001) 1813.

[6] Bryant, G. W. and Jaskolski, W., J. Phys. Chem. B **109** (2005) 19650.

[7] Bandaranayake, R. J., Wen, G. W., Lin, J. Y., Jiang, H. X., and Sorensen, C. M., Applied Physics Letters **67** (1995) 831.

[8] Ricolleau, C., Audinet, L., Gandais, M., and Gacoin, T., Thin Solid Films **336** (1998) 213.

[9] Banerjee, B., Jayakrishnan, R., and Ayyub, P., J. Phys.: Condens. Mat. **12** (2000) 10647.

[10] He, J. et al., J. App. Phys. **95** (2004) 6381.

[11] Murray, C. B., Norris, D. J., and Bawendi, M. G., J. Am. Chem. Soc. **115** (1993) 8706.

[12] Bautista-Hernandez, A., Loaiza-Gonzalez, G., Meza-Montes, L., and Pal, U., Sol. Energy Mater. Sol. Cell **79** (2003) 539.

[13] Tomasulo, A. and Ramakrishna, M. V., J. Med. Chem. **49** (1996) 3612.

[14] Nanda, J., Kuruvilla, B. A., and Sarma, D. D., Phys. Rev. B **59** (1999) 7473.

[15] Troparevsky, M., Kronik, L., and Chelikowsky, J. R., Phys. Rev. B **65** (2001) 033311.

[16] Joswig, J., Sringborg, M., and Seifert, G., J. Phys. Chem. **104** (2000) 2617.

[17] Joswig, J., Seifert, G., Niehaus, T. A., and Sringborg, M., J. Phys. Chem. B **107** (2003) 2897.

[18] Frenzel, J., Joswig, J., Seifert, G., and Sringborg, M., J. Phys. Chem. C **111** (2007) 10761.

[19] Sarkar, P. and Springborg, M., Phys. Rev. B **68** (2003) 235409:1.

[20] Roy, S. and Springborg, M., J. Phys. Chem. **107** (2003) 2771.

[21] Pal, S., Goswami, B., and Sarkar, P., J. Chem. Phys. **123** (2005) 044311:1.

[22] Wen, B. and Melnik, V. N., App. Phys. Let. **92** (2008) 261911.

[23] Pokrant, S. and Whaley, K. B., Eur. Phys. J. D **6** (1999) 255.

[24] Hill, N. A. and Whaley, K. B., J. Chem. Phys. **99** (1993) 3707.

[25] Hill, N. A. and Whaley, K. B., J. Chem. Phys. **100** (1994) 2831.

[26] Yokojima, S., Meier, T., and Mukamel, S., J. Chem. Phys. **106** (1997) 3837.

[27] von Grünberg, H. H., Phys. Rev. B **55** (1997) 2293.

[28] Leung, K., Pokrant, S., and Whaley, K. B., Phys. Rev. B **57** (1998) 235409:1.

[29] Leung, K. and Whaley, K. B., J. Chem. Phys. **110** (1999) 11012.

[30] Troparevsky, M. and Chelikowsky, J. R., J. Chem. Phys. **114** (2001) 943.

[31] Eichkorn, K. and Ahlrichs, R., Chem. Phys. Lett. **288** (1998) 235.

[32] Andersen, K. E., Fong, C. Y., and Pickett, W. E., Journal of Non-Crystalline Solids **299-302** (2002) 1105.

[33] Deglmann, P., Ahlrichs, R., and Tsereteli, K., J. Chem. Phys. **116** (2002) 1585.

[34] Goswami, B., Pal, S., Sarkar, P., Seifert, G., and Springborg, M., Phys. Rev. B **73** (2006) 205312:1.

[35] Wang, L. W. and Zunger, A., Phys. Rev. B **53** (1996) 9579.

[36] Puzder, A., Williamson, A. J., Gygi, F., and Galli, G., Phys. Rev. Lett. **92** (2004) 217401:1.

[37] Yu, M. et al., App. Phys. Let. **88** (2006) 231910.

[38] Keith, J. B., Fennick, J. R., Junkermeier, C. E., Nelson, D. R., and Lewis, J. P., Comput. Phys. Commun. (Accepted).

[39] Junkermeier, C. E., Lewis, J. P., and Bryant, G. W., Phys. Rev. B **77** (2008) 205125.

[40] Junkermeier, C. E., Bryant, G. W., and Lewis, J. P., Phys. Rev. B (Submitted to PRB).

[41] Demkov, A. A., Ortega, J., Sankey, O. F., and Grumbach, M. P., Phys. Rev. B **52** (1995) 1618.

[42] Lewis, J. P. et al., Phys. Rev. B **64** (2001) 195103:1.

[43] Jelínek, P., Wang, H., Lewis, J. P., Sankey, O. F., and Ortega, J., Phys. Rev. B **71** (2005) 235101:1.

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

[45] Foulkes, W. M. C. and Haydock, R., Phys. Rev. B **39** (1989) 12520.

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

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

[48] Sankey, O. F. et al., Int. J. Quant. Chem. **69** (1998) 327.

[49] Finnis, M., J. Phys.: Condens. Mat. **2** (1990) 331.

[50] Polatoglou, H. M. and Methfessel, M., Phys. Rev. B **37** (1988) 10403.

[51] Read, A. and Needs, R., J. Phys.: Condens. Mat. **1** (1989) 7565.

[52] Zaremba, E., J. Phys.: Condens. Mat. **2** (1990) 2479.

[53] Polatoglou, H. M. and Methfessel, M., Phys. Rev. B **41** (1990) 5898.

[54] Robertson, I. and Farid, B., Phys. Rev. Lett. **66** (1991) 3265.

[55] Farid, B., Heine, V., Engel, G. E., and Robertson, I., Phys. Rev. B **48** (1993) 11602.

[56] Martin, R. M., *Electronic Structure: Basic Theory and Practical Methods*, Cambridge University Press, New York, 2004.

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

[58] Perdew, J. P., Phys. Rev. B **33** (1986) 7406.

[59] Lee, C., Yang, W., and Parr, R. G., Phys. Rev. B **37** (1988) 785.

[60] Becke, A. D., Physical Review A: General Physics **38** (1988) 3098.

[61] Perdew, J. P., Jackson, K. A., Pederson, M. R., Singh, D. J., and Fiolhais, C., Phys. Rev. B **46** (1992) 6671.

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

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

[64] Perdew, J. P., Burke, K., and Wang, Y., Phys. Rev. B **54** (1996) 16533.

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

[66] Trouiller, N. and Martins, J. L., Phys. Rev. B **43** (1991) 1993.

[67] Fuchs, M. and Scheffler, M., Comp. Phys. Comm. **119** (1999) 67.

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

[69] Horsfield, A., Phys. Rev. B **56** (1997) 6594.

[70] Hellwege, K. H., Madelung, O., and Landolt-Bönstein, editors, *Numerical Data and Functional Relationships in Science and Technology, Vol 17, Pt. b*, Springer, Berlin, 1982.

[71] Zunger, A. and Freeman, A. J., Phys. Rev. B **17** (1978) 4850.

[72] Cardona, M., Weinstein, M., and Wolff, G. A., MRS Proceedings 2003 **140** (1965) A633.

[73] Pradhan, B., Sharma, A., and Ray, A. K., J. Nanosci. and Nanotech. **5** (2005) 1130.

[74] Lippens, P. E. and Lannoo, M., Phys. Rev. B **39** (1989) 10935.

[75] Sapra, S., Shanthi, N., and Sarma, D. D., Phys. Rev. B **66** (2002) 205202:1.

[76] Lewis, J. P. et al., Phys. Stat. Sol. (b) **233** (2002) 90.

[77] Siegel, A. F., *Statistics and Data Analysis: An Introduction*, John Wiley & Sons, New York, 1988.

[78] Lifshitz, E. et al., Chem. Phys. Lett. **288** (1998) 188.

[79] Lifshitz, E., Dag, I., Litvin, I. D., and Hodes, G., J. Phys. Chem. B **102** (1998) 9245.

[80] Herron, N., Wang, Y., and Eckert, H., J. Am. Chem. Soc. **112** (1990) 1322.

[81] Raji, P., Sanjeeviraja, C., and Ramachandran, K., Cryst. Res. Technol. **39** (2004) 617.

[82] Nose, S., J. Chem. Phys. **81** (1984) 511.

[83] Hoover, W. G., Phys. Rev. A **31** (1985) 1695.

[84] Martyna, G. J., Tuckerman, M. E., Tobias, D. J., and Klein, M. L., Mol. PHys. **87** (1996) 1117.

[85] consortium, T. C., *Car-Parrinello Molecular Dynamics: An ab initio Electronic Structure and Molecular Dynamics Program*, MPI für Festkörperforschung and IBM Research Laboratory, Zürich, 2008.

[86] Elliott, S. R., *Physics of amorphous materials*, Longman Scientific & Technical, New York, 1990.

[87] Vorokh, A. S. and Rempel, A. A., Physics of the Solid State **49** (2007) 148.

[88] Tu, Y., Jacobsson, S. P., and Laaksonen, A., Phys. Rev. B **74** (2006) 205104.

[89] Soler, J. M. et al., J. Phys.: Condens. Mat. **14** (2002) 2745.

[90] Frauenheim, T. et al., Phys. Stat. Sol. **217** (2000) 41.

[91] Truong, T. N. et al., J. Chem. Inf. Model. **46** (2006) 971.

[92] Delley, B., *DMol methodology and applications*, Springer-Verlag New York, New York, 1991.

[93] Chevreau, H., Junkermeier, C. E., Keith, J. B., and Lewis, J. P., Manual: Fireball module, 2005, http://fireball.phys.wvu.edu/LewisGroup/index-fireball.htm.

[94] Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P., *Numerical Recipes*, Cambridge University Press, Cambridge, 1992.

[95] Lewis, J. P., Ordejón, P., and Sankey, O. F., Phys. Rev. B **55** (1997) 6880.

[96] Monkhorst, H. J. and Pack, J. D., Phys. Rev. B **13** (1976) 5188.

[97] Anderson, D. G., J. A.C.M. **12** (1965) 547.

[98] Eyert, V., J. Comp. Phys. **124** (1996) 271.

[99] Freund, J. E. and Simon, G. A., *Modern Elementary Statistics*, Simon & Schuster Company, Englewood Cliffs, New Jersey, 1992.

[100] Davis, J. C., *Statistics and Data Analysis in Geology*, John Wiley & Sons, New York, 1986.

[101] Venables, W. N. and Smith, D. M., *An Introduction to R*, 2006.

[102] Tufte, E. R., *The Visual Display of Quantitative Information*, Graphics Press, Cheshire, Conn., 2001.

[103] Yang, S. H. and Berry, R. J., Mat. Res. Soc. Symp. Proc. **769** (2003) H6.15.1.

Appendix A

xyzSTATS Output File Types

atoms_with_m_nghbrs_extended.txt

Average_Bond_lengths_by_bond.txt

Average_NNN_bond_length_by_time_step.txt

Average_NN_bond_length_by_time_step.txt

Average_angle_by_part.txt

Average_angle_by_step.txt

COM_all.txt

MDvscomparisonatoms01.txt

NNN_Ayres_timestep_SUBSYSTEM.txt

NNN_raw_SUBSYSTEM.txt

NN_Ayres_timestep_SUBSYSTEM.txt

NN_raw_SUBSYSTEM.txt

answer_all2.xyz

diffusion.txt

ave_pos_all.xyz

gr_average.txt

gr_raw_SUBSYSTEM.xyz.txt

logfile.txt

neighbor_info.txt

rad_dis_MD.txt

rad_xi_squared.txt

Index

relaxation, 2, 22, 27, 29, 30, 33
FIREBALL, 3–6, 11, 12, 37, 40, 54, 56–59, 61–66, 69, 91, 93, 106
FIREBALLUI, 3, 54, 56, 59, 62–66, 68, 69, 71
FIREBALL, 53, 67
FIREBALLUI, 55, 63, 67, 68, 70

band gap, 1, 4, 11, 12, 17, 18, 20–22, 25, 27, 29, 30, 32–35
binned data, 5, 14, 41, 42, 76
    histogram, 14, 16, 41, 48, 75, 76, 79, 88, 89
bottom of the conduction band (BCB), 18, 21, 25, 27–31, 33, 35, 101, 106
box plot, 42, 45
bulk probability, 13, 25, 27, 29, 31, 32, 34, 35, 98, 99, 101

capping layer, 2, 25
charge density, 5, 6, 61
Chi-squared (ξ^2), 14, 16, 44, 46, 92
Chi-squared (ξ^2), 5, 14, 15, 42, 45, 49, 76, 87, 91
cutoff radius, 7, 10, 68

dangling bonds, 4, 17, 18, 20–22, 25, 27, 28, 35, 95, 97–101, 104, 106
density functional theory, 5, 10, 11, 37, 53
DOGS, 8, 18–21, 33

enthalpy, 3
exchange-correlation, 9, 10, 57

Fdata, 56–59, 62, 65–69
    integral tables, 10

generalized gradient approximation (GGA), 9, 58
goodness-of-fit, 14
ground state, 3, 6, 8, 12, 96

Harris-Foulkes, 6–8, 11, 17, 19–22, 28
Hartree, 6, 8, 9
heavy hole, 11
high-throughput, 66, 68
highest occupied molecular orbital (HOMO), 20, 21, 32
histogram, 14, 16, 41, 48, 75, 76, 79, 88, 89
    binned data, 5, 14, 41, 42, 76

Kohn-Sham, 6–8

lattice constant, 11
light hole, 12
local minima, 3
local-density approximation (LDA), 9, 57
lowest unoccupied molecular orbital (LUMO), 20, 21, 32, 33

mean, 3, 8, 12, 14, 15, 41, 76
molecular dynamics, 3, 5, 14, 15, 22, 37–49, 58, 59, 64, 67, 69, 70, 73–75, 77–81, 87, 89–92, 96

nanocrystal
    relaxed, 2, 4, 20–31, 33–35
nanocrystal
    relaxed, 121
nanoheterostructure, 2, 79, 99, 103, 104
    capping layer, 2, 25
    quantum dot, 2, 79, 99, 103, 104
    quantum-dot quantum-well, 2

nanoparticle, 1–4, 12, 14–17, 23, 24, 37, 38, 40–42, 44–46, 48–51, 79, 89, 90, 95–99, 101, 104, 106
  nanocrystal, 1, 13, 17, 20–27, 29, 30, 32–34, 37, 38, 51, 67, 69, 70, 89, 90, 108–110
  relaxed, 38, 40, 41
nearest neighbor, 13, 23, 25, 37, 45, 47–49, 80–82, 88, 93, 99
next-nearest neighbor, 38, 45, 47–49, 80, 88, 89, 93
Nosé-Hoover, 40, 59
Number of accessible atoms (W), 5, 13, 33, 34, 99, 101, 103, 104

occupation number, 7–9
orbitals, 1, 5, 7, 8, 11, 20, 21, 27, 54, 57, 62, 68

pseudopotential, 5, 6, 9, 10, 57
  Fuchs and Scheffler, 9
  Hamann, 9, 68
  semilocal, 9
  Troullier-Martins, 9

quantum dot, 2, 79, 99, 103, 104
  capping layer, 2, 25
quantum well, 2

radial distribution, 75, 79, 87, 89, 90
radial distribution function (g(R)), 50, 75, 120
reconstruction, 2, 34
relaxation, 2, 4, 20–26, 28, 30, 31, 33–35, 38, 40, 41, 95, 97, 104, 121

Schrödinger equation, 9
self-consistant field (SCF), 8, 22, 61
stoichiometric, 23, 99

top of the valence band (TVB), 21, 25, 27–30, 32, 33, 35, 99, 101, 106
total energy per atom, 38

variation, 6, 8

wurtzite, 2, 3, 13–15, 37, 38, 40–42, 44–51, 82, 90–92

xyzSTATS, 4, 74, 76–78, 81–84, 87, 90, 91, 93, 96, 97

zinc-blende, 2, 3, 11, 12, 14, 17, 18, 22–27, 29, 30, 32–34, 38, 40–42, 45, 47–50, 82, 89–91, 101, 102, 105, 107