THE REAL SPACE FINITE ELEMENT HARTREE-FOCK METHOD AND THE THERMO-MECHANICAL PROPERTIES OF CARBON NANOTUBES

BY
ROUHOLLA ALIZADEGAN

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Theoretical and Applied Mechanics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2011

Urbana, Illinois

Doctoral Committee:

Professor K. Jimmy Hsia, Chair, Director of Research
Professor Iwona M. Jasiuk
Professor Nancy R. Sottos
Assistant Professor Eric Pop

ABSTRACT

This thesis consists of two parts. The first part aims to explore the application of the popular method of the finite element method (FEM) in the electronic structure theory. The finite element method is a very general numerical technique in mathematics for solving partial differential equations (PDEs) and it has been widely applied in computational mechanics and engineering in general, but it has not been extensively used in science for electronic structure calculations. Currently most electronic structure calculations rely on well-established and fast basis-set alternatives. However, there are serious shortcomings with the standard global basis-set methods such as basis saturation and ill-conditioning of the matrices as the basis-set size is increased. In this dissertation we exploit new strategies that rely on the divide-and-conquer (DC) as well as the enriched/generalized FEM (GFEM) and face-based smoothed FEM (FS-FEM) methods to solve the electronic structure problems. The linear-scaling DC partitioning scheme has been used to scale up the method for larger systems with facile parallelization among many processors utilizing locality assumptions. GFEM and FS-FEM techniques have been proposed to deal with the inner core singularity and to improve the quality of the solutions without considerable added computational cost. While these results are highly encouraging, still more research needs to be conducted in order to be able to decisively determine the best method of tackling the numerical solution of the electronic

structure of atoms and molecules. Based on these preliminary results, it is anticipated that yet more elegant hybrid techniques may exist.

In the second part of the thesis, special attention has been paid to carbon nanotubes (CNTs) and their thermo-electro-mechanical properties. Application of CNTs and other carbon-based materials such as graphene in science and technology has been constantly on the rise in the past two decades for example as wires, switches, transistors or other nano-electro-mechanical systems (NEMS) and nanostructures. Here, several of the more fundamental mechanical, chemical, heat transport and thermal properties of the CNTs for these applications and for microscopy purposes (in particular, atomic force microscopy or AFM) have been computationally as well as experimentally studied. Properties such as stability and collapse propagation in CNTs, dispersibility and thermal coupling to the substrate have been the focus of attention. The origins of the difficulty of the dispersion of CNT solutions have been explained and quantitative suggestions have been made to solve this problem. The thermal footprint of CNTs on SiO_2 substrate has been extracted to predict the thermal conductance from CNT to SiO_2. AFM tip-CNT interactions have been thoroughly investigated and recommendations for the correct interpretation of AFM images of individual CNTs have been given. Energetics of collapse and inflation of CNTs on SiO_2 have been studied and upper-bound estimates for the collapse/inflation propagation speeds have been obtained. These studies provide some computational tools and rather in-depth theoretical insight into the mechanisms at play at the nano-scale and should lead to a better understanding for the design and analysis of future carbon-based nanodevices and nanostructures.

ACKNOWLEDGMENTS

The author wishes to express sincere appreciation to his advisor, Professor Jimmy Hsia, for his patience and support in conducting this multidisciplinary research in fundamental science and engineering with unique and sometimes exceptional challenges. Special thanks to Professor Todd Martinez (Stanford University) and Professor Eric Pop (University of Illinois) whose collaboration and constructive comments added valuable insight and inspired excellence. Professor Nancy Sottos is hereby gratefully acknowledged for providing TA support and sharing valuable experience and wisdom during the course of completion of this degree. The author would like to acknowledge the generous help of Dr. Ting Zhu (Georgia Institute of Technology) and Dr. Sulin Zhang (Pennsylvania State University) for allowing him to use their computational facilities. Some of the simulations in this thesis were carried out on the computers at the National Center for Supercomputing Applications (NCSA) at the University of Illinois at Urbana-Champaign (UIUC) which is hereby recognized and appreciated. In addition, the author would like to acknowledge the financial support from the National Science Foundation (NSF) and also from the Department of Mechanical Science and Engineering (MechSE) at the University of Illinois. Last but not least, he would like to express his heartfelt gratitude to his parents whom without their continuous and nonstop support and encouragement, this work would have never ended.

To my beloved parents,

Mohammadtaghi and Hoorieh

who brought me to this world and taught me how to soar.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

In 1929, British physicist Paul A. M. Dirac stated that "The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known…" [1]. Even though now we know that this statement is an over simplification of the laws of physics, it fairly accurately applies to that part of the laws of physics that explain the day-to-day physical and chemical phenomena. Thus, according to this statement, it only remains to find methods to enable us to apply these physical laws to basically every aspect of the physical world, from atoms and molecules to wires and transistors and even the living cells. Many new discoveries in science were made possible because of advances in technologies. Physics and chemistry are no exception. Recently, new understandings in physics and chemistry, and the mathematical theories or models of these discoveries are aided by, among others, the availability of advanced computational tools. The developments in simulation and computational tools have had tremendous impact on the advances of science. Electronic structure calculations with the advent of efficient techniques and powerful computers is one example of such impact.

Electronic structure calculations have been one the most computationally intensive computations ever carried out by researchers. Millions of lines of code have been written by different research groups around the world in different quantum chemistry and physics software packages and hundreds of millions of CPU hours have been spent solving them. However, finding methods that meet both the challenge of reliable physics vs. reasonable computation time is an open question. There are three basic *ab initio* methods to calculate electronic structures: the plane wave (PW) and grid methods, localized atomic(-like) orbitals method, and atomic sphere method [2]. The finite-element method (FEM) belongs to the first category. In the FEM, the *shape functions* are strictly local piecewise polynomials in real-space. The method is completely general and its convergence is guaranteed by variational principles, and thus can be controlled systematically. The FEM method in theory can achieve a highly variable resolution in real-space via p-adaptivity (i.e. using higher-order shape functions where needed) and h-adaptivity (i.e. locally using finer elements), and it is well-suited for parallel computing implementations. The method thus combines all the major advantages of both the grid-based and basis-oriented approaches and is particularly promising for large-scale, accurate *ab initio* calculations [3]. These methods can be implemented with either k-space formulation [3-7] or the real-space-grid formulation. The real-space viewpoint to calculating the electronic properties is more suited for the cases which are not regular lattices, such as in the case of deformed carbon nanotubes (CNTs) and perhaps biological macromolecules.

There is a myriad of electronic structure techniques available in the literature which each one has its respective advantages and disadvantages, power points and weak points. The following list is not a comprehensive list of the different electronic structure methods by any means:

- Tight-Binding method (TB or LCAO)
- Hückel/Extended Hückel
- Hartree-Fock method (HF)
- Configuration Interaction method (CI)
- Coupled-Cluster method (CC)
- Møller-Plesset perturbation theory (MP)
- Density Functional Theory (DFT)
    - Local Density Approximation (LDA)
    - Generalized Gradient Approximation (GGA)
    - Orbital-Free Density-Functional Theory (OFDFT)
- Quantum Monte Carlo (QMC)
- Multireference Configuration Interaction method (MR-CI)
- Etc.

The semi-empirical tight-binding (TB) approach which is probably the simplest of these methods, is also referred to as the linear combination of atomic orbitals method (LCAO). It is considered semi-empirical because a number of free parameters for a given material are fit to either experimental data or first principles calculations in order to obtain a realistic model. It is based on the assumption that the total cohesive energy of

a crystalline material can be written as the sum of two contributions: one due to the bonding valence electrons associated with each atom, and the other due to repulsive sources including ion-ion interactions. The total cohesive energy per atom for an arrangement of N atoms is given by: [8]

$$E_{tot} = \frac{1}{N}\left(\sum_{j=1}^{N} E_j^{elec} + \frac{1}{2}\sum_{i=1}^{N}\sum_{k=1}^{N} \phi^{rep}(r_{ik}) \right), \tag{1.1}$$

where the first term is the electronic part of the energy, and the second term is the repulsive part of the energy. The repulsive energy is usually written as a simple pair potential term accounting for interactions between all atoms in the arrangement as a function of their separation distance. The tight-binding method uses the bonding electron orbitals associated with each of the atoms for example in a semiconductor material to represent the energetics of the structure. The eigenstates of the Hamiltonian of the material are written in an atomic-like basis-set (Figure 1.1) and the exact many-body Hamiltonian operator is replaced with a parameterized Hamiltonian matrix. For molecular structures, TB provides the energy states corresponding to different bonding and anti-bonding electron orbital states. These energy states can be used to determine the electrical properties of the structure. When combined with a repulsive empirical pair-wise potential, the electronic states can sometimes be used to compute the mechanical properties of the structure as well.

The tight-binding method of modeling materials lies between the very accurate, very expensive *ab initio* methods and the fast but limited empirical methods such as the

Terssoff-Brenner potential. [9-11] When compared with other *ab initio* methods, tight-binding is typically two to three orders of magnitude faster, but suffers from a reduction in transferability due to the approximations made, and when compared with empirical methods, TB is two to three orders of magnitude slower, but the quantum mechanical nature of bonding is retained, ensuring that the true nature of bonding may correctly be described for example in non-equilibrium structures.

The density functional theory (DFT) [12] is another frequently used method especially in the solid-state physics community to calculate the electronic structure of molecules because of its relatively low computational cost. However, DFT methods fail to describe non-covalent interaction energies and reaction barriers because they suffer from self-interaction problems and do not incorporate long-range correlation effects. Second-order Møller–Plesset perturbation (MP2) theory [13] is the simplest method to account for electron correlation at an *ab initio* level. However, the computational cost of MP2 calculations is considerably higher and its practical applications are limited to molecules of moderate size and therefore it is not very popular. The most accurate quantum chemistry techniques available today are the full configuration-interaction and quantum Monte-Carlo which are prohibitively expensive and their application is limited to a few atoms at the present time.

Among all electronic structure calculation methods, the quantum chemistry Hartree-Fock (HF) *ab initio* method is very widely studied and mature. It has been successfully used to study the chemical and physical properties of a wide range of atoms, molecules, and compounds. HF directly works with the electronic wavefunctions

of the different orbitals whereas DFT works with electronic charge density as the basic variable. In DFT method, while the Hohenberg-Kohn theorem guarantees the existence of a universal functional of the ground state electron charge density distribution, $n_0(\mathbf{r})$, that defines the total energy of the electronic system and all the other properties of the ground state, it doesn't tell anything specific about this universal functional form [14-15]. Although there are exact solutions for the free electrons case, in general these unique functional relations are not known and therefore for general classes of problems they should be developed by laborious empirical or heuristic methods. On the other hand, HF is essentially free from any assumptions on the form of the density functional and, although it may be computationally more demanding for being an orbital-based method, it includes an exact expression for the electron exchange which is usually the most important portion of the exchange-correlation (XC) energy. There has been efforts to incorporate this unique feature of the Hartree-Fock theory to DFT by using *hybrid functionals* [16-18]. Hybrid functionals such as PBE0 [19] have proven to be superior to GGAs in DFT. The accuracy of HF technique is between the crude Hückel model [20] which neglects electron repulsions and the very accurate but very expensive quantum chemistry methods such as quantum Monte Carlo (QMC). The HF technique can be viewed as the first level variational expansion of the fermionic many-body Schrödinger equation into an anti-symmetrized product of one-electron wavefunctions using Slater determinants [21]. There are many post-HF methods that systematically build upon the accuracy of the HF technique, among them configuration interaction (CI) method [22],

the coupled-cluster (CC) method [23-24] and Møller-Plesset perturbation theory (MP) [13] can be mentioned. Figure 1.2 compares the accuracy vs. the typical size of the problems that these different methods can currently solve.

In this thesis the HF has been chosen for the quantum chemical calculations for the reasons mentioned above. An effort has been made to explore the application of the popular method of the Finite Element Method (FEM) in the electronic structure theory (Chapter 2). The Finite Element Method is a very general numerical technique in mathematics for solving partial differential equations (PDEs) and it has had a tremendous amount of success in computational mechanics and engineering in general, but its application in electronic structure calculations has been rather limited. One of the reasons for this low popularity rate in quantum science is the availability of well-established and fast basis-set alternatives that make competition very tough in this field. Also, there are certain challenges with regard to the treatment of nuclear singularity in FEM methods which make their use rather time-consuming and inefficient. However, there are serious shortcomings with the standard global basis set methods such as basis saturation and ill-conditioning of the matrices as the basis set size is increased. These challenges must be resolved before electronic structure calculations can be reliably and prevalently used for the accurate solution of arbitrarily large systems. Therefore, many researchers around the world are still trying to address these outstanding issues and come up with methods that alleviate some of these problems. The divide-and-conquer (DC) as well as the enriched FEM and face-based smoothed FEM (FS-FEM) methods developed and discussed in detail in this work (Chapter 3) are steps in that direction.

Chapter 4 focuses on a FEM-based implementation of the non-equilibrium Green's function (NEGF) method that in theory could be used for accurate modeling of quantum transport phenomena in molecular systems. However, due to a very high computational demand, its applicability is rather limited.

In the second part of this thesis (i.e. Chapters 5-8) an especial attention has been paid to carbon nanotubes (CNTs) as the original motivation of this work was a project on the conductivity and electro-mechanical couplings in carbon nanotubes which later on led to a detailed study of the electronic structure of other molecular systems. Carbon nanotubes are currently used in industry only in nano-composites and some other limited applications, but due to their extraordinary mechanical and electrical properties, they are highly regarded as one of the prospective building-blocks of future nano-electro-mechanical systems (NEMS) and devices. As an evidence, the research on CNTs and other carbon based materials such as graphene in science and technology has been constantly on the rise in the past two decades for example as wires, switches, transistors or other nano-electronic devices and structures. Here, several of the more fundamental mechanical, chemical, transport and thermal properties of the CNTs for these applications, as well as for microscopy purposes have been computationally as well as experimentally studied. Properties of CNTs such as dispersibility (Chapter 5), thermal coupling to the substrate (Chapter 6), atomic force microscopy (AFM) characterization (Chapter 7) and stability and collapse propagation (Chapter 8) have been the focus of attention. The origins of the difficulty of the dispersion of CNT solutions have been explained and quantitative suggestions have been made to solve this problem. The

thermal footprint of CNTs on SiO_2 substrate has been extracted to predict the thermal conductance from the CNT to SiO_2. AFM tip-CNT interactions have been thoroughly studied and recommendations for the correct interpretation of AFM images of individual CNTs (both for single-walled and double-walled CNTs) have been given. Energetics of collapse and inflation of CNTs on SiO_2 substrate have been studied and upper-bound estimates for the collapse/inflation propagation speeds have been obtained. These studies provide some computational tools and rather in-depth theoretical insight into the mechanisms at play at the nano-scale and should lead to a better understanding for the design and analysis of future carbon-based nanodevices and nanostructures. These distinct studies are presented as separate chapters in this thesis as mentioned above.

1.1. Figures and tables

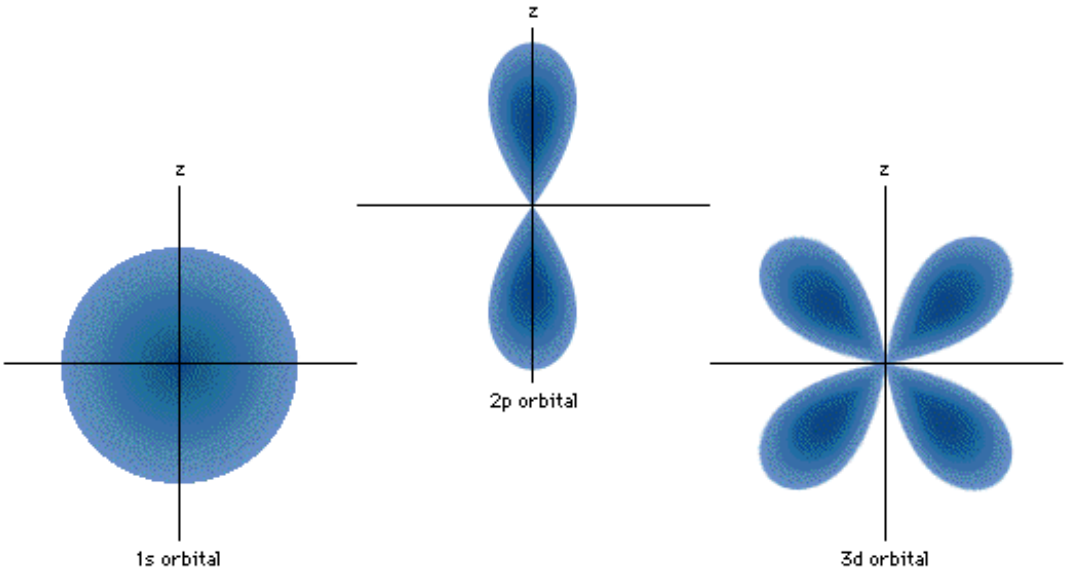


Figure 1.1. Schematic of the atomic orbitals used in the tight-binding approach.

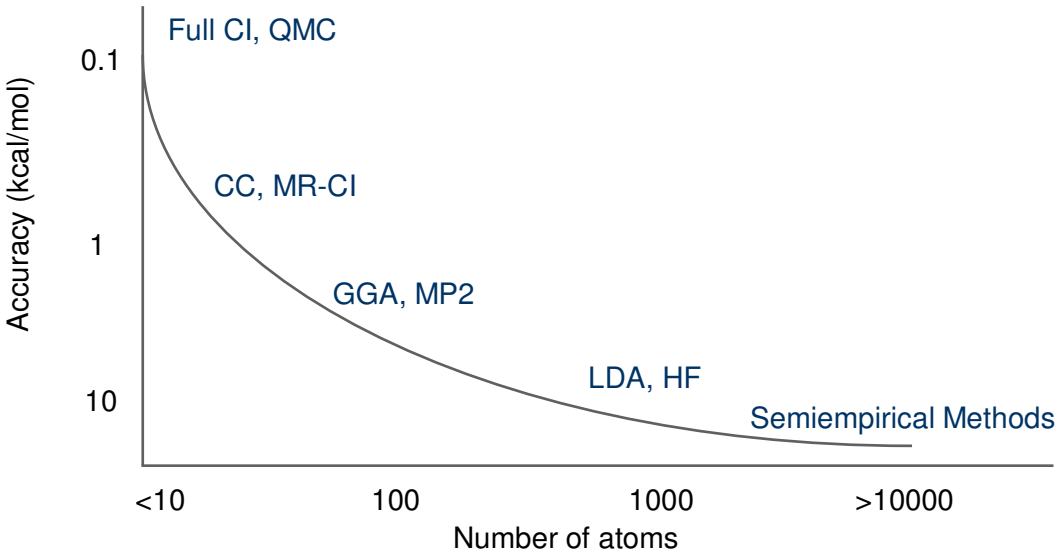


Figure 1.2. Comparison of the accuracy of the different typical electronic structure calculations.

CHAPTER 2: ELECTRONIC STRUCTURE CALCULATIONS USING THE FEM-HF METHOD

2.1. Introduction

Much computational effort has been devoted to the solution of the electronic Schrödinger equation for molecular systems, as this yields the potential energy surfaces that form the basis for further dynamical and/or statistical mechanical calculations. Most common methods for solving the electronic Schrödinger equation start from a one-electron basis set composed of smooth nonorthogonal functions, generally atom-centered Gaussians. This choice has the advantage of allowing one to encapsulate much of the important physics of the solutions in the basis set. For example, the cusp at the nuclei is approximated by contraction of many Gaussian functions and the basis functions are chosen to provide a good representation of the solutions for free atoms. The primary disadvantage of this approach is that it is awkward to carry these calculations to convergence since the nonorthogonal nature of the basis functions leads to ill-conditioned matrices as the basis set size is increased. Grid-based methods provide an alternative approach wherein convergence can be achieved more systematically. Examples include wavelet, finite element (FE), and finite difference (FD) methods [25].

The basic formulation of the finite element method (FEM) and its advantages are summarized by Pask, *et al.* [3] which can be used to obtain solution with prescribed resolution in real-space. These real-space methods can be implemented either with Bloch-periodic boundary conditions in k-space [3-7] or with nonperiodic wavefunctions. The latter viewpoint to calculating the electronic properties is best suited for nonperiodic cases such as bent or squashed carbon nanotubes or isolated molecules.

The Hartree-Fock (HF) method forms the starting point for most *ab initio* wavefunction-based approaches to solve the electronic Schrödinger equation. It also encompasses most of the computational steps involved in density functional theory (DFT) approaches, since the Kohn-Sham (KS) equations differ from the HF equations only in the presence of an additional one-electron potential derived from the exchange-correlation functional K_{xc}. Within the local density (LDA) and generalized gradient (GGA) approximations, the KS equations do not require the calculation of the exact exchange that is an important component of HF. However, most modern functionals are hybrids that do include some admixture of exact exchange. [16-17] The computational effort for solving the KS equations with hybrid functionals is nearly the same as that for the HF equations, with the added need for numerical integration of the exchange-correlation contribution. In this paper, we focus on the HF method, but emphasize that extension to KS approaches is conceptually straightforward.

There have been a number of previous efforts aimed at using FEM in conjunction with HF or DFT methods. Flores and coworkers [26-27] used higher order polynomials as the basis functions in the solution of the two dimensional atomic HF equations and

called it the p-version of the finite-element method (p-FEM). They showed that the method works well for some representative atoms. Heinemann, *et al.* reported highly accurate calculations for light atoms and diatomic molecules with up to 6^{th} order Lagrange polynomials. [28-30] Taking advantage of the axial symmetry of these atoms and molecules, they used prolate spheroidal coordinates that dramatically reduced the computational cost. However, this technique severely limited the range of the problems one could solve. In separate work, [31] they studied the effect of spin-polarization using the Hartree-Fock-Slater method. Lavor, *et al.* [32] proposed a global optimization method with nonconvex polynomials which could be considered as a variant of the finite-element approach. In a series of papers [33-39] Sundholm, *et al.* studied the application of the finite-element multiconfiguration Hartree-Fock calculations to different atoms and chemical compounds. Ackermann and Roitzsch [40-41] used a multigrid adaptive finite element method in cylindrical coordinates for the solution of some symmetric single-electron systems. Quiney, *et al.* [42] studied the relativistic Dirac equation in the algebraic approximation and compared the finite basis set and finite element molecular Dirac-Hartree-Fock calculations. White and coworkers [43] studied the application of second-order orthonormal shape functions with a uniform grid to some one and two-electron systems. Tsuchida and coworkers in a series of publications [44-47] reported adaptive higher-continuity elements within the pseudopotential DFT-LDA framework and applied their method to simulate periodic systems up to 512 atoms. It is noted that they used transformation in curvilinear coordinates rather than Cartesian coordinates as used by others. Pask and coworkers [3-

6] provided a nice and up-to-date survey of FEM methods for electronic structure of periodic systems. They used C^0–continuity cubic serendipity elements on a uniform grid within the pseudopotential DFT-GGA framework.

In this chapter, we describe a real-space finite element implementation of the HF method for polyatomic molecules. We introduce a new, accelerated calculation method of the exact exchange by an auxiliary local exchange fitting related to the X-α approximation. Standard C^0–continuity eight-node hexahedral and four-node tetrahedral elements have been used in a fully adaptive meshing scheme without symmetry assumptions. Using locality concepts, we apply a divide-and-conquer method which enables facile parallelization and should lead to reduced scaling for the electronic structure of more complex systems such as carbon nanotubes (CNTs). In section 2 we present the finite element formulation of the Hartree-Fock self-consistent field method along with the details of the implementation of the boundary conditions (BCs). In section 3 we present some representative results and discussion. Section 4 contains summary and conclusions.

2.2. The self-consistent field (SCF) method

In the Hartree-Fock self-consistent field theory, the Hamiltonian matrix, **H**, is formally replaced by the Fock matrix, **F**, which is the effective one-electron Hamiltonian. Here a brief derivation of this method is included for the sake of completeness. We start from the time-independent Schrödinger equation:

$$\mathbf{H}\left|\mathbf{\Phi}\right\rangle = E\left|\mathbf{\Phi}\right\rangle, \tag{2.1}$$

where $\left|\mathbf{\Phi}\right\rangle$ are the wavefunctions, E is the energy and

$$\mathbf{H} = -\sum_{i=1}^{N}\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M}\frac{1}{2M_A}\nabla_A^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{r_{AB}}. \tag{2.2}$$

In this equation, N is the total number of electrons, M is the number of atoms, Z are atomic numbers and r are distance between two particles. The first sigma refers to the kinetic energy of the electrons, the second sigma refers to the kinetic energy of the nuclei, the third term refers to the nuclear-electronic attraction, the fourth term refers to the electron-electron repulsion and the last term refers to the nuclear-nuclear repulsion. The wavefunction is a collection of spin-orbitals which are orthonormal and incorporate the spin state as well as the spatial state and obey the Pauli Exclusion Principle which states that the total wavefunction for two identical fermions (particles with half-integer spin) is anti-symmetric. One way to make sure that the wavefunctions satisfy the Pauli Exclusion Principle is to write the overall wavefunction as a Slater determinant:

$$\Psi = \frac{1}{\sqrt{N!}}\begin{vmatrix} \psi_1(\mathbf{r}_1) & \overline{\psi_1(\mathbf{r}_1)} & \cdots & \psi_k(\mathbf{r}_1) & \overline{\psi_k(\mathbf{r}_1)} \\ \psi_1(\mathbf{r}_2) & \overline{\psi_1(\mathbf{r}_2)} & \cdots & \psi_k(\mathbf{r}_2) & \overline{\psi_k(\mathbf{r}_2)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \psi_1(\mathbf{r}_N) & \overline{\psi_1(\mathbf{r}_N)} & \cdots & \psi_k(\mathbf{r}_N) & \overline{\psi_k(\mathbf{r}_N)} \end{vmatrix}, \tag{2.3}$$

where N is again the total number of electrons. The single-particle spin-orbitals are orthonormal and incorporate the spin states as well as the spatial states. The spin-orbitals with overbars correspond to the β spin state and the ones without overbars correspond

to the α spin state. The procedure usually followed is the restricted Hartree-Fock (RHF) approach although restricted open-shell Hartree-Fock (ROHF) and unrestricted Hartree-Fock (UHF) approaches could be implemented in a similar fashion [48]. In the RHF approach, the molecular orbitals (MOs) occur in pairs in which the same given spatial orbital appears twice in the determinant with different spin factors, whereas in UHF approach different spatial orbitals are assigned to different spins. For large systems, there are always more electrons in pairs with opposite spin compared to the partially occupied orbitals. Here we assume that the Aufbau principle [22] holds valid, i.e. we have implicitly made the assumption that the ground state single-determinant wavefunction is the one obtained by occupying the lowest-energy solutions of the Hartree-Fock equations. Of course for more complex systems with multiply degenerate highest occupied molecular orbitals (HOMOs), this assumption loses its plausibility. An orbital obtained as a solution from SCF calculations whose energies are higher than those of HOMO is called a virtual orbital. The first one of such orbitals is the lowest unoccupied molecular orbital (LUMO). These orbitals are not variationally correct approximations to the excited state orbitals.

Having this form, the ground state energy functional is obtained by:

$$E[\Psi_0] = \langle \Psi_0 | \mathbf{H} | \Psi_0 \rangle . \tag{2.4}$$

Taking the variation of this functional,

$$
\begin{aligned}
E[\Psi_0 + \delta\Psi] &= \langle \Psi_0 + \delta\Psi | \mathbf{H} | \Psi_0 + \delta\Psi \rangle \\
&= E[\Psi_0] + \langle \delta\Psi | \mathbf{H} | \Psi_0 \rangle + \langle \Psi_0 | \mathbf{H} | \delta\Psi \rangle + \cdots \\
&= E[\Psi_0] + \delta E + \cdots
\end{aligned}
\tag{2.5}
$$

where δE is called the first variation in E. In order for E to be a stationary point with respect to variations in Ψ_0, it's necessary and sufficient to have $\delta E = 0$. There is also a normalization requirement on Ψ_0 that originates from charge conservation and is described as:

$$
\langle \Psi_0 | \Psi_0 \rangle = 1 . \tag{2.6}
$$

In basis-set techniques, each orbital is represented as a linear combination of basis functions with unknown coefficients. The problem of calculating the wavefunctions is hence equivalent to computing these coefficients. Thus, each orbital is expanded in the form

$$
\psi_k = \sum_{\mu=1}^{M} C_{\mu k} \phi_\mu , \tag{2.7}
$$

where ϕ_μ are the basis functions, $C_{\mu k}$ are the unknown coefficients and M is the total number of such functions. Two options for basis sets are atom-centered Slater type orbitals (STO) and Gaussian type orbitals (GTO). STOs are in the general form below

$$
\phi^{STO}(r, \theta, \varphi) = A Y_{l,m}(\theta, \varphi) r^{n-1} \exp(-\alpha r) , \tag{2.8}
$$

where A is a normalization constant and $Y_{l,m}(\theta,\varphi)$ is spherical harmonics in the azimuth and elevation spherical coordinates and l, m and n are the quantum numbers of the orbital. GTOs are in the following form

$$\phi^{GTO}(r,\theta,\varphi) = AY_{l,m}(\theta,\varphi)r^{l}\exp(-\alpha r^{2}). \tag{2.9}$$

Although STOs have the advantage of representing the nuclear-electron cusp correctly, they are no longer used extensively (a notable exception is the ADF program package [49] based on numerical integration) because of the difficulty of evaluating the required two-electron repulsion integrals. The alternative GTO basis leads to integrations which can be carried out analytically, but with undesirable behavior for the small and large values of r. This is compensated by using many GTOs, which in turn leads to a large number of integrals that must be evaluated. [20] Unfortunately, the GTOs are nonorthogonal and the use of many basis functions can lead to linear dependence and ill-conditioning problems which make systematic convergence difficult. It can be shown that for the RHF procedure, the condition $\delta E = 0$ together with the normalization condition enforced through the Lagrange Multipliers method leads to:

$$\hat{f}\psi_k(\mathbf{r}) = \varepsilon_k\psi_k(\mathbf{r}), \tag{2.10}$$

where ε_k are the eigenvalues and ψ_k are the eigenfunctions, $\hat{f}$ is the Fock differential operator defined by,

$$\hat{f} = -\tfrac{1}{2}\nabla^2 + V_0(\mathbf{r}) + V_d(\mathbf{r}) + V_x(\mathbf{r}), \tag{2.11}$$

where V_0 contains the effect of all interactions in the electron potential except those arising from the other electrons, and V_d is the Hartree potential which is the electrostatic potential generated by the charge distribution due to the wavefunction, and V_x is the exchange potential. More detail on these terms will be given in the next section. A subtlety is that the eigenvalues ε_k of the Fock operator are not the energies of single electron orbitals. They are identified as Lagrange multipliers in the derivation of the HF equations as mentioned above. Therefore the sum of the state energies is not the total energy of the atom or molecule, although they are related by:

$$\begin{aligned} E_{tot} &= \tfrac{1}{2}\sum_k [\varepsilon_k + < \psi_k \mid \hat{h} \mid \psi_k >] + \tfrac{1}{2}\sum_{I,J\neq I} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \\ &= \sum_k [\varepsilon_k - \tfrac{1}{2} < \psi_k \mid V_d + V_x \mid \psi_k >] + \tfrac{1}{2}\sum_{I,J\neq I} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \end{aligned}, \qquad (2.12)$$

where E_{tot} is the total energy of the system, $\hat{h}$ is the sum of the first two terms in the Fock operator expression and Z_I and Z_J are the atomic numbers of atoms I and J at positions $\mathbf{R}_I$ and $\mathbf{R}_J$, respectively. The operator <…> represents the expectation values in the Dirac bra-and-ket notation. Of course one can still use the Koopman's theorem [22] to calculate approximate ionization potentials and electron affinities. Koopman's theorem simply states that the difference between the sums of the eigenvalues of the HF equation for the ground state and the excited state configuration is equal to the real energy difference between those two states.

2.2.1 The HF-based finite element formulation

In the subsequent sections we delineate the real space FEM approach for the HF equations which promises to make systematic convergence much easier. Several real-space approaches have been used to solve electronic-structure problems such as finite difference [50-61], finite element [3-6, 44-47, 62-63], meshfree [7] and wavelet [64-66] approaches. The finite difference method uses numerical stencils for the individual terms of the differential equation of interest projected on a real-space grid. The accuracy of the solution depends on the grid density. Additionally, there is no associated variational principle in FD methods. Despite these deficiencies, the FD methods [54-59] have been the preferred tool for real-space calculations and, at least with generalized finite differences, allow for variable spatial resolution to handle singular functions such as Coulombic potentials. [3-6] The finite-element method has several advantages over the FD methods by utilizing localized bases called *shape functions*. The shape functions are polynomials that have compact support, i.e., they vanish identically outside some region. Variable spatial resolution in FEM is facilitated by both p-adaptivity (using higher-order shape functions) and h-adaptivity (using finer elements) with a standard recipe. There have been extensive investigations on the generation of the FEM mesh structure. [67] Software packages are readily available for automatic mesh generation for any domain with high quality and performance. [68-69]

An excellent reference on the general FE method exists. [70] Thus, we only briefly outline the FEM in the context of HF here. Our objective is to solve the normalized single-electron equation for spin-orbitals

$$-\tfrac{1}{2}\nabla^2\psi_k + V\psi_k = \varepsilon_k\psi_k \quad \text{in } \Omega, \tag{2.13}$$

where ψ_k and ε_k are the kth molecular orbital and its energy, Ω is the computational domain, and $V = V_0 + V_d + V_x$ is the potential where V_0 contains electron-nuclear attraction and any external potential, V_d is the Hartree potential generated by the charge distribution due to the wavefunction, and V_x is the non-local exchange potential. We adopt atomic units throughout: the unit of distance is the Bohr radius a_0, masses are expressed in the electron mass m_e and the charge is measured in unit charges, e. The Hartree potential V_d can be obtained by solving Poisson's equation:

$$\nabla^2 V_d(\mathbf{r}) = -4\pi n(\mathbf{r}), \tag{2.14}$$

where the density of electrons $n(\mathbf{r})$ is given by

$$n(\mathbf{r}) = \sum_{j=1}^{N} \left|\psi_j(\mathbf{r})\right|^2, \tag{2.15}$$

and where N is the total number of electrons. The exchange potential $V_x(\mathbf{r})$ can be computed using the following relation

$$V_x(\mathbf{r})\psi_k(\mathbf{r}) = -\sum_{j\neq k} \left(V_x\right)_{jk} \psi_j(\mathbf{r}), \tag{2.16a}$$

where

$$\left(V_x\right)_{jk} = \int \psi_j^*(\mathbf{r}')\psi_k(\mathbf{r}')\frac{1}{|\mathbf{r}'-\mathbf{r}|}d^3r' \quad . \tag{2.16b}$$

This is an integrable singularity and can be treated by direct numerical integration as long as the points of singularity are excluded. However, the non-local nature of the exchange operator makes its evaluation very time-consuming and thus we seek an alternative. A statistically averaged exchange potential can be introduced as an approximation to the exact exchange potential:

$$V_x(\mathbf{r}) = -3/2\,\alpha\left[3n(\mathbf{r})/\pi\right]^{1/3}, \tag{2.17}$$

where α is a fitting parameter chosen in such a way to match the exchange energy when calculated from Eqs. (2.16). This approximation, often referred to as the X-α method, was originally envisioned by Slater [71] to reduce the computational demand, and would be used for a fixed value of α, which Slater hoped to be universal. The X-α method was an early form of DFT, but was found not to be sufficiently accurate for a fixed α. In contrast to Slater's original idea, we use this method as a device to minimize computational effort. Specifically, we calculate the exact exchange from Eqs. (2.16) once every 10-15 iterations. An optimal value of α is then determined from such calculations and used for the following 10-15 iterations, when the exact exchange is evaluated once more and a new optimal value for α is determined. Since the calculation of exchange is the most time consuming computational step in our FEM implementation of HF, this leads to considerable savings of computing time. Once the iterative

procedure nears convergence, we compute the exchange contribution without approximation to ensure that the correct solution is obtained.

Taking the inner product of the differential equation in Eq. (2.13) with an arbitrary *test function* ψ' to form an equivalent integral equation, we have

$$\int_{\Omega} \psi'[-\tfrac{1}{2}\nabla^2\psi_k + V\psi_k - \varepsilon_k\psi_k]d\Omega = 0\,, \tag{2.18}$$

and integrating by parts we get

$$\int_{\Omega} \tfrac{1}{2}\nabla\psi'\cdot\nabla\psi_k d\Omega - \int_{\Gamma} \tfrac{1}{2}\psi'\nabla\psi_k\cdot\hat{\mathbf{n}}d\Gamma + \int_{\Omega}\psi'(V_0 + V_d + V_x - \varepsilon_k)\psi_k d\Omega = 0\,, \tag{2.19}$$

where Γ is the boundary and $\hat{\mathbf{n}}$ is the outward unit normal at each point. Assuming natural, Neumann or free boundary conditions (i.e. no external flux) the second term drops out on the surface. We cast this into the weak formulation and perform the standard discretization procedure of the finite elements, i.e. introduce the ansatz:

$$|\psi\rangle \approx \mathbf{N}^T\mathbf{a} = \sum_i N_i(\mathbf{r})a_i\,, \tag{2.20}$$

where the ket vector $|\psi\rangle$ is the quantum state of interest, **N** are the shape functions and **a** are the nodal degrees of freedom (DOFs), corresponding to the coefficients *C* in the standard basis set expansion of Eq. (2.7). In this paper, we use standard C^0–continuity linear isoparametric brick (eight-node hexahedral) and four-node tetrahedral elements for the shape functions. Isoparametric means that the same shape functions are used to specify the relation between the global and local coordinate systems and to describe the variation of the field variables. The shape functions for these elements are shown in

Figure 2.1, which also gives the defining equations for these elements. Although our choice of elements does not yield a high accuracy per degree of freedom, it does give us considerable flexibility in using adaptive meshing strategies. Substituting the wavefunction ansatz into Eq. (2.19), we obtain the following matrix equation for determining the orbitals:

$$\mathbf{H}_k \mathbf{a}_k = \mathbf{v}_k , \quad (k=1,\ldots,N/2), \tag{2.21}$$

where

$$\begin{aligned} \mathbf{H}_k &= \int_\Omega \left\{ \tfrac{1}{2} \nabla \mathbf{N} (\nabla \mathbf{N})^T + (V_0 + V_d - \varepsilon_k) \mathbf{N}\mathbf{N}^T \right\} d\Omega \text{ , and} \\ \mathbf{v}_k &= \sum_{j \neq k} \left\{ \int_\Omega (V_x)_{jk} \mathbf{N}\mathbf{N}^T \, d\Omega \right\} \mathbf{a}_j . \end{aligned} \tag{2.22}$$

Eq. (2.21) can be solved iteratively as a sequence of linear equations with updated estimate of the eigenvalue ε_k in each iteration [72]. We use a bi-conjugate gradient stabilized iterative solver [73] for the linear equations along with the reverse Cuthill-McKee node numbering algorithm, [74] to solve Eq. (2.21) for each orbital k in every iteration. The orbital energies ε_k in Eq. (2.22) are obtained from the Rayleigh-Ritz expression for $|\psi_k>$ from the previous iteration, i.e. $\varepsilon_k = <\psi_k | \hat{F} | \psi_k > / <\psi_k | \psi_k >$ where $\hat{F}$ is the Fock operator. The standard Gram-Schmidt procedure is used to ensure orthonormality of all the orbitals. The Coulomb potential is obtained with the following matrix equation for the Poisson equation (2.14).

$$\mathbf{Lc} = \mathbf{f}\ , \tag{2.23}$$

where

$$\mathbf{L} = \int_{\Omega} \nabla\mathbf{N}(\nabla\mathbf{N})^T \, d\Omega \text{, and}$$
$$\mathbf{f} = 4\pi \int_{\Omega} \mathbf{N} \cdot n d\Omega \,. \tag{2.24}$$

As mentioned earlier, the natural boundary conditions are already incorporated in this derivation. It remains to show how to apply the essential, Dirichlet or value boundary conditions for prescribed values at the boundaries. This, for instance, will be used in the development of the divide-and-conquer technique. First, we rewrite Eq. (2.21) in the partitioned form:

$$\begin{bmatrix} \mathbf{H}_{ff} & \mathbf{H}_{fp} \\ \mathbf{H}_{pf} & \mathbf{H}_{pp} \end{bmatrix} \begin{Bmatrix} \mathbf{a}_f \\ \mathbf{a}_p \end{Bmatrix} = \begin{Bmatrix} \mathbf{v}_f \\ \mathbf{v}_p \end{Bmatrix}, \tag{2.25}$$

where the subscripts *f* and *p* in the sub-matrices mean *free* and *prescribed,* respectively. By straightforward algebraic manipulations it can be shown that this is equivalent to solving the following equations

$$\mathbf{H}_{ff}\mathbf{a}_f = \mathbf{v}_f - \mathbf{H}_{fp}\mathbf{a}_p \text{, and}$$
$$\mathbf{H}_{pf}\mathbf{a}_f + \mathbf{H}_{pp}\mathbf{a}_p = \mathbf{v}_p \,. \tag{2.26}$$

By the proper choice of trial space satisfying the required inhomogeneous Dirichlet boundary conditions and test space satisfying homogeneous boundary conditions in the construction of the weak formulation, the second condition was not

imposed. In the case of homogeneous boundary conditions (i.e. $\mathbf{a}_p = \mathbf{0}$) the first condition simplifies to $\mathbf{H}_{ff}\mathbf{a}_f = \mathbf{v}_f$.

As is well-known, the Fock operator depends on the orbitals $|\psi_k\rangle$. Thus, solution of the Hartree-Fock equation requires a self-consistent iterative approach. The computational procedure used in the current paper is depicted in the flowchart in Figure 2.2. As the first step, a reasonable initial guess is provided. Based on this initial guess, one can calculate the electronic density and solve the Poisson equation, and then calculate the Fock operator and solve for a new set of orbitals and their eigenvalues. The energies can then be calculated. This process is repeated until the values of the calculated energies become stationary. The criterion for convergence has been chosen such that the absolute change in the total energy of the system in the successive iterations to be less than a small tolerance. The direct inversion of iterative subspace (DIIS) method is used to accelerate convergence of this self-consistent field procedure. [75] As mentioned above, we use the statistically-averaged exchange potential for most iterations, updating α every 10-15 iterations, and switching to exact exchange towards the end of the iterative procedure.

2.2.2. The divide-and-conquer method

One way of solving the electronic structure problem for large domains is the local approximation, i.e., assuming that the charge density is not affected by the potential changes far away from the site considered. A number of linear scaling approaches based

on these ideas have been introduced in electronic structure theory. [76] The remaining problem with such approaches is the size of the prefactor associated with the reduced scaling. In many cases, linear scaling is not observed until the system size is very large. The very localized nature of the elements in FEM promises to reduce this prefactor significantly, and we present such an approach here.

In the divide-and-conquer method [77] the domain is partitioned into sub-domains β via partition functions $p^{\beta}(\mathbf{r})$ that have the following property:

$$\sum_{\beta} p^{\beta}(\mathbf{r}) = 1 . \tag{2.27}$$

For instance, in the Hirshfeld-type partitioning [78] of the electronic density in a molecule, each subsystem contains precisely one atom. Then the total density of the system is written as

$$\rho(\mathbf{r}) = \sum_{\beta} \rho^{\beta}(\mathbf{r}) , \tag{2.28}$$

where

$$\rho^{\beta}(\mathbf{r}) = 2 p^{\beta}(\mathbf{r}) \sum_{m} f_{FD}(\varepsilon_m - \mu, T) \left| \psi_m^{\beta}(\mathbf{r}) \right|^2 . \tag{2.29}$$

ψ_m^{β} is the portion of ψ_m that falls within the subdomain β. f_{FD} is the usual Fermi-Dirac distribution function given by

$$f_{FD}(\varepsilon_m - \mu, T) = \frac{1}{1 + e^{(\varepsilon_m - \mu)/k_B T}} . \tag{2.30}$$

T is the temperature and μ is the chemical potential of the system. k_B is the Boltzmann's constant. Charge conservation dictates that

$$N = \int_{\Omega} \rho(\mathbf{r}) d\mathbf{r} \,, \tag{2.31}$$

where as before N is the total number of the electrons.

In this scheme, each sub-domain β^i interacts only with its physically near-by sub-domains or in other words, its immediate neighbors, $\{\beta^j\}$. Rewriting Equation (2.21) in the form

$$\mathbf{H}_k^{ii}\mathbf{a}_k^i = \mathbf{v}_k^i - \sum_j \mathbf{H}_k^{ij}\mathbf{a}_k^j \,, \tag{2.32}$$

for the sub-domain β^i, the global orbitals can be obtained simply by the assembly of the orbitals in each sub-domain. Figure 2.3 shows the general structure of the global Hamiltonian matrix after this partitioning step. The dark gray block in this figure refers to the local Hamiltonian matrix of the corresponding sub-domain, $\mathbf{H}_k^{ii}$. $\sum_j \mathbf{H}_k^{ij}\mathbf{a}_k^j$ can be viewed as the correction to the local Hamiltonian matrix of the sub-domain β^i from its interaction with its neighboring sub-domains. $\mathbf{H}_k^{ij}$ is the interaction matrix between the sub-domains β^i and β^j and is determined in the assembly process of the global Hamiltonian matrix. Each sub-domain can be pertinently handled by one processor using efficient symmetric sparse solvers. These local sub-matrices are then used to get the global response. In other words, we take advantage of the *local* characteristics of the

global Hamiltonian. In order to construct the Hamiltonian matrix in each sub-domain, it is required to take into account the values of the field variables in the adjacent sub-domains. For instance, one can consider a radius of influence around each sub-domain and incorporate everything within that cut-off range in the assembly process. There is good reason to believe that in many cases only the first or second neighbors will suffice for reaching a monolithic solution. This technique of eliminating the long-range interactions is by nature linear with respect to the total number of DOFs, although in practice, long-range effects still indirectly exist via the neighbor-to-neighbor interactions and therefore the efficiency is not necessarily linear. More discussion on the performance of this method is given in section 3. It should be noted that for an infinitely large molecule, both the Hartree potential and the nuclear potential blow up at every point due to the slow-decaying nature of these potentials. However, for a neutral molecule, the sum of these two potentials is bounded. This can be used for example in a modified Poisson equation as:

$$\nabla^2\left(V_0(\mathbf{r})+V_d(\mathbf{r})\right)=-4\pi n(\mathbf{r})+4\pi\sum_I Z_I\delta(\mathbf{R}_I-\mathbf{r}), \tag{2.33}$$

where, again, n(r) is the density of electrons at each point r in space, Z_I is the atomic number of atom I at position R_I, and δ is the Dirac's delta function. The solution to this term can be easily treated by the corresponding Green's function. Once the combined value of V_0+V_d at each step is determined, it can be used in Eq. (2.22) as before.

2.3. Results and discussion

In order to demonstrate our FEM-HF approach, we have computed the total energy of selected light atoms and small molecules (namely H, He^+, He, Be, Ne, H_2^+, BH and HF) as shown in Table 2.1. These energy values are compared with the Hartree-Fock limit or the corresponding theoretical/experimental values. A typical FEM mesh for an atom, with small cubic elements near the nucleus and concentric mesh geometrically expanding towards the domain boundary, as shown in Figure 2.4, has been used to capture the higher gradients of the electron density near the nuclei. For the diatomic molecules, in the axial direction there is a transition region that smoothly connects the two regions with finer elements around each nucleus. Both the domain size and the mesh density can be adjusted by changing the number of elements and the ratio of the lengths of neighboring elements. Linear isoparametric brick elements have been used in the current computations. Homogeneous boundary conditions have been assumed in all calculations.

The choice of mesh configuration in the calculations of all atoms and light molecules is determined by the convergence of the computations. We have used an adaptive meshing strategy [79] that ensures higher accuracy in places where more rapid variation in the field variables is expected. In this approach, the representative element dimension d is inversely proportional to the second gradient of the electron density via the following criterion:

$$d \propto \left[\max\left(\left| \frac{\partial^2 \rho}{\partial x^2} \right|, \left| \frac{\partial^2 \rho}{\partial y^2} \right|, \left| \frac{\partial^2 \rho}{\partial z^2} \right| \right) \right]^{-1}, \quad d < d_0 . \tag{2.34}$$

We start this adaptive meshing procedure by using an approximate electron density distribution from a simple tight binding analysis. After that, the mesh can be constructed from the most recent electron density distribution. We do at most one remeshing beyond this step as the meshing process is expensive and usually the quality of the first mesh is quite good. Figure 2.5 shows the electronic configuration of the krypton atom as an example of the results obtained with this scheme.

The size of the core region shown in Figure 2.4 is typically chosen to be ~10% of the total computation domain. For a hydrogen atom, this choice of FEM mesh gives almost a perfect fit to the analytical solution of the ground state, as shown in Figure 2.6. The resulting charge density distribution for hydrogen fluoride is shown as an example in Figure 2.7. The total energies calculated from our method have less accuracy than one might desire. This is a common problem with FEM approaches in electronic structure theory and arises because of the singular nature of the Coulomb potential. Although we do have a finer mesh in the regions near the nuclei, this is not sufficient to achieve chemical accuracy (typically considered to be on the order of a millihartree error). The results do improve as we increase the density of the mesh, as shown in Table 2.1. However, the important quantities in chemistry are energy differences and not necessarily absolute energies. Thus, it is important to investigate the degree to which the error incurred by our FEM approach is systematic. In order to answer this question, we

have calculated the potential energy curve for HF using both our FEM approach and the conventional atom-centered Gaussian basis set method (using the aug-cc-pv5z basis set). The results are shown in Figure 2.8. Importantly, our FEM approach yields a smooth potential energy curve and furthermore is in good agreement with the finite basis set results (we have shifted the FEM potential curve by a constant ~20eV representing the systematic error incurred by integration of the core region). The use of pseudopotentials would alleviate much of the difficulties with respect to the mesh density needed in the core region, since pseudopotentials remove the Coulomb singularity near the nuclei.

The same approach could be used to study the electronic structure of the multi-atom molecules. As an example, the internal rigid rotation of the ethylene molecule (C_2H_4) was considered. In this case a tetrahedral mesh is more convenient. The unstructured mesh used for solving this problem is shown in Figure 2.9a. Their corresponding FEM energy values are shown in Figure 2.9b and compared to the Gaussian basis-set results. These results generally agree very well with each other, although some nonsmoothness arising from the nonsystematic error associated with integration of the core region is now visible. For comparison purposes, the FEM energies have been shifted ~180kcal/mol such that the FEM and aug-cc-pVTZ energies are matched at the planar ethylene geometry.

Enhanced convergence is a major benefit of the FEM approach. We have investigated two aspects of convergence for the FEM-HF approach. First, we consider the convergence with decreasing element size, i.e. increasing number of elements. We define a convergence norm as the error in the total energy as a function of the number of

divisions on each side of the computation domain. Results for helium atom are shown in both Table 2.1 and Figure 2.10a. Figure 2.10a clearly shows that the convergence norm follows a power law with the number of divisions (the number of elements increases as the third power of the number of divisions in each direction). Secondly, we consider the effect of the finite computational domain. The true orbitals are expected to extend to infinity in space, but they are forced to be identically zero on the boundaries of the chosen computational domain in our FEM-HF approach. Since the magnitude of electron density decreases exponentially far from the center of an atom or molecule, one expects the domain truncation error resulting from this forced truncation to be quite small as long as the computational domain is sufficiently large. Indeed, we found domain truncation errors to be negligible for a computational domain whose boundaries are farther than $20a_0$ away from each nucleus in every direction. In order to demonstrate this size dependence, the convergence of the total energy as a function of the total domain size is shown in Figure 2.10b for the helium atom from which exponential convergent behavior is evident.

Another important practical issue is the scaling of the required computer time with the number of DOFs. Scaling up the calculations to larger, more complex systems requires a high computational efficiency (a task which is quite formidable with the original non-local approach). The computational efficiency of the current method is very good. All the calculations in Table 2.1 have been carried out on a single processor personal computer using an implementation of our method within Matlab version 7.4.0.287. We have performed FEM-HF simulations on graphene sheets with over 1000

electrons with a full account of all the electrons in the system on a cluster of 16 processors in less than 48 hours. [80] The wall clock time required as a function of the number of DOFs for beryllium atom is shown in Figure 2.11. We have used the DC formalism described above with four domains. The systems studied are too small to expect significant computational advantage from decreased scaling, but we do benefit from parallelization across the domains as mentioned above. Although the FEM method as described above formally has computational effort that scales linearly with the number of degrees of freedom for a given problem, Figure 2.11 shows that our implementation exhibits nearly quadratic scaling. As discussed in the supporting information (Appendix A), we have determined that this is due to the increased Matlab internal communications overhead of handling large sparse matrices. Implementation of the method in Fortran or C should lead to the expected linear scaling behavior.

Several features of the current computations are noted. First, no pseudopotentials have been used in the current formulation, i.e. all electrons are included explicitly, enabling the explicit calculation of the exact exchange potential. Implementation of pseudopotentials is expected to improve the accuracy of the calculations significantly, since the Coulomb singularity at the nuclei is then removed or softened. Second, we have not used symmetry in any of the cases shown. Therefore this approach, in principle, can be used for a truly *ab initio* analysis of any arbitrary three-dimensional molecular system with a reasonable quantum resolution.

One of our future goals is to use this technique to reliably study phenomena such as electro-mechanical coupling of systems such as carbon nanotubes. In a separate

article it will be shown that the current method can be used to study the transport properties of the quantum systems of interest with facile convergence. We can then use the reduced scaling local approximation to study the effects of defects and mechanical deformation on the properties of individual graphene sheets and carbon nanotubes under tension, torsion, bending and general deformation conditions at finite temperatures. [80-81]

2.4. Conclusions

We have developed a method of using the finite-element method (FEM) to calculate the Hartree-Fock (HF) *ab initio* electronic-structures in three dimensions. Traditional basis set approaches to the HF method were discussed and compared to this FEM implementation. Although the absolute accuracy of the method is hampered by the singularities in the Coulomb potential at nuclear positions, the method was nevertheless shown to produce smooth potential curves due to our adaptive meshing scheme. Higher absolute accuracy could be obtained by using pseudopotentials to alleviate the singularities at the nuclei. We introduced a new acceleration scheme for the exchange contribution which uses an approximate X-α formulation to rapidly compute an approximate exchange for many of the self-consistent field iterations. We also formulated the method in the context of a divide-and-conquer approach to improve the scaling of the implementation. Larger systems need to be investigated to demonstrate reduced scaling from this approach, but the method lends itself easily to parallelization

even for small systems and the implementation we described here has already been parallelized (with each divide-and-conquer domain apportioned to a separate processor). The Hartree-Fock method forms the basis for many methods including electron correlation effects and the extension of this work to calculations of electronic structure of more complex molecules using post-HF methods seems feasible. Implementation in the context of density functional theory, including hybrid functionals which require exact exchange, can also be implemented.

2.5 Figures and tables

Table 2.1. Total energy (a.u.) of the selected chemical species as a function of the number of elements in each direction compared to the reference values. For selection of the computation domain, please refer to the discussion on the effect of domain size on convergence in the text.

Species	Number of elements in each direction (Number of DOFs)			Reference values
	16 (2299)	20 (5850)	24 (11913)	
H	-0.481766	-0.487039	-0.490062	-0.5*
He^+	-1.949650	-1.965756	-1.975213	-2.0*
He	-2.739113	-2.802768	-2.839044	-2.8616800 [21]
Be	-13.76410	-14.09187	-14.27383	-14.573023 [21]
Ne	-112.8331	-118.7391	-122.3812	-128.54710 [21]
H_2^+	-0.586220	-0.592524	-0.595744	-0.6029**
BH	-21.62545	-23.24638	-24.18116	-24.8088515 [29]
HF	-90.32122	-93.75640	-95.26689	-100.071 [22]

* Exact theory

** Experimental

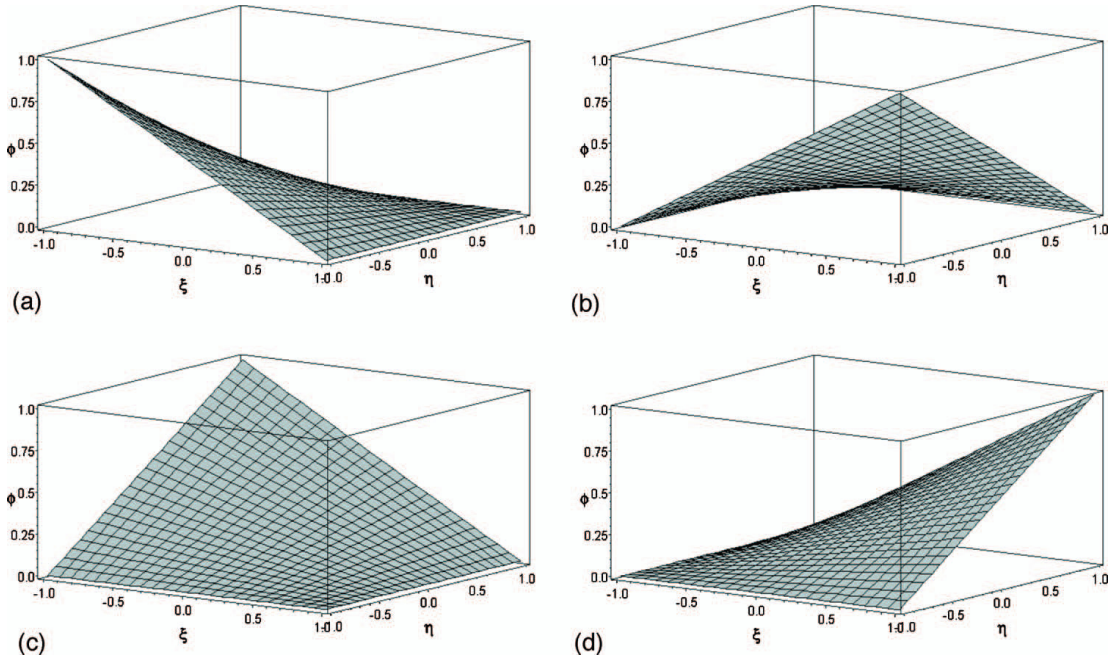


Figure 2.1. The linear isoparametric quadrilateral shape functions in (ξ,η) natural coordinates are:

a) $N_1 = (1-\xi)(1-\eta)/4$, b) $N_2 = (1+\xi)(1-\eta)/4$, c) $N_3 = (1-\xi)(1+\eta)/4$, d) $N_4 = (1+\xi)(1+\eta)/4$.

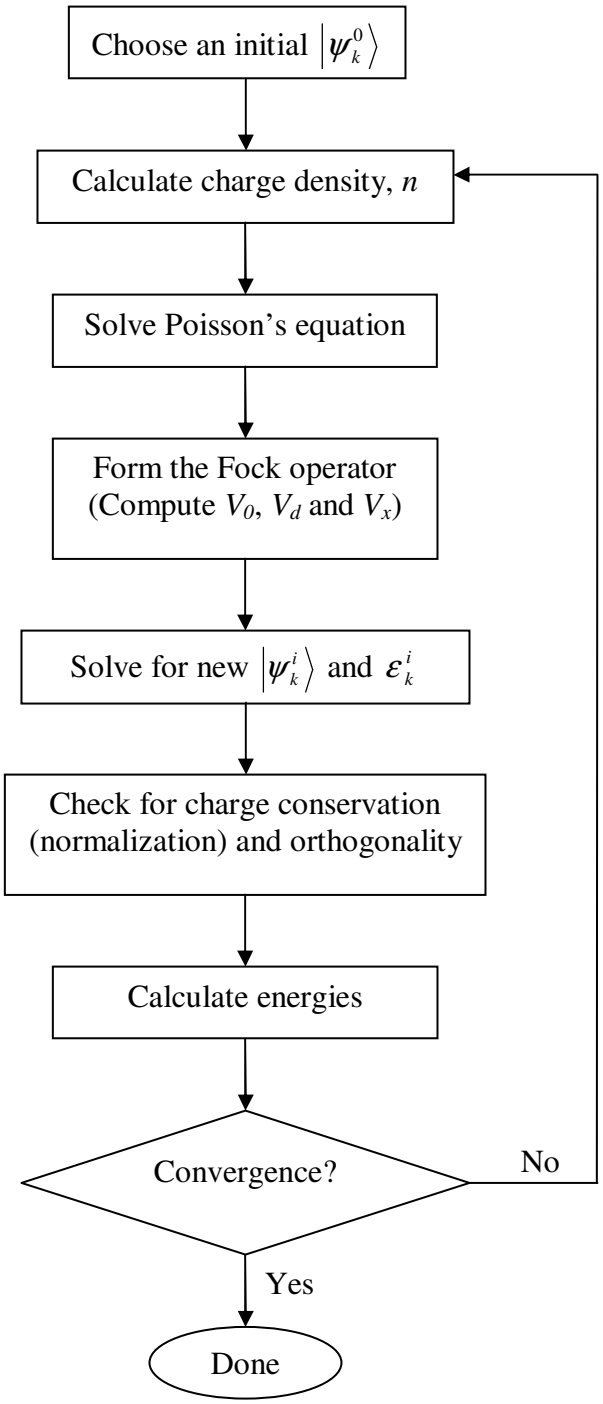


Figure 2.2. The flowchart of the self-consistent field (SCF) iterative approach.

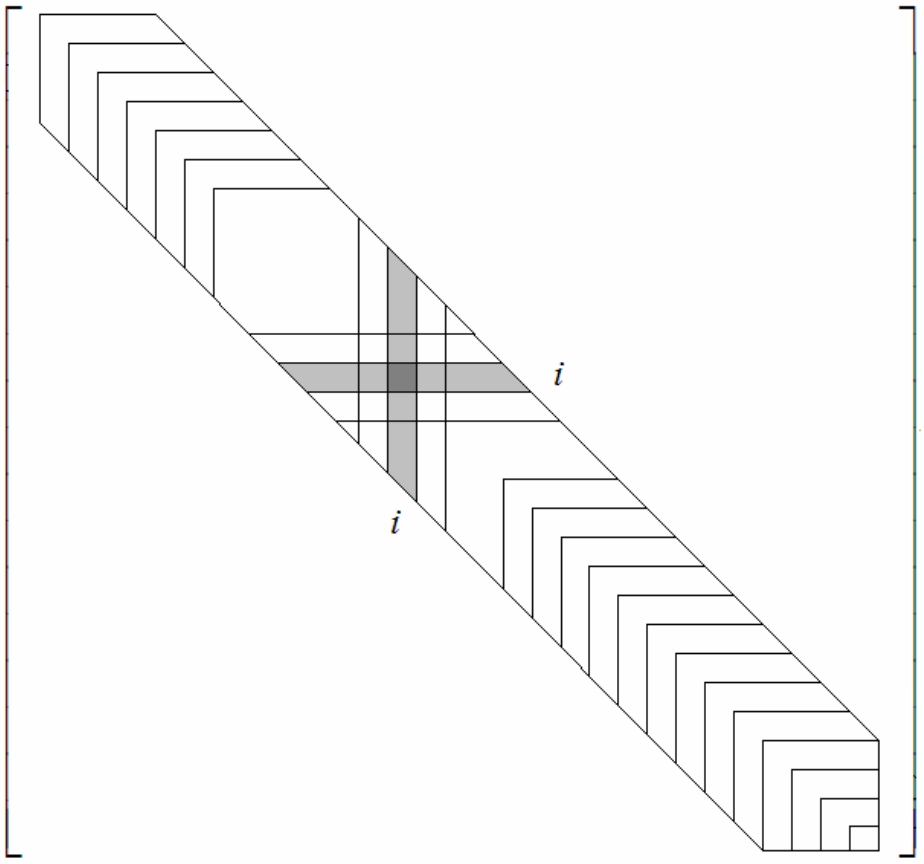


Figure 2.3. Schematic of the Hamiltonian matrix structure after partitioning. The gray shaded area shows the sub-matrices corresponding to the sub-domain β^i.

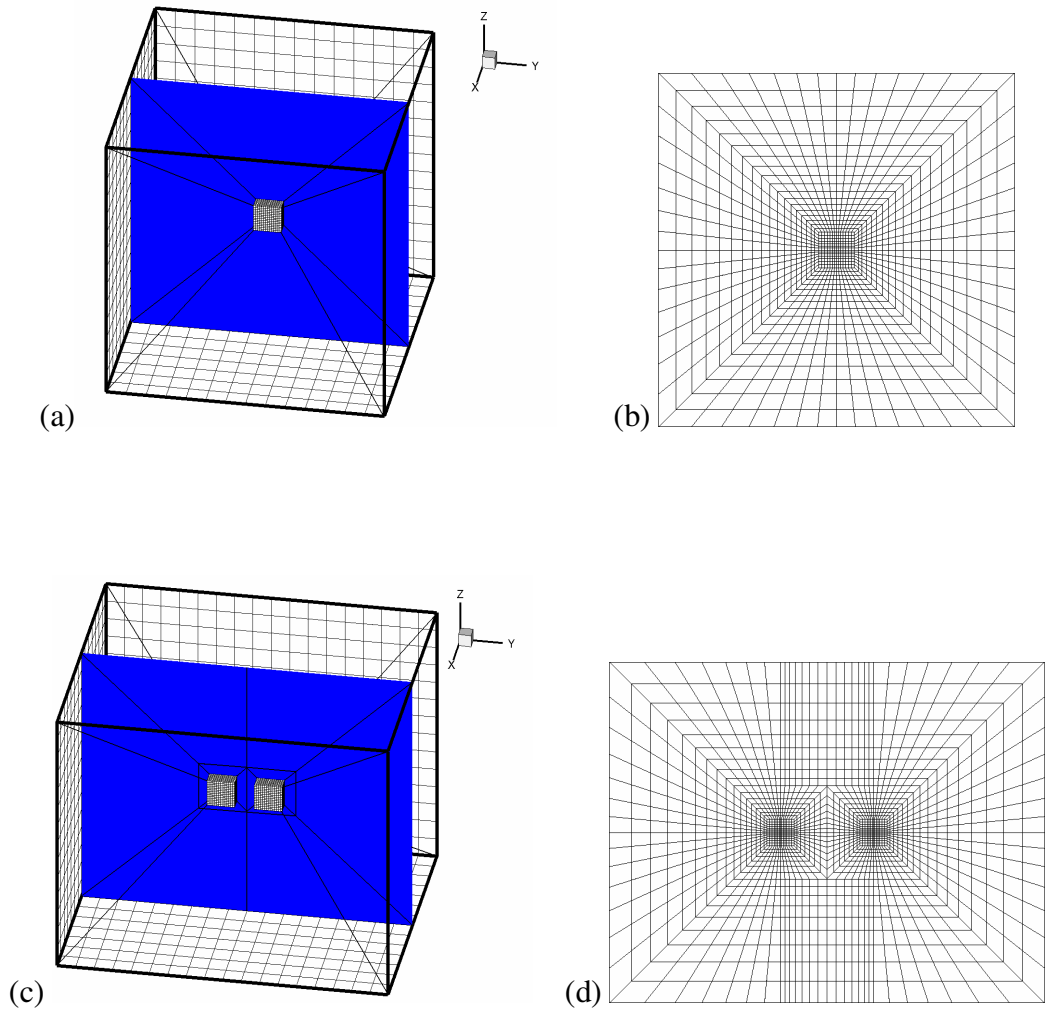


Figure 2.4. Schematic of the typical concentric-cubes mesh configurations used at the location of each nucleus; (a) the 3D mesh and (b) the corresponding slice for atoms, (c) the 3D mesh and (d) the corresponding slice for diatomic molecules.

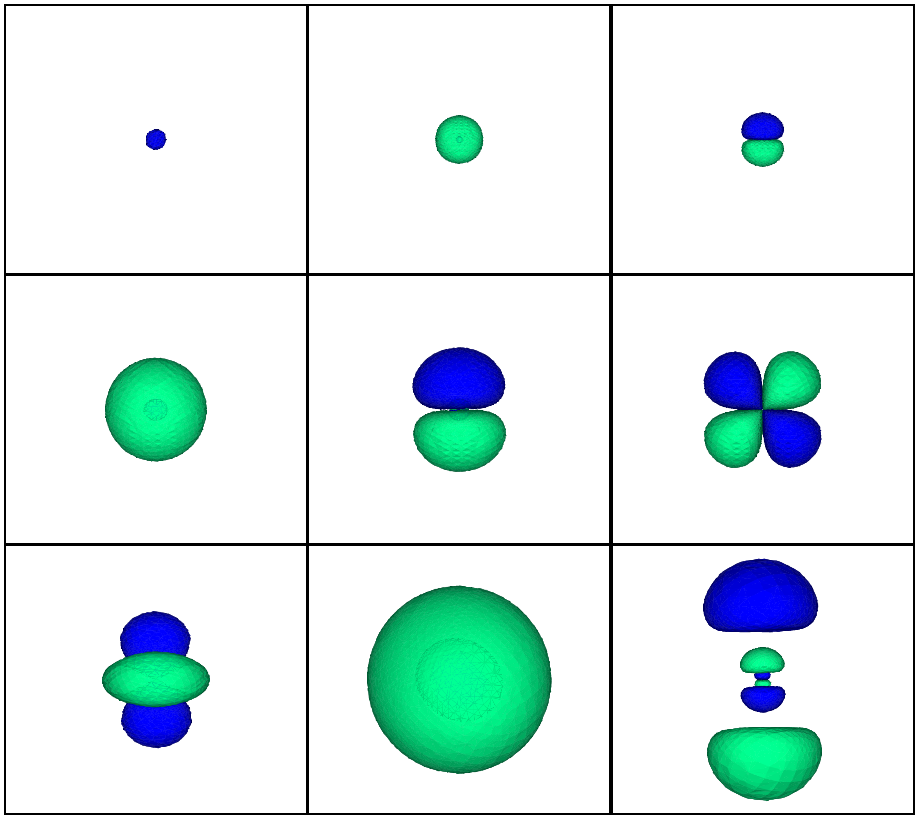


Figure 2.5. Computed atomic orbitals using FEM for the krypton atom with the atomic number of 36 and the atomic configuration of: [Kr] $1s^2$ $2s^2$ $2p^6$ $3s^2$ $3p^6$ $3d^{10}$ $4s^2$ $4p^6$ From left to right and from top to bottom the boxes correspond to the iso-surfaces of the following orbitals: 1s, 2s, $2p_x$, 3s, $3p_x$, $3d_{xy}$, $3d_{zz}$, 4s and $4p_x$. All the orbitals are drawn to the same scale. Therefore, their true size could be compared to each other from this figure. Different elements have different size and shape s, p, d, and f atomic orbitals.

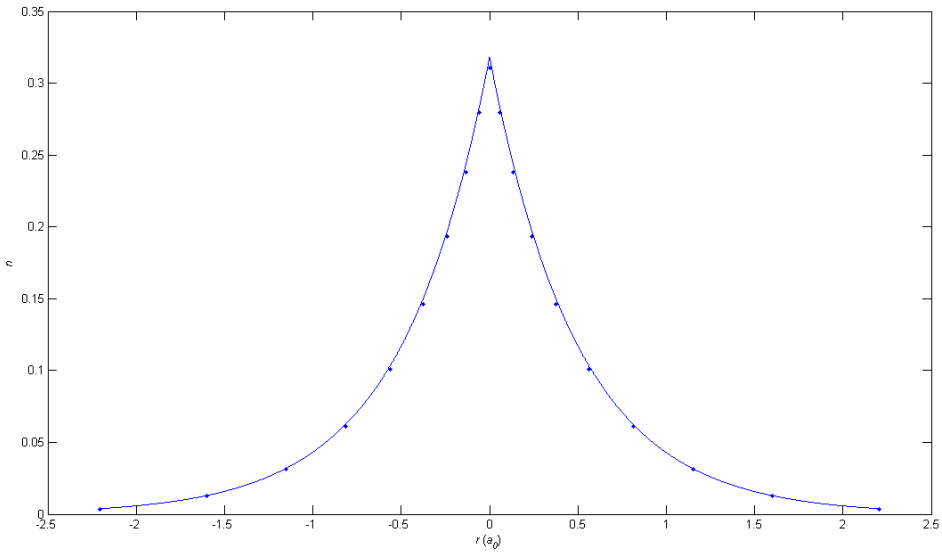


Figure 2.6. Radial charge density distribution for hydrogen atom from the finite-element solution with 24 divisions on each side (the dots). The curve represents the exact theoretical solution.

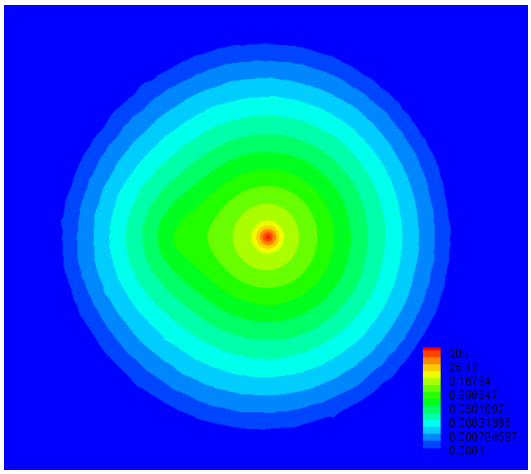


Figure 2.7. Axisymmetric charge density distribution for hydrofluoric acid using 24 divisions in the axial direction. The density contours clearly show the H^+F^- polarity of the molecule.

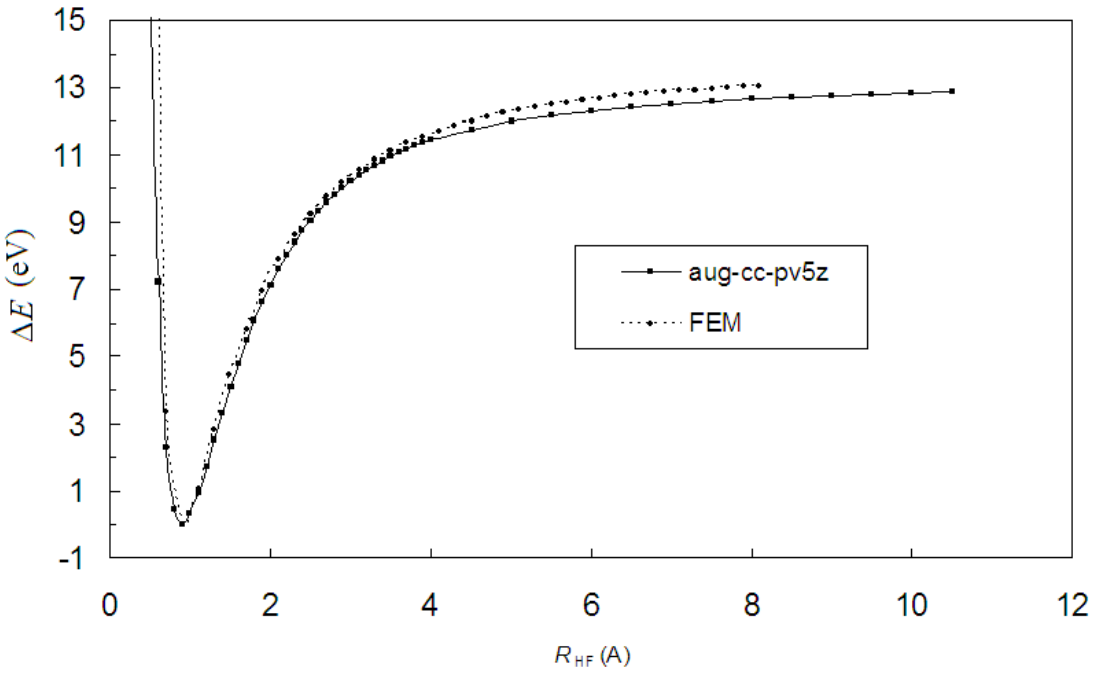
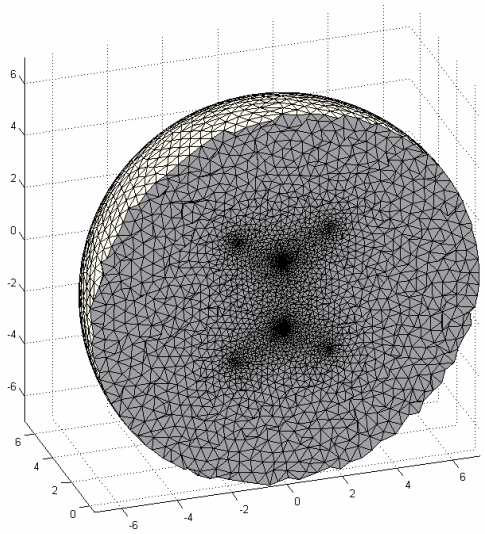
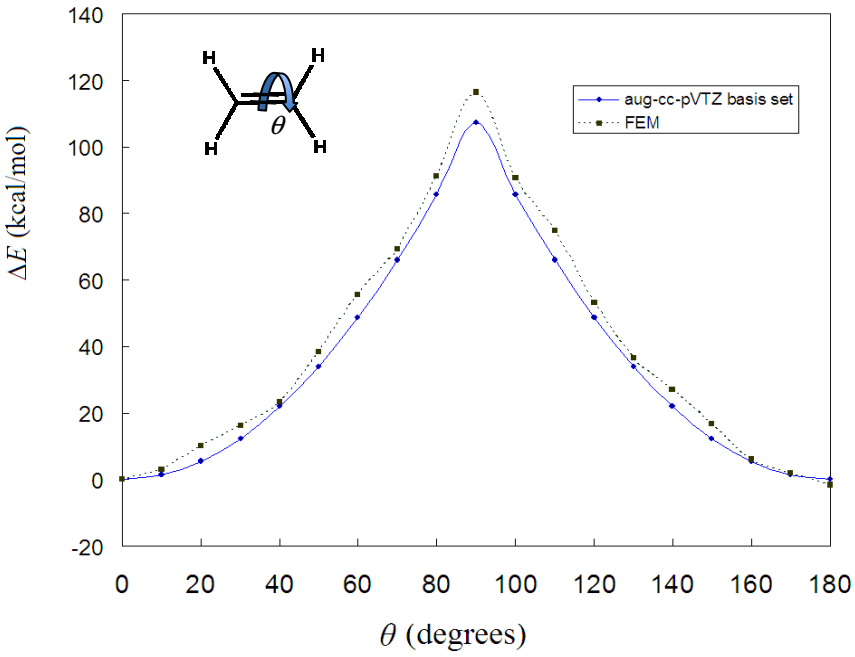


Figure 2.8. Bond dissociation curve for hydrofluoric acid from the Gaussian basis set and the FEM. The FEM results have been uniformly shifted to agree with the Gaussian basis set results at the equilibrium distance.

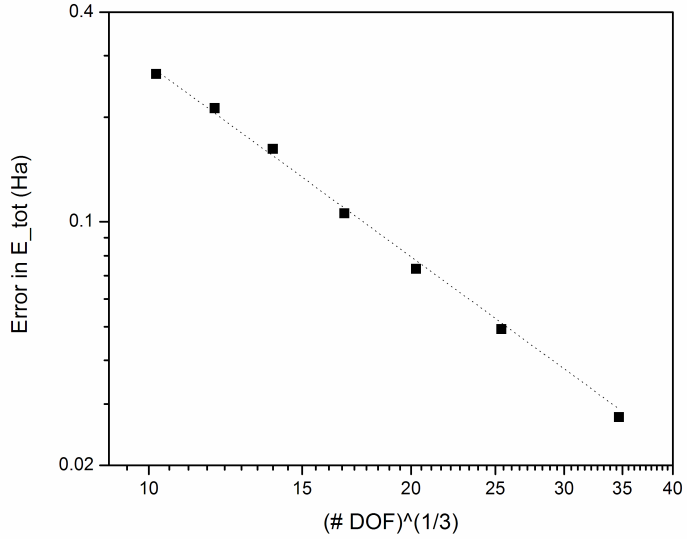


a)

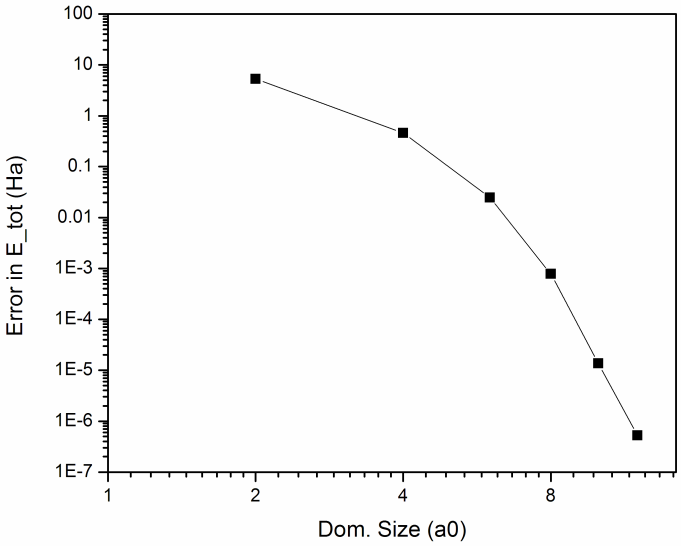


b)

Figure 2.9. a) Typical mesh configuration and b) total energy profile for rigid internal rotation of ethylene (see inset). The FEM results have been uniformly shifted to agree with the Gaussian basis set results at the planar ethylene geometry.



(a)



(b)

Figure 2.10. Convergence behavior of the finite-element solution for the helium atom; the error in the total energy as a function of (a) the average number of divisions in each direction and (b) the total domain size.

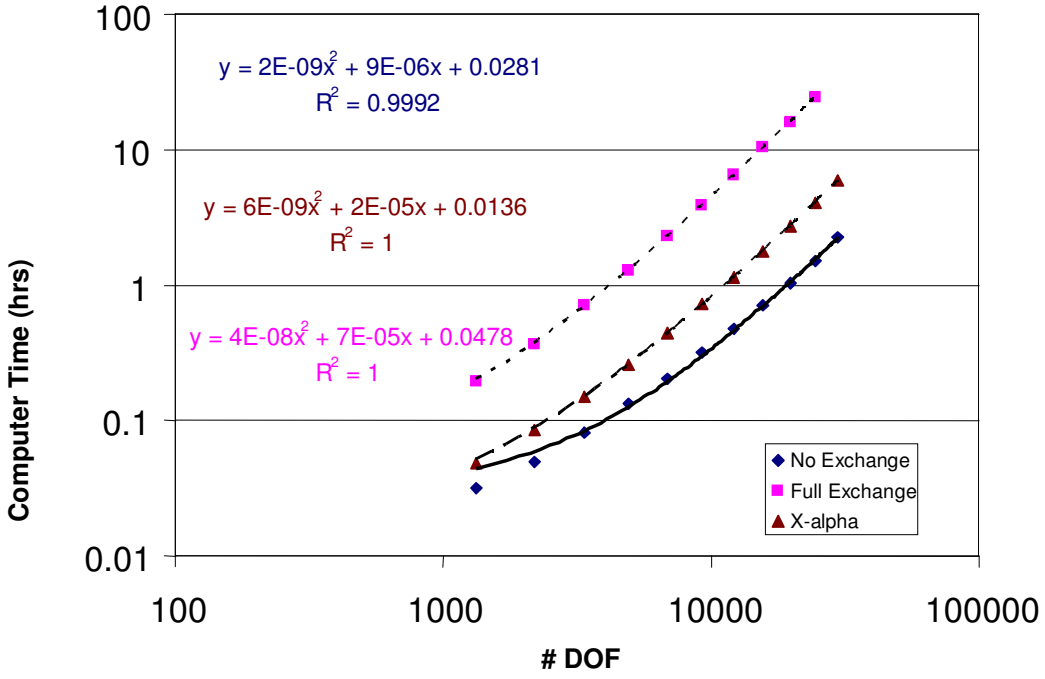


Figure 2.11. Computer time as a function of the number of degrees of freedom for the Be atom.

CHAPTER 3: APPLICATION OF FACE-BASED SMOOTHED FINITE-ELEMENT AND ENRICHMENT FUNCTIONS IN ELECTRONIC STRUCTURE CALCULATIONS

3.1. Introduction

Much computational effort has been devoted to the solution of the electronic Schrödinger equation for molecular systems, as this yields the potential energy surfaces that form the basis for further dynamical and/or statistical mechanical calculations. Since the seminal contribution of Roothaan, quantum chemistry methods are traditionally expressed using finite basis sets comprised of smooth and continuous functions (atom-centered Gaussians) to describe the electronic degrees of freedom. Although this approach has proven quite powerful, it is not well-suited for large basis sets because of linear dependence problems and ill-conditioning of the required matrices. Most common methods for solving the electronic Schrödinger equation start from a one-electron basis set composed of smooth nonorthogonal functions, generally atom-centered Gaussians. This choice has the advantage of allowing one to encapsulate much of the important physics of the solutions in the basis set. For example, the cusp at the nuclei is approximated by contraction of many Gaussian functions and the basis functions are

chosen to provide a good representation of the solutions for free atoms. The primary disadvantage of this approach is that it is awkward to carry these calculations to convergence since the nonorthogonal nature of the basis functions leads to ill-conditioned matrices as the basis set size is increased. Accurately representing the rapidly varying and complicated orbitals and reconstructing them from many Gaussian functions is also not very efficient. The major reason that these methods have been so successful up to now is the availability of analytical integration formulas that procure ultra-fast evaluation of these two-center integrals. Grid-based methods provide an alternative approach wherein convergence can be achieved more systematically. Examples include wavelet, finite element (FE), and finite difference (FD) methods [25]. In Ref. [80], a divide-and-conquer real-space finite element implementation of the restricted HF (RHF) method for polyatomic molecules was described. There, standard C^0–continuity eight-node hexahedral and four-node tetrahedral elements have been used in a fully adaptive meshing scheme without symmetry assumptions. Even though this local-basis approach to electronic structure theory was shown to have a linear time scaling which allows for systematic convergence and facile parallelization and promises to provide a rigorous and accurate way towards the full *ab initio* analysis of materials at larger scales, the computational burden of the assembly and integration of the exceedingly small elements near the core region has proven to be a great detriment to the computational efficiency of the FEM methods compared to the global basis set methods and it's clear that a successful treatment of this problem is a key issue in any future application of the FEM methods in the field of computational quantum chemistry. In

order to demonstrate the origins of the challenge facing the numerical treatment of this problem, the ground state of the 1D infinite square well potential with a point charge Z at the origin, i.e. the solution to the following equation

$$\left[-\tfrac{1}{2}\frac{\partial^2}{\partial x^2}-\frac{Z}{|x|+A}\right]\psi(x)=E\psi(x) \quad \text{in } -B\le x\le B \tag{3.1}$$

with A=1e-10, B=1 and Z=0-4 is shown in Figure 3.1. As it can be seen, the ground state wavefunction is sharply peaked at the core, and this is where the nuclear charge is still moderately low and the addition of higher occupied orbitals can only push the core orbitals further in and make the situation even worse.

For treatment of the nuclear singularity, several options were considered at the early stages of carrying out this research including: the pseudopotential approach, the Model-Potential approach [82], the perturbation solution, the variational matching approach [83], the enriched finite element method [84] and the faced-based smoothed finite element method [85]. Each of these methods have their own advantages and disadvantages which a detailed discussion of them goes beyond the scope of this dissertation.

In this chapter, we use the concept of generalized/extended finite element method (GFEM/XFEM) within the partition of unity framework [86-89] to remedy this problem. The idea in generalized/extended finite element method is to use a priori knowledge about the solution of a problem to obtain an improved numerical approximation. This knowledge is, for example, in the form of known analytical solutions at the limit of very

close to the crack tip (in fracture mechanics) or very close to the nuclei (in quantum chemistry). From the other hand, it's known that FEM gives higher energy values than the true energies. In structural mechanics nomenclature, it's sometimes said that FEM stiffness matrices are "too stiff". This property makes the convergence slow and occasionally introduces spurious modes in modal analyses. To remedy these problems, several solutions have been proposed such as the strain smoothing technique for meshfree methods [90], the smoothed finite element method (SFEM) [91-92], the node-based smoothed finite element method (NS-FEM) [93] and the face-based smoothed finite element method (FS-FEM) [85]. The FS-FEM has been reported to be numerically very efficient with superior convergence characteristics for modal analyses [94]. In this article, an effort has been made to take advantage of these good characteristics of FS-FEM enriched with exponential enrichment functions at the points of singularity and explore the accuracy and efficiency of this approach for some selected test atoms and molecules compared to our previous results using standard FEM. Here, for comparison reasons, we focus on the Hartree-Fock (HF) method which avoids the problem of self-interaction via the inclusion of all the electrons explicitly, enabling the explicit calculation of the exact exchange potential. Moreover, HF does not entail any fitting parameters or empirical functionals which makes it ideal for benchmark numerical studies. However, we'd like to emphasize that extension to Møller–Plesset perturbation theory (MP), coupled cluster (CC) or the Kohn-Sham (KS) approaches such as the density functional theory (DFT) approaches including the local density (LDA) and

generalized gradient (GGA) approximations with hybrid functionals [16-17] is conceptually straightforward.

There have been numerous efforts aimed at using FEM in conjunction with HF or DFT methods. For a detailed literature review of these methods, we refer the interested reader to [80] and the references therein. In section 2 we present the generalized finite element formulation of the self-consistent field method. In section 3 we present some representative results and discussion. Section 4 contains summary and conclusions.

3.2. The generalized and face-based smoothed finite element formulation of the self-consistent field

The derivation of the FEM within the context of RHF is given elsewhere. [80] Thus, we only mention the major resulting equations here. The normalized single-electron equation for spin-orbitals is

$$-\tfrac{1}{2}\nabla^2\psi_k + V\psi_k = \varepsilon_k\psi_k \quad \text{in } \Omega, \tag{3.2}$$

where ψ_k and ε_k are the k-th molecular orbital and its energy, Ω is the computational domain, and $V = V_0 + V_d + V_x$ is the potential where V_0 contains electron-nuclear attraction and any external potential, V_d is the Hartree potential generated by the electron-electron repulsion, and V_x is the non-local exchange potential due to the exchange symmetry of the wavefunctions of indistinguishable particles. The Hartree potential V_d can be obtained by solving Poisson's equation:

$$\nabla^2 V_d(\mathbf{r}) = -4\pi \sum_{j=1}^{N} n_{jj}(\mathbf{r}), \tag{3.3}$$

where $n_{jk}(\mathbf{r})$ is defined by

$$n_{jk}(\mathbf{r}) = \psi_j^*(\mathbf{r})\psi_k(\mathbf{r}), \tag{3.4}$$

and where N is the total number of the electrons. The exchange potential $V_x(\mathbf{r})$ can be computed using the following relation

$$V_x(\mathbf{r})\psi_k(\mathbf{r}) = -\sum_{j=1}^{N/2} (V_x)_{jk}\psi_j(\mathbf{r}), \tag{3.5a}$$

where

$$\nabla^2 (V_x)_{jk} = -4\pi n_{jk}(\mathbf{r}). \tag{3.5b}$$

We adopt atomic units throughout. The corresponding weak form of Eq. (3.2) is

$$\int_\Omega \tfrac{1}{2}\nabla\psi' \cdot \nabla\psi_k d\Omega - \int_\Gamma \tfrac{1}{2}\psi'\nabla\psi_k \cdot \hat{\mathbf{n}} d\Gamma + \int_\Omega \psi'(V_0 + V_d + V_x - \varepsilon_k)\psi_k d\Omega = 0, \tag{3.6}$$

where Γ is the boundary and $\hat{\mathbf{n}}$ is the outward unit normal at each point. Assuming natural, Neumann or free boundary conditions (i.e. no external flux) the second term drops out on the surface. After performing the standard discretization procedure of the finite elements, we introduce the ansatz:

$$|\psi\rangle \approx \mathbf{N}^T\mathbf{a} = \tilde{\mathbf{N}}^T\tilde{\mathbf{a}} + \tilde{\tilde{\mathbf{N}}}^T\tilde{\tilde{\mathbf{a}}} = \sum_i \tilde{N}_i(\mathbf{r})\tilde{a}_i + \sum_p \sum_i \tilde{\tilde{N}}_i^p(\mathbf{r})\tilde{\tilde{a}}^p, \tag{3.7}$$

where the ket vector $|\psi\rangle$ is the quantum state of interest, $\mathbf{N}^T =< \tilde{\mathbf{N}}^T, \tilde{\tilde{\mathbf{N}}}^T >$ are the shape functions and $\mathbf{a}^T =< \tilde{\mathbf{a}}^T, \tilde{\tilde{\mathbf{a}}}^T >$ are the nodal degrees of freedom (DOFs). $\tilde{\mathbf{N}}$ and $\tilde{\mathbf{a}}$ correspond to the regular shape functions and DOFs, and $\tilde{\tilde{\mathbf{N}}}$ and $\tilde{\tilde{\mathbf{a}}}$ correspond to the enriched shape functions and DOFs, respectively. This expansion ensures the completeness of the solution via inclusion of the linear shape functions explicitly in the basis. In this paper, we use standard C^0–continuity linear isoparametric four-node tetrahedral elements for the regular shape functions. Isoparametric means that the same shape functions are used to specify the relation between the global and local coordinate systems and to describe the variation of the field variables. This choice of elements caters considerable flexibility in using adaptive meshing strategies. For the choice of enriched shape functions, we used the following relations:

$$\tilde{\tilde{N}}_i^p(\mathbf{r}) = S(\mathbf{r}) \times \left(\phi_p(\mathbf{r}) - \sum_i \tilde{N}_i(\mathbf{r}) \times \phi_p(\mathbf{r}_i) \right), \text{ (for } p=1,\ldots,P), \tag{3.8}$$

where ϕ_p are the Slater type atom-centered basis functions, P is the total number of such functions and $\mathbf{r}_i$ is the position of the node corresponding to $\tilde{N}_i(\mathbf{r})$. S is a blending function which is zero far from the point of singularity and is 1 close to the points of singularity so that the compatibility of elemental solutions is guarantied among the enriched and non-enriched elements. The linear form of this function is:

$$S(r)=\begin{cases} r<r_{cut}-t & 1 \\ r_{cut}-t\le r\le r_{cut} & (r_{cut}-r)/t\,, \\ r_{cut}<r & 0 \end{cases} \tag{3.9}$$

where r_{cut} is the radius around the points of singularity below which all elements are enriched and t is the thickness of the blending region, typically on the order of the element size. Substituting the wavefunction ansatz into Eq. (3.3), we obtain the following matrix equation for determining the orbitals:

$$\mathbf{H}_k\mathbf{a}_k=\mathbf{v}_k\,,\quad (k=1,\ldots,N/2), \tag{3.10}$$

where

$$\begin{aligned}\mathbf{H}_k&=\int_\Omega\left\{\tfrac{1}{2}\nabla\mathbf{N}^{FS}(\nabla\mathbf{N}^{FS})^T+(V_0+V_d-\varepsilon_k)\mathbf{N}\mathbf{N}^T\right\}d\Omega\,,\\ \mathbf{v}_k&=\sum_{j=1}^{N/2}\left\{\int_\Omega(V_x)_{jk}\mathbf{N}\mathbf{N}^T\,d\Omega\right\}\mathbf{a}_j\,.\end{aligned} \tag{3.11}$$

In FS-FEM the calculation of the face-based smoothed gradients, $\nabla\mathbf{N}^{FS}$, is different from that of the conventional FEM. For a detailed discussion of this topic the interested reader should consult reference [85], but a short version is given below for the sake of completeness:

$$\nabla\mathbf{N}^{FS}=\left\{\left(Vol(e^i)\cdot\nabla\mathbf{N}^i\right)\oplus\left(Vol(e^j)\cdot\nabla\mathbf{N}^j\right)\right\}\Big/Vol(e^i\cup e^j)\,. \tag{3.12}$$

In this equation, $\nabla\mathbf{N}^i$,$Vol(e^i)$ and $\nabla\mathbf{N}^j$,$Vol(e^j)$ are respectively the regular gradient matrices and volumes of two adjacent elements e^i and e^j (Figure 3.2). The symbol $\oplus$

in this notation is the assembly operator which simply means that the values corresponding to the same DOFs add up together. Eq. (3.10) can be solved iteratively as a sequence of linear equations with updated estimate of the eigenvalue ε_k in each iteration. As the first step in the computational procedure, a reasonable initial guess is provided. Based on this initial guess, one can calculate the electronic density and solve the Poisson equations, and then calculate the Fock operator and solve for a new set of orbitals and their eigen-energies. This process is repeated until the values of the calculated energies become stationary. The criterion for convergence has been chosen such that the absolute change in the total energy of the system in the successive iterations to be less than a small tolerance. Homogeneous boundary conditions have been assumed in all calculations. The iterative subspace (DIIS) method is used to accelerate the convergence of this self-consistent field procedure. [75, 95] Also, it is found out that it's usually best to first obtain a solution with a relatively coarse mesh and then map that solution on a finer mesh and iterate to convergence. This way, many of the iterations far away from the stationary solution are quickly marched through. Sometimes for degenerate systems, a case of pseudo-instability occurs in which wavefunctions do not become stationary. One way to handle this situation is to use a dummy charge to split the degeneracy and then gradually remove the dummy charge.

3.3. Results and discussion

Structured meshes are typically more accurate than the unstructured meshes with the same order of interpolation due to their symmetric error cancellation. In this paper, we have used tetrahedral elements with an adaptive structured meshing strategy that ensures higher accuracy in places where more rapid variation in the field variables is expected. In this approach, the representative element dimension, d, should satisfy the following simple criterion based on electron density:

$$d < \kappa / \rho, \qquad d_{\min} < d < d_{\max}, \tag{3.13}$$

where κ is a constant. If d is too large, the element size is halved until this condition is met. We start this adaptive meshing procedure by using an approximate electron density distribution from a simple tight binding analysis. After that, the mesh can be constructed from the most recent electron density distribution. This procedure results in a mesh similar to the typical mesh shown in Figure 3.3 due to the exponential type nature of the electron density distribution. We do at most one remeshing beyond this step as the meshing process is expensive and usually the quality of the first mesh is quite good. In the case of the enriched simulations, a $d_{\min}$ is imposed so that the element density at the cores does not get unnecessarily too high since the inverse density criterion tends to result in a proliferation of very small elements at the core which can cause numerical issues. Nonetheless, linear elements with the size of a fraction of a thousandths of a Bohr radius, a_0, are commonly required.

In order to demonstrate the different computational methods discussed above, we have computed the total energy of the He atom with these different methods and the results are shown in Figure 3.4. With the same number of Gauss integration points per element used, all these methods have a computational time that scales linearly with the total number of DOFs, however, enriched simulations are more accurate than their regular counterparts and FS-FEM is significantly more accurate than standard FEM. Impressively, FS-FEM is even more accurate than the enriched FEM. We have calculated the potential energy curve for hydrogen fluoride (HF) using the different FEM approaches and the conventional atom-centered Gaussian basis set method (using the aug-cc-pv5z basis set). To speed up these simulations, we implemented a resume functionality that linearly maps the last electronic configuration to the initial configuration of a higher density mesh, thus avoiding many of the initial iteration steps. The results are shown in Figure 3.5. FEM approaches are in good agreement with the finite basis set results. Our previous results with linear hexahedral elements indicate that these elements generally perform better than their tetrahedral counterparts, although as mentioned above, meshing is not as flexible with hexahedral elements for multi-atom molecules with arbitrary geometry. It's also worthwhile to mention that we have not assumed any symmetries here and therefore this approach, in principle, can be used for a truly *ab initio* analysis of any arbitrary three-dimensional molecular system with a reasonable and adjustable quantum resolution.

3.4. Summary and Conclusions

We have applied an enrichment method to the standard and face-based smoothed finite-element method (FS-FEM) to calculate the Hartree-Fock (HF) *ab initio* electronic-structures in three dimensions. Our results suggest that FS-FEM is significantly more accurate than the standard FEM in dealing with modal analyses even without using enrichment. The proposed inclusion of Slater-type enrichment functions drastically improves the quality of the solutions both in regular FEM and FS-FEM without any substantial adverse effect on the computational efficiency of the corresponding methods. However, still a relatively large number of elements is needed at the core in order to capture the correct cusp behavior of the Coulomb potential. Further research is needed to be able to conclusively determine which approach will be the most efficient and versatile one in dealing with the problem of the core singularity in the FEM approaches. It might well be possible in the future to devise a more elegant global basis set/FEM hybrid method which eliminates much of the current difficulties. Such a method would indeed be the holy grail of quantum chemistry and physics calculations. The Hartree-Fock method forms the basis for many other electronic structure methods including electron correlation effects and the extension of this work to calculations of electronic structure of more complex molecules using post-HF methods seems feasible. There's no conceptual or practical obstacle with the implementation of the method in the context of density functional theory, including hybrid functionals which require the calculation of exact exchange.

3.5. Figures and tables

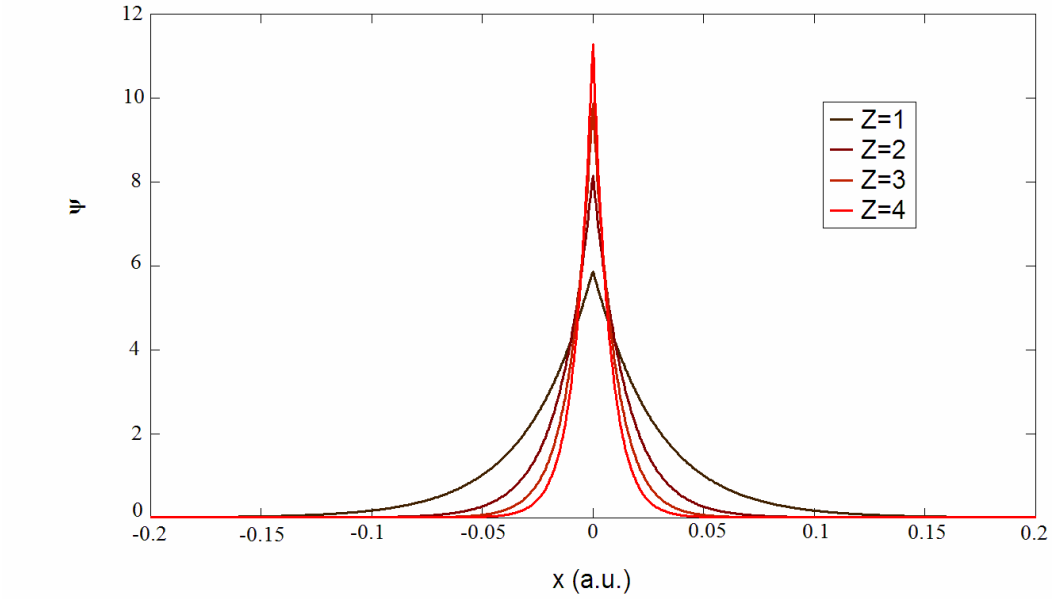


Figure 3.1. Ground state of the 1D infinite square well potential with a point charge Z in the middle for some different values of Z.

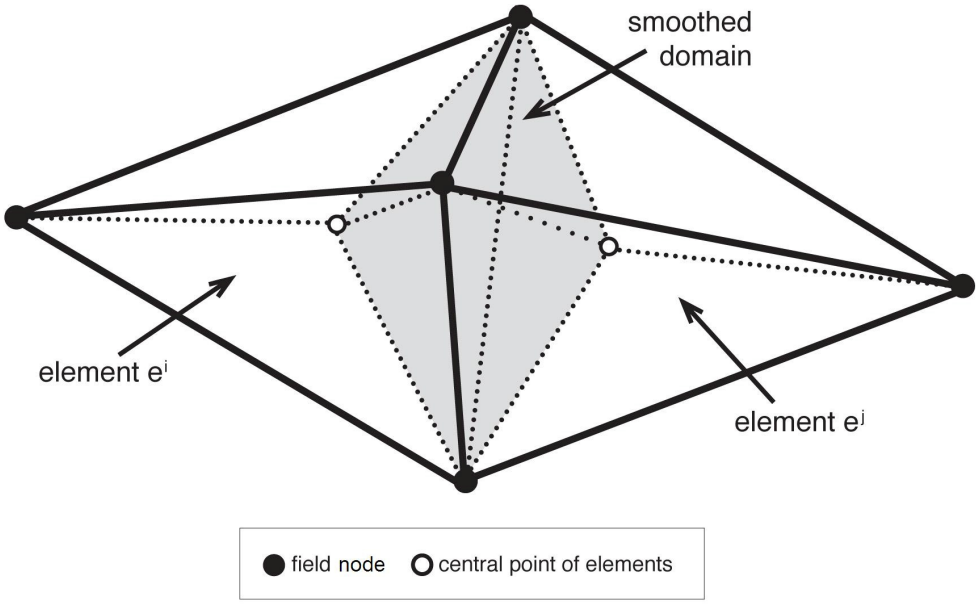


Figure 3.2. The definition of elements e^i and e^j and the smoothed domain in FS-FEM.

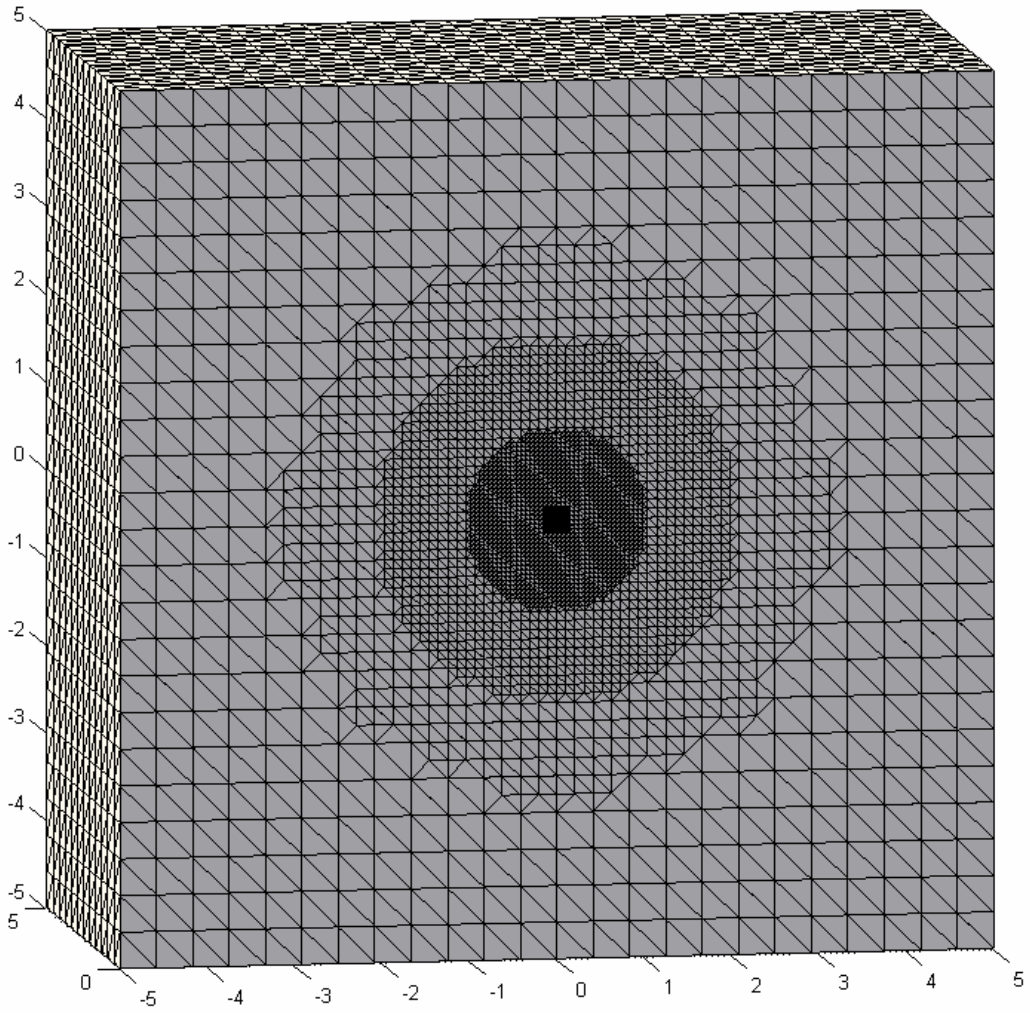


Figure 3.3. An example of the adaptive structured mesh. Element sizes are halved at regular intervals.

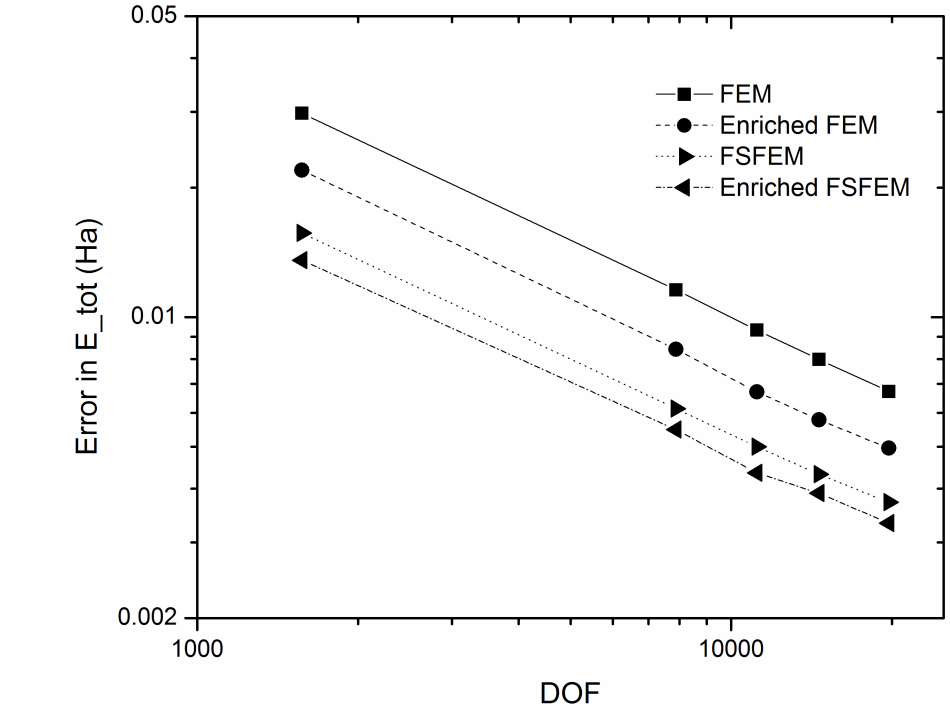


Figure 3.4. Convergence behavior of the FEM, enriched FEM, FS-FEM and enriched FS-FEM solutions for the helium atom; the error in the total energy as a function of total number of DOF.

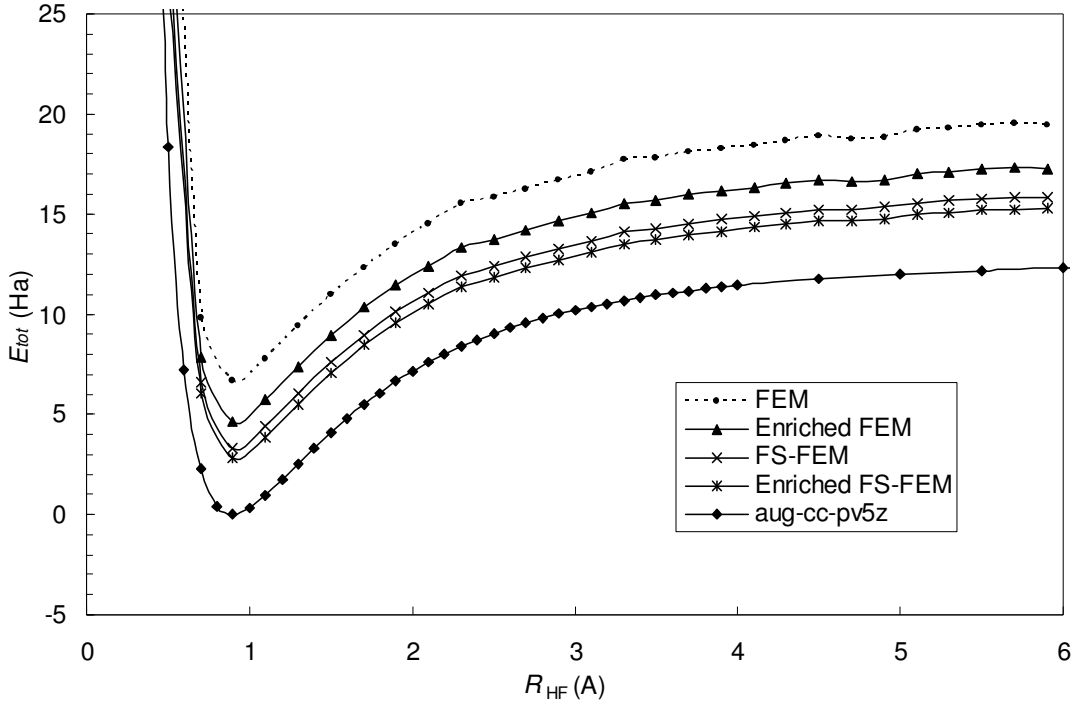


Figure 3.5. Bond dissociation curve for hydrofluoric acid from the Gaussian basis set and the different FEM approaches.

CHAPTER 4: ELECTRONIC STRUCTURE SIMULATIONS OF ELECTRO-MECHANICAL COUPLING PROBLEMS USING THE FINITE ELEMENT METHOD

4.1. Introduction

As electronic devices continue to shrink from micron size to submicron and nanoscales, new approaches are required to precisely take into account the associated quantum transport phenomena. Developments of future nano and molecular electronics systems (nanowires, nanotransistors and nano-electro-mechanical systems - NEMS) demand reliable predictions of the behavior of these systems at the quantum level. This requires a thorough knowledge of the electronic structure of these systems. This is a major challenge to physics and computational science community. Electronic structure calculations have been one the most computationally intensive simulations ever carried out by researchers. Millions of lines of code have been written by different research groups around the world in different quantum chemistry and physics software packages and hundreds of millions of CPU hours have been spent solving them. However, finding methods that meet both the challenge of reliable physics yet reasonable computation time is an open question. Furthermore, existing techniques do not always work well, e.g.

in the cases of adsorption/desorption of H_2 molecule on Si(001) or of simulations of experimentally observed quantum corrals shown in Figure 4.1 [96].

Another major challenge in electronic structure simulations is the electro-mechanical coupling phenomena. In some nano materials/structures, as the material deforms mechanically, its electronic properties change dramatically. An example of such phenomena was observed in carbon nanotubes (CNT) by Tombler *et al*. [97] Using tapping-mode atomic force microscopy (AFM) technique to bend and stretch a single walled CNT (Figure 4.2a), Tombler, *et al*. found that the electrical conductance of the nanotube could change by two orders of magnitude (Figure 4.2b), i.e., a metallic nanotube will become a semiconducting one after deformation. Moreover, this change of electric conductance is completely recoverable once the large mechanical deformation is removed. This unique electromechanical characteristics makes carbon nanotubes a good candidate for nano-electro-mechanical systems (NEMS). But being able to simulate and predict such coupling phenomena remains elusive to the scientific research community.

There are three basic *ab initio* methods to calculate electronic structures: the plane wave (PW) and grid methods, localized atomic(-like) orbitals method, and atomic sphere method [2]. The finite-element-based methods (FEM) belong to the first category. In the FEM, the *shape functions* are strictly local piecewise polynomials in real-space. The method is completely general and its convergence is guaranteed by variational principles, and thus achieved systematically. FEM method is the overwhelmingly preferred method for mechanical deformation simulations of materials, and is well suited

for parallel computing. The method thus combines all the major advantages of both the grid-based and basis-oriented approaches and is particularly promising for large-scale, accurate *ab initio* calculations of electro-mechanical coupling problems [3-6, 80].

Recently we have developed a reduced scaling local density three-dimensional FEM Hartree-Fock (HF) method to calculate electronic structures [80]. The Hartree-Fock (HF) theory is one of the cornerstones of electronic structure calculations. Significant portion of current qualitative as well as quantitative knowledge in theoretical chemistry has been gained by this technique. There are many post-HF methods that systematically build upon the accuracy of the HF technique, among them the configuration interaction (CI) method [22], the coupled-cluster (CC) method [23-24] and Møller-Plesset perturbation theory (MP) [13]. HF is essentially free from any assumptions on the forms of the density functional and includes an exact expression for the electron exchange which is the most important portion of the exchange-correlation (XC) energy. There have been efforts to incorporate this unique feature of the Hartree-Fock theory into the density functional theory (DFT) by using hybrid functionals [16-17]. Hybrid functionals (such as PBE0) have proven to be superior to the generalized gradient approximations (GGAs) in DFT.

4.2. Electronic structure calculations using FEM-HF method

The original HF theory was first introduced by D. R. Hartree [98] and is known as the self-consistent field (SCF) method. This method was later improved by V. A. Fock [99] to include the effects of electron exchange. In the Hartree-Fock self-consistent field

theory, the Hamiltonian matrix is formally replaced by the Fock matrix, which is the effective one-electron Hamiltonian. Writing the wavefunction as a product of spin-orbitals (which depend on the spatial and spin coordinates of one electron), the following equation for these spin-orbitals similar to the Schrödinger equation is obtained:

$$\hat{f}\psi_k(\mathbf{r}) = \varepsilon_k \psi_k(\mathbf{r}), \tag{4.1}$$

where ε_k are the eigenvalues and ψ_k are the eigenfunctions, $\hat{f}$ is the Fock differential operator defined by,

$$\hat{f} = -\tfrac{1}{2}\nabla^2 + V_0(\mathbf{r}) + V_d(\mathbf{r}) + V_x(\mathbf{r}), \tag{4.2}$$

where V_0 contains the effect of all interactions in the electron potential except those arising from the other electrons, and V_d is the Hartree potential which is the electrostatic potential generated by the charge distribution due to the wavefunction,

$$V_d(\mathbf{r}) = \sum_{j \neq k} \int \psi_j^*(\mathbf{r}')\psi_j(\mathbf{r}') \frac{1}{|\mathbf{r}' - \mathbf{r}|} d^3 r'. \tag{4.3}$$

Rather than calculating this integral directly, V_d is usually obtained by solving Poisson's equation invoking the Green's functions theory:

$$\nabla^2 V_d(\mathbf{r}) = -4\pi n(\mathbf{r}), \tag{4.4}$$

where the density of electrons n(r) is given by

$$n(\mathbf{r}) = \sum_{j=1}^{N} \left|\psi_j(\mathbf{r})\right|^2 , \tag{4.5}$$

where N is the total number of electrons. The $V_x(\mathbf{r})$ term in Equation (4.2) is the exchange potential and can be computed using the following relation

$$V_x(\mathbf{r})\psi_k(\mathbf{r}) = -\sum_{j\neq k} \left(V_x\right)_{jk} \psi_j(\mathbf{r}) , \tag{4.6a}$$

where

$$\left(V_x\right)_{jk} = \int \psi_j^*(\mathbf{r}')\psi_k(\mathbf{r}') \frac{1}{\left|\mathbf{r}'-\mathbf{r}\right|} d^3r' . \tag{4.6b}$$

It is noted that, when the Schrödinger equation is solved for a system, the result is a set of eigenvalues and their corresponding eigenvectors that are the quantum mechanical descriptions of all the allowed states of that system. The Hamiltonian operator is independent of the eigenfunctions ψ_k whereas the Fock operator depends on ψ_k via electron repulsion and correlations. Therefore, unlike in solving the Schrödinger equation, the solution of the Hartree-Fock equation requires a self-consistent iterative approach. The computational procedure is as follows. As the first step, a reasonable initial guess has to be provided. This initial guess can simply be the linear combination of the atomic orbitals (LCAO) of the system [100]. Based on this initial guess, one can calculate the electronic density and solve the Poisson's equation and thence, calculate the Fock operator and solve for a new set of eigenstates and eigenvalues. Then the

energies can be calculated and this process is repeated until the energies become stationary.

The FEM procedure to solve electronic structure problems is briefly outlined here. The goal is to solve the normalized single-electron equation in Ω which is the computation domain. Taking the inner product of the differential Equation (4.1) with an arbitrary *test function* ψ' to form an equivalent integral equation and integrating by parts (assuming natural boundary conditions), and performing the standard discretization procedure of the finite elements, we arrive at the following matrix equation for determining the wavefunctions:

$$\mathbf{H}_k \mathbf{a}_k = \mathbf{v}_k , \quad (k=1,\ldots,N/2), \tag{4.7}$$

where

$$\mathbf{H}_k = \int_\Omega \left\{ \tfrac{1}{2} \nabla \mathbf{N} (\nabla \mathbf{N})^T + (V_0 + V_d - \varepsilon_k) \mathbf{N}\mathbf{N}^T \right\} d\Omega \text{ , and}$$

$$\mathbf{v}_k = \sum_{j \neq k} \left\{ \int_\Omega (V_x)_{jk} \mathbf{N}\mathbf{N}^T \, d\Omega \right\} \mathbf{a}_j , \tag{4.8}$$

where **N** are the shape functions and **a** are the nodal degrees of freedom (DOFs). The standard Gram-Schmidt orthogonalization process is used to orthogonalize all the wavefunctions. We also require that the obtained wavefunctions to be normalized, i.e. $< \psi_k \mid \psi_{k'} >= \delta_{kk'}$ where $\delta_{kk'}$ is the Kronecker delta function. The quantity represented by the angled parentheses is the inner product of the sates *k* and *k'* in the Dirac bra-and-

ket notation. We found out that the bi-conjugate gradient stabilized iterative solver is particularly suitable for solving this equation.

4.3. The non-equilibrium Green's function (NEGF) methodology

The non-equilibrium Green's function (NEGF) formalism is widely used in nanodevice analyses. This technique can be applied to study the effect of mechanical deformation/defects on the conductivity of molecules and small carbon nanotubes (CNTs) to be used in NEMS and nanoelectronics products. Since their identification in 1991, CNTs have been the focus of extensive research. CNTs unique mechanical and electrical properties suggest potential applications in broad areas of science and technology. Since certain electrical properties of CNTs are of great prospective technological importance, they have been studied from many different viewpoints by different researchers. We would like to model and simulate the electronic and transport properties of carbon nanotubes with the finite element *ab initio* technique described in the previous section and compare the results with the other theoretical results and eventually with the experimental results. In particular, we are interested in taking into account the effects of mechanical deformations, defects and higher temperatures. Potential results of these simulations can be used in designing CNT-based electronic components and smart composite materials with electro-mechanical coupling to be used in sensors.

We employ the Born-Oppenheimer approximation which states that the electronic wavefunction depends only on the positions of the nuclei and not on their momenta. One

critical step in the atomistic studies of the electrical property change due to mechanical deformation or defects is the determination of the atomic positions. A real-space tight-binding (TB) energy minimization scheme based on the Xu *et al.* parameterization for carbon [101] and Polak-Ribiere conjugate gradient method (CGPR) were used for this purpose in this study. The semi-empirical tight-binding approximation or LCAO (linear combination of atomic orbitals) is quite good for the inner electrons of atoms, but it is not a good description of the conduction electrons themselves. Especially that the conventional TB fails for charged systems, systems with defects and dangling bonds and therefore cannot be directly used for this purpose. Although there has been some progress in tackling these kinds of systems with TB in the recent years, for example see Ref. [102].

The typical system considered here is shown schematically in Figure 4.3, which consists of two contacts as source and drain (boxes 1 and 2) and a molecule attached in between. The electric current of such a system in the coherent transport regime is given by [103]

$$I = (e/2\pi\hbar)\int_{-\infty}^{+\infty} dE\tilde{I}(E), \tag{4.9}$$

where $\tilde{I}_1$ is the inflow minus outflow for the first terminal:

$$\tilde{I}_1 = \mathrm{Tr}[\mathbf{\Sigma}_1^{\mathrm{in}}\mathbf{A}] - \mathrm{Tr}[\mathbf{\Gamma}_1\mathbf{G}^{\mathrm{n}}]\ . \tag{4.10}$$

In this equation:

$$\mathbf{\Sigma} = \mathbf{\Sigma}_1 + \mathbf{\Sigma}_2 ,$$

$$\mathbf{\Sigma}_{1,2} = \boldsymbol{\tau}_{1,2}\mathbf{G}_{1,2}\boldsymbol{\tau}_{1,2}^{\mathbf{H}} ,$$

$$\mathbf{\Sigma}^{\text{in}} = \mathbf{\Sigma}_1^{\text{in}} + \mathbf{\Sigma}_2^{\text{in}} ,$$

$$\mathbf{\Sigma}_{1,2}^{\text{in}} = \mathbf{\Gamma}_{1,2} f_{1,2} , \tag{4.11}$$

$$\mathbf{\Gamma}_{1,2} = i[\mathbf{\Sigma}_{1,2} - \mathbf{\Sigma}_{1,2}^{\mathbf{H}}] ,$$

$$\mathbf{A} = i[\mathbf{G} - \mathbf{G}^{\text{H}}] ,$$

$$\mathbf{G}^{\text{n}} = \mathbf{G}\mathbf{\Sigma}^{\text{in}}\mathbf{G}^{\text{H}} .$$

$\mathbf{G}$ is the retarded electronic Green's function of the molecule which is given by

$$\mathbf{G}[(E + i0^{+})\mathbf{S} - \mathbf{H} - \mathbf{\Sigma}] = \mathbf{I} , \tag{4.12}$$

and the advanced Green's function $\mathbf{G}^{\text{H}}$ is the Hermitian adjoint of the retarded Green's function. $\mathbf{H}$ and $\mathbf{S}$ are the Hamiltonian matrix and the overlap matrix of the conductor (channel), respectively. $\mathbf{\Gamma}_{1,2}$ are called the broadening matrices, $\mathbf{A}$ the spectral function and $\mathbf{G}^{\text{n}}$ the correlation function. E is the energy of the external excitation (i.e. the incident electrons) and 0^{+} is an infinitesimal positive number whose precise value is unimportant. $\mathbf{\Sigma}_{1,2}$ are the self-energies for the contacts and $\mathbf{\Sigma}_{1,2}^{\text{in}}$ are the inscattering functions, where the subscripts 1, 2 refer to the left and right contacts. It should be noted that in the coherent transport regime the inelastic phase-breaking processes within the

molecule are assumed to be negligible. The matrices $\boldsymbol{\tau}_{1,2}$ describe the coupling between the contacts and the molecule. $f_{1,2}$ are the usual Fermi functions of the contacts:

$$f_{1,2} = \frac{1}{e^{(E-\mu_{1,2})/k_B T} + 1}. \tag{4.13}$$

Finally, the applied bias can be computed via $eV = \mu_2 - \mu_1$.

4.4. The current state of the project

In Reference [80] it has been shown that this three-dimensional HF-FEM method is accurate and robust for the selected test atoms and molecules. These results also show how reduced scaling could be achieved through locality arguments, as well as the enforcement of the boundary conditions. An example of the calculations of large, complex molecules, namely a graphene sheet and a carbon nanotube is given in Figure 4.4 to illustrate the capability of the method. In order to take advantage of the facile convergence characteristics of the FEM, it was originally intended to extend this methodology to determine the electrical conductivity of CNT systems such as the one shown in Figure 4.5 and then study the electro-mechanical coupling properties discussed above at the Hartree-Fock level of theory [104-108]. The aforementioned Green's function methodology requires the inversion of very large complex sparse matrices given by Equation (4.12) many times. We have implemented this method in a Matlab script. Unfortunately this FEM code was computationally too demanding and slow and limited to the amount of the memory available on one computer as it's not been

parallelized. Therefore we were not able to solve any realistic systems with it as we intended. However, while continuing working on this project and trying to improve its efficiency, another group was able to solve this problem from a very similar approach with DFT [109].

Studying the phenomena of electro-mechanical coupling requires interdisciplinary approaches involving ideas and concepts drawn from physics, chemistry, mechanical engineering, electrical engineering, materials science, and computer science. To accelerate progress in this area, cyber-tools and cyber-infrastructure play an extremely critical role. Pursuing research in exciting and interdisciplinary areas requires truly outstanding facilities and active researchers to collaborate with that can promote research and development in these emerging technologies such as the calculations of electro-mechanical coupling phenomena in nanoscale materials and structures. Therefore the availability of powerful computational resources plays a decisive role in the chances of success of these types of projects in the future.

4.5. Figures and tables

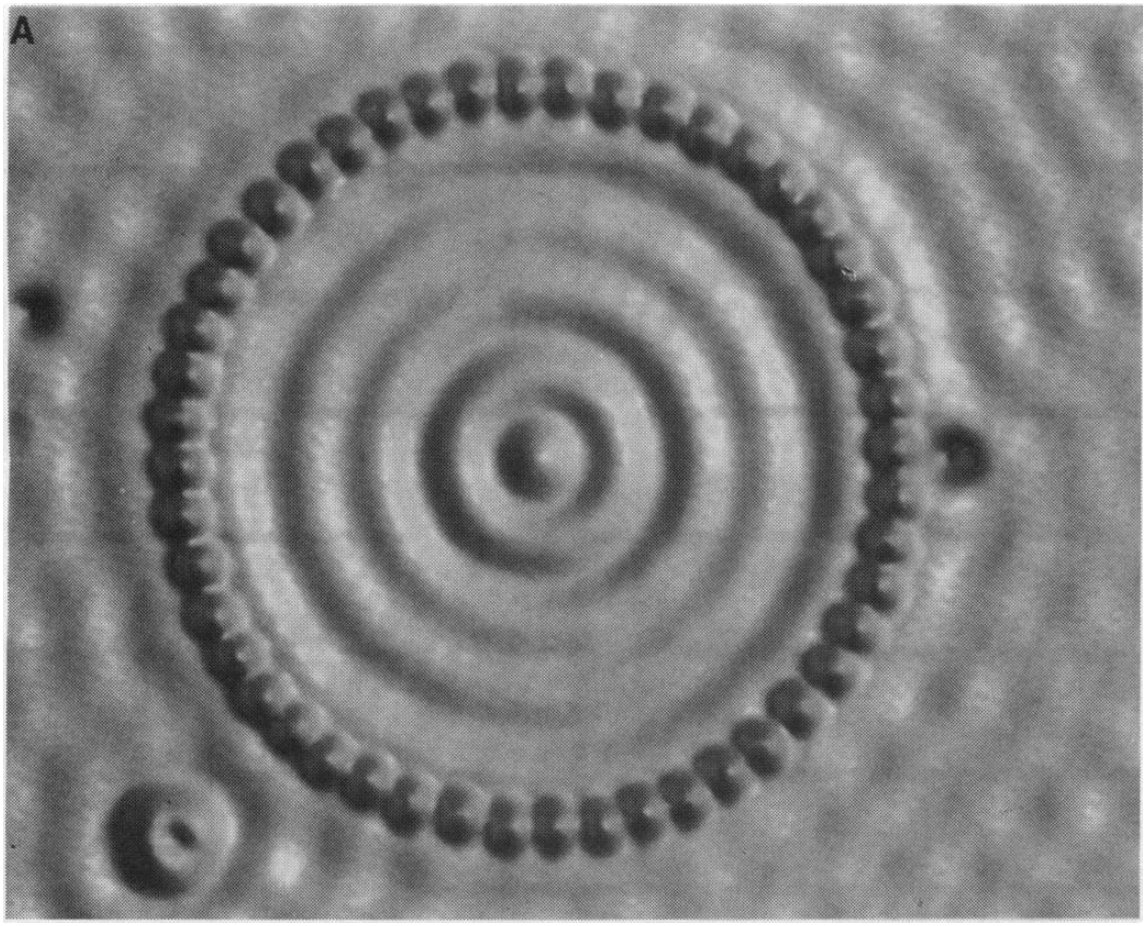
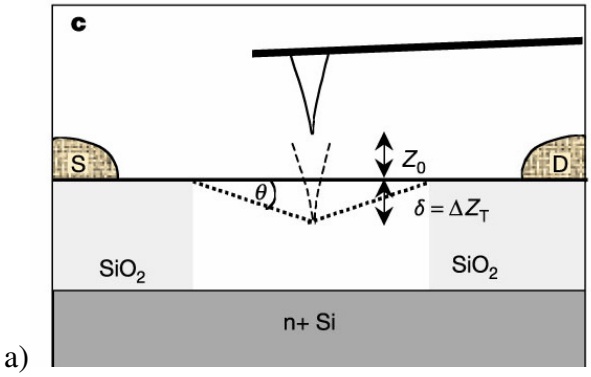
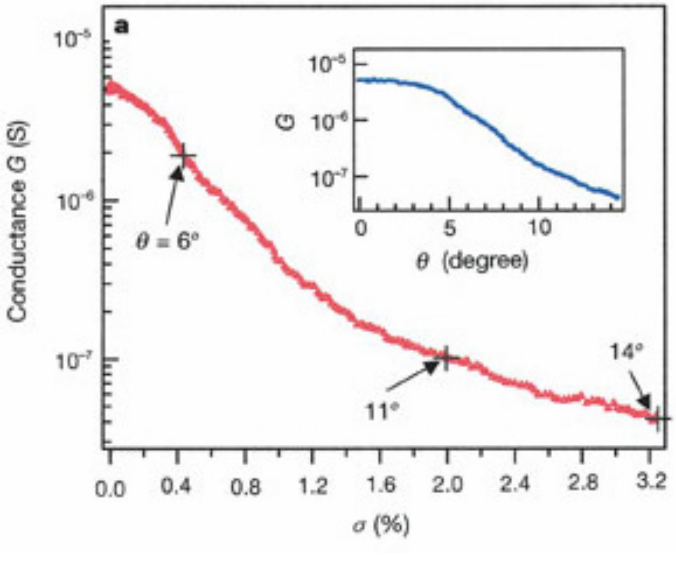


Figure 4.1. 48-atom Fe ring constructed on the Cu(111) surface. Average diameter of ring (atom center to atom center) is 142.6 Å. (Ref. [96])



a)



b)

Figure 4.2. (a) An atomic force microscope pushing a suspended single-wall carbon nanotube attached to two metal electrodes. (b) Electrical conductance versus mechanical deformation for a single-wall carbon nanotube. (Ref. [97])

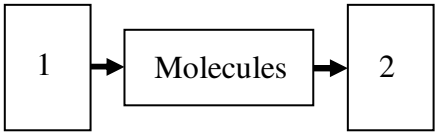


Figure 4.3. Schematic of the source, conductor and drain.

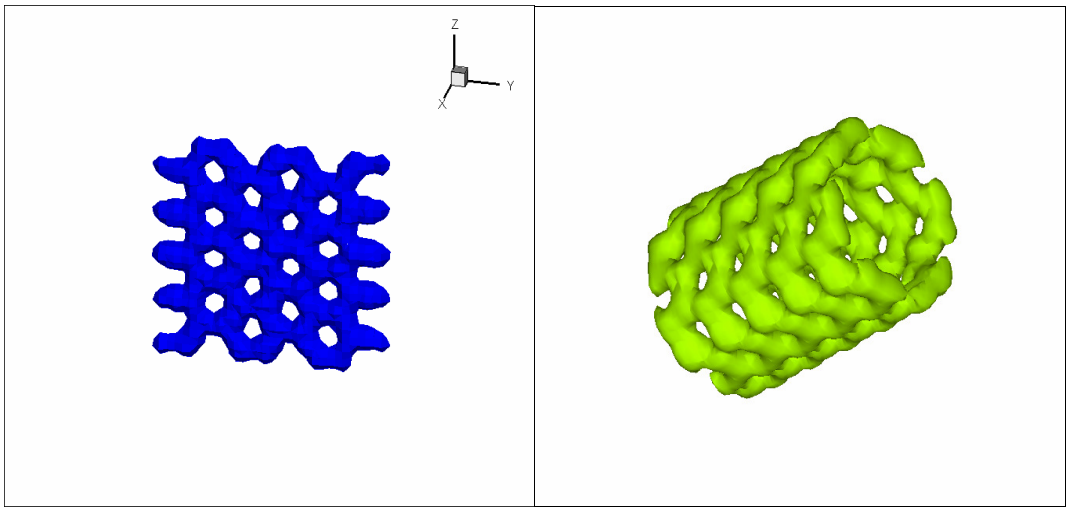


Figure 4.4. The electronic structure of a graphene sheet (left box) and a (6,6) armchair carbon nanotube (right box).

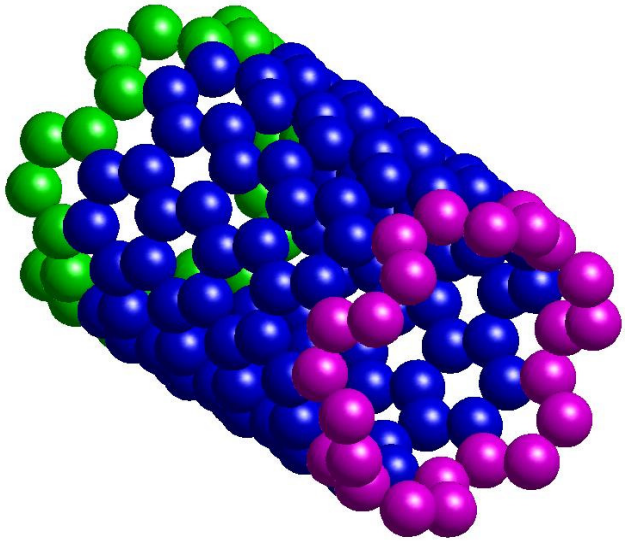


Figure 4.5. Definition of source (green), conductor (blue) and drain (purple) in a typical (6,6) armchair CNT.

CHAPTER 5: THE EFFECT OF DIFFERENT PHYSICAL PARAMETERS ON THE DISPERSIBILITY OF CARBON NANOTUBE BUNDLES

5.1. Introduction

Dispersion of carbon nanotubes is a very important step in all the diverse prospective applications of individual carbon nanotubes (CNTs) such as nanowires, the "smoothest bearing" and "gigahertz oscillators" and thus, has been a focus of extensive research. [110-117] Dispersion of single-wall carbon nanotubes (SWCNTs) is an important first step toward many potential applications that harness the unique electronic, thermal, optical, and mechanical properties of the individual tube. Furthermore, conventional applications such as the use of SWCNTs as conductive fillers in composites will benefit from a well-dispersed system that will exhibit the same conductivity with a smaller percentage of filler. Well-dispersed SWCNTs in a suitable solvent is also the preferred state for chemical modification of SWCNTs. Dispersibility of CNT bundles has an intimate relationship with tribological behavior of the CNTs. Tribology is concerned with the adhesion and friction behavior when two material surfaces are brought into close contact with each other. The classical relation between the frictional force and the normal pressure was described by the Amonton's law which

dates back to some 300 years ago. The Amonton's law simply stated that the frictional force increases monotonically (or more restrictively, linearly) with the normal pressure. The mechanistic explanation of the Amonton's law can be offered by a mechanism of interlocking asperities that leads to a contact area which increases with the normal pressure. That is basically due to the substantial increase in the area of intimate contact as the pressure increases with an increase in the stored elastic energy. Even if the surface pairs in contact appear to be perfectly smooth, such as between two clean crystal surfaces, small hydrocarbon or biological molecules would be absorbed on any surface exposed to air. These molecules can arrange to lock two contacting surfaces, and consequently induce static friction similar to the prediction of the Amonton's law.

The discovery of carbon nanotubes in 1991 by Iijima [118] presents new possibilities of extraordinary technological usages. With a supreme mechanical behavior such as 1/6 of the weight of the steel but an extension elastic modulus around 1TPa and an effective radial modulus in the range of 0.3 to 4 GPa (namely a factor of one thousand less than axial Young's modulus) the carbon nanotubes exhibit excellent strength, toughness and a formidable failure strain. [119] The diameter of a single wall carbon nanotube (denoted by D hereafter) ranges from 0.4 nanometer to several nanometers, and that for a multiwalled nanotube could be from 2nm to 100nm. Two aligned carbon nanotubes adhere to each other when brought into a pressureless contact. The aligned nanotube bundles however, exhibit a detachment work that declines with the contacting pressure. In other words, in contrast to the Amonton's law, the frictional force between carbon nanotubes has a zero slope and a downward curvature with respect

to the normal pressure and would decline as the normal pressure increases. This phenomenon falls into the general domain of *contact and adhesion problems* and can be viewed in that context. [120-123]

5.2. Methodology and Results

In this section we present the effect of various geometric parameters on the equilibrium configuration of the SWCNTs and eventually on dispersion properties of CNT bundles. For each parameter a computational setup has been constructed that is explained in its corresponding part. For simplicity, the attention herein is focused on the contact, adhesion and friction of identical single-wall carbon nanotubes (SWCNT) bundles. All nanotubes share the same uniform diameter D, and their centers are separated by the same distance L. For the case of non-cylindrical (i.e. deformed) CNTs, L represents the distance of the center of mass of the adjacent tubes.

5.2.1. The effect of distance

To study the effect of the centerline-to-centerline distance of CNTs on their dispersion behavior, we consider only infinitely long parallel single-wall carbon nanotubes as shown in Figure 5.1. Based on the Lennard-Jones potential one can obtain an ideal equilibrium distance for aligned carbon nanotubes without respect to local deformations as follows. The Lennard-Jones (LJ) 6-12 potential is given by the equation below

$$V = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right]. \tag{5.1}$$

In this well-known equation, ε is an energy parameter, σ is a length parameter and r is the interatomic distance. For long parallel cylindrical single-wall carbon nanotubes the collective van der Waals interaction per unit length of two tubes with diameters D_1 and D_2 can be approximated by the triple integral

$$V_{vdw} = \int_0^{\infty}\int_0^{\pi}\int_0^{2\pi} \frac{64\varepsilon D_1 D_2}{27 a_{C-C}^4}\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] d\theta_1 d\theta_2 dz\,, \tag{5.2}$$

in which a_{C-C} is the C-C bond length of carbon nanotubes; roughly equal to 1.44 Å and r is given by

$$r = \sqrt{\left[L + \frac{1}{2}\left(D_2\cos\theta_2 - D_1\cos\theta_1\right)\right]^2 + \left[\frac{1}{2}\left(D_2\sin\theta_2 - D_1\sin\theta_1\right)\right]^2 + z^2}\,. \tag{5.3}$$

This integral can be readily evaluated numerically. (In fact the z part integration can be carried out analytically as has been done by Benedict, *et al.* [124]). For numerical purposes a finite value for the infinite integration bound (for instance 5-10 angstroms) suffices. This integral has been evaluated for a (10,10) armchair CNT with a diameter of 13.728 Å and the result is shown in Figure 5.2. It should be mentioned that this integral generally tends to give lower estimates of the total interaction potential due to ignoring the effect of local spikes of closely positioned atoms, but this difference is reduced by assuming a relaxed configuration of the neighboring nanotubes and also as L is increased so that it can safely be used to represent the collective van der Waals (vdW) interaction

between two aligned CNTs. To demonstrate this point, the same problem has been solved by direct calculation of the contributions of individual atoms and as it can be observed in the same figure, the results generally agree with that of the theoretical analysis, though the former is obtained with a much higher computational effort. Based on this mathematical model a binding energy can be calculated for the typical parameter values given in the Table 5.1. It should be noted that these values are for an ideal cylindrical geometry which will be shown in the next section that *do not* necessarily agree well with the other more sophisticated computational and experimental results.

For a unit-long strip of infinitesimal width *w* over - and parallel to - an infinite sheet of graphene (*xz*-plane) at a height *y* the collective van der Waals (vdW) interaction per unit length of the tubes can be approximated in the similar manner by

$$V_{vdw} = \int_0^\infty \int_0^\infty \frac{256\varepsilon w}{27 a_{C-C}^4} \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^{6} \right] dxdz \,, \tag{5.4}$$

in which a_{C-C}=1.44 Å and $r = \sqrt{x^2 + y^2 + z^2}$. This can easily be carried out to get

$$V_{vdw} = 32\varepsilon w \sigma^2 \pi \left[2(\sigma / y)^{10} - 5(\sigma / y)^4 \right] \Big/ 135 a_{C-C}^4 \,, \tag{5.5}$$

which has a similar form as the original LJ potential – except with different exponents and pre-factors. If a nanotube is modeled as a membrane with curvature modulus *k* and arbitrary cross-section shape defined parametrically by $f(t) = x(t)\vec{i} + y(t)\vec{j}$ where $t_1 \le t \le t_2$, then the curvature energy is

$$E_c = \frac{k}{2}\int \frac{ds}{\left[\rho(t)\right]^2} = \frac{k}{2}\int_{t_1}^{t_2} \frac{\left[x'y'' - y'x''\right]^2}{\left\{x'^2 + y'^2\right\}^{\frac{5}{2}}} dt\,, \tag{5.6}$$

where $\rho(t)$ is the radius of curvature and the total energy is just the sum of these two potential energies (i.e. vdW energy and curvature energy). In this equation, primes represent differentiation with respect to t. Thus using the standard techniques of variational calculus, we have to minimize the following functional

$$\begin{aligned} E &= \int_{t_1}^{t_2} \left\{ \frac{k}{2} \frac{\left[x'y'' - y'x''\right]^2}{\left(x'^2 + y'^2\right)^{\frac{5}{2}}} - \frac{32\varepsilon\sigma^6\pi\left(x'^2 + y'^2\right)^{\frac{1}{2}}(-2\sigma^6 + 5y^6)}{135\,y^{10}a_{C-C}^4} \right\} dt \\ &= \int_{t_1}^{t_2} F(y, x', y', x'', y'')dt \end{aligned}\,, \tag{5.6}$$

subject to

$$\begin{aligned} &\pi D = \int ds = \int_{t_1}^{t_2} \sqrt{x'^2 + y'^2}\, dt\,, \\ &x(t_1) = x(t_2)\,,\ y(t_1) = y(t_2)\,,\ x'(t_1) = x'(t_2)\,,\ y'(t_1) = y'(t_2)\,, \text{ etc.} \end{aligned} \tag{5.7}$$

The first equation is for preserving the total perimeter length of the tube and the rest are for ensuring closedness and smoothness of the resulting curve. Applying the Euler's theorem in the parametric form to this problem yields

$$F_{x'} - \frac{d}{dt}F_{x''} = C\,, \tag{5.8}$$

where C is an arbitrary constant that has to satisfy the inextensibility constraint as well as the closedness and smoothness requirements along with the other constants.

Substituting F and simplifying and finally numerically solving the resulting sixth order nonlinear differential equation with shooting method (taking t_1=0 and t_2=1 for convenience and k = 1.4 eV from Ref. [125]) we find the family of solutions plotted for a few typical values of D i.e. with different radii in Figure 5.3(a). These curves at least qualitatively are the same as those shown in Refs. [126-127]. It should be mentioned that an alternative way to accomplish this task is to consider the different deformation paths and use shape optimization techniques, but here variational approach deemed to be more appropriate. Since in this simplified analysis the intra-molecular interactions are neglected, no bifurcation type behavior such as collapse could be captured. This phenomenon will be discussed more fully in the ensuing chapters. Adhesion forces and energies can be calculated in the same fashion as was done in Eq. (5.2) and the fact that $F(y) = -\partial V / \partial y$. For example, the force between two parallel, infinitely long infinitesimal axial strips of width w can be shown to be

$$F(r) = \frac{\varepsilon w^2 \pi}{36 a_{C-C}^4} \left[-231 \left(\frac{\sigma}{r} \right)^{12} + 160 \left(\frac{\sigma}{r} \right)^{6} \right]. \tag{5.9}$$

From this equation, the vertical and horizontal components of the theoretical adhesion forces are calculated and shown for a typical CNT in Figure 5.3(b). From this figure it can vividly be inferred that almost all the vdW interaction is localized at the ridges of the tube and almost no interaction occurs in between. Binding energies are also calculated and listed in Table 5.3.

5.2.2. The effect of local deformation

Brenner's (1990) empirical multi-body interatomic potential for carbon [10] has been widely used in the study of CNTs, and is summarized below.

$$V(r_{ij}) = V_R(r_{ij}) - B_{ij}V_A(r_{ij}), \tag{5.10}$$

where r_{ij} is the length between two atoms i and j, V_R and V_A are the repulsive and attractive pair terms given by

$$\begin{aligned} V_R(r) &= \frac{D^{(e)}}{S-1} e^{-\sqrt{2S}\beta(r-R^{(e)})} f_c(r) \\ V_A(r) &= \frac{D^{(e)}}{S-1} S e^{-\sqrt{2/S}\beta(r-R^{(e)})} f_c(r) \end{aligned}, \tag{5.11}$$

where $D^{(e)}$ =6.000 eV, S =1.22, β=21 nm^{-1}, and $R^{(e)}$=0.1390 nm. The function f_c in the preceding equation is a smooth cut-off function given by

$$f_c(r) = \begin{cases} 1 & r < R^{(1)} \\ \frac{1}{2}\left\{1+\cos\left[\frac{\pi(r-R^{(1)})}{R^{(2)}-R^{(1)}}\right]\right\} & R^{(1)} < r < R^{(2)} \\ 0 & r > R^{(2)} \end{cases} \tag{5.12}$$

where the effective range of the cut-off function is defined by $R^{(1)}$ = 0.17 nm and $R^{(2)}$ = 0.20 nm. The term B_{ij} in Eq. (5.10) represents a multi-body coupling effect (i.e., the contribution from atoms other than i and j), and is given by

$$B_{ij} = \left[1+\sum_{k\neq i,j} G(\theta_{ijk}) f_c(r_{ik})\right]^{-1/2}, \tag{5.13}$$

where $k(\neq i, j)$ denotes the other carbon atoms, f_c is given in Eq. (5.12), θ_{ijk} is the angle between $i - j$ and $i - k$ bonds in Figure 5.4, and the function G is given by

$$G(\theta) = a_0 \left[1 + \frac{c_0^2}{d_0^2} - \frac{c_0^2}{d_0^2 + (1 + \cos\theta)^2} \right], \tag{5.14}$$

with $a_0 = 0.00020813$, $c_0 = 330$ and $d_0 = 3.5$.

Based on this potential we can also obtain the strain energy versus distance from quasi-static molecular mechanics (MM) calculations by an energy minimization technique called Polak-Ribiere Conjugate Gradient method (CGPR) since it is not possible to obtain any analytical or semi-analytical form for the deformed configuration of the nanotubes as was done in the previous section. In Figure 5.5 the total internal deformation potential energy per unit length of the tubes (which in turn is composed of many-body interaction and pair potential parts) for the same CNT is shown as a function of centerline-centerline distance. In order to exclude any finite length effects in the computations, a plane strain boundary condition on the outer rings of the nanotubes in the cell ranging from a few angstroms long to several times the diameter with frequent pair table updates has been enforced. Based on this model a binding energy can also be calculated for the typical parameter values given in the Table 5.2 for a few armchair nanotubes (more extensive tables can be found e.g. in Refs. [121, 128]). As it is apparent from this figure, the variation in deformation potential is small for relatively long ranges of L and thus it can be safely neglected compared to the van der Waals potential contribution i.e. the effect of local deformation is insignificant for large values of L, but

for closer ranges of contact it becomes more important as is also expected intuitively and it must be taken into account. It should be noted that generally the behavior of the potential function with deformation vs. distance is quite different from that of the simplified circular geometry; the tubes tend to flatten out locally at the contact surface with a local equilibrium distance close to the graphite inter-layer spacing (0.335 nm) with a much higher van der Waals interaction which in turn is responsible for the difficulty associated with the dispersion of carbon nanotube bundles. In fact we did similar calculations taking into account full van der Waals interactions between the individual atoms for a (30,30) armchair nanotube with a diameter D~4nm and found a binding energy on the order of γ=0.42 nN at an equilibrium distance of L=42.6 Å which is slightly higher than the experimental result of 0.36 nN given in Ref. [129] and roughly twice lower than the projected analytical value of 0.81 nN given in Ref. [128]. A typical relaxed deformed configuration of these nanotubes in shown in Figure 5.6.

5.2.3. The effect of diameter and finite length

Carbon nanotubes have exceptionally high length to diameter aspect ratios sometimes on the order of 10^7. This point plus their lateral flexibility makes them particularly susceptible to clotting. In order to quantify this phenomenon we may proceed as follows: using the same analytical expression for the total potential energy as in the previous section, first we can investigate the effect of having CNTs with different radii. It should be mentioned that D in that expression itself is a function of the chiral numbers (n,m) as given by

$$D = a/\pi \sqrt{n^2 + m^2 + nm}\,, \tag{5.15}$$

where a=2.49 Å is the lattice constant for carbon nanotubes. The total potential energy per unit length of the CNT as a function of D is calculated and shown in Figure 5.7. It is clear from this figure that as the diameter of the nanotubes increase, the total potential energy – which may be thought of as representing a measure of binding energy – decreases which means that dispersion becomes more difficult. This is somewhat in contrast to the experimental observations where bigger SWCNTs and MWCNTs are much easier to disperse than smaller diameter SWCNTs. Micron-long MWNT composites have been made successfully with the MWCNTs completely dispersed, whereas to best of our knowledge, the only group that has been able to completely disperse small SWCNTs is Prof. Strano's group who used to be at the Chemical Engineering Department at the University of Illinois. We believe the reason for this inconsistency is that for bigger nanotubes, even though the binding energy is higher, due to a larger surface to binding energy ratio, they are more exposed to dispersants and other chemical agents and thus more prone to being dispersed. Another reason that could be mentioned specifically in the case of MWCNTs is their higher bending stiffness due to a generally greater cross sectional moment of inertia which mechanically prevents them from easily tying up together and thus forming clots and clusters. Obviously shorter CNTs are less likely to wrap around each other and form knots.

5.2.4. The effect of orientation

For studying the effect of orientation of the nanotubes on dispersion, only computational approach is viable since no straight-forward analytical solution is possible as in the previous cases. The geometrical configuration used here is shown in Figure 5.8(a) at the equilibrium distance. The CNTs are rotated relative to each other while maintaining the ability to adjust their equilibrium distance. The angular dependence of the overall van der Walls interaction for a typical (10,10) SWCNT is shown in Figure 5.8(b). In this figure, θ is the alignment angle. For other types of CNTs, similar curves exist. It can be observed in this curve that there is a plateau of high negative overall vdW interaction for the small values of rotation angle and also, this potential is a monotonously increasing function of the alignment angle, therefore $\theta=0^o$ corresponds to the minimum potential energy and $\theta=90^o$ corresponds to the maximum potential energy. For sufficiently small θ (say less than 10^o for this particular arrangement) the tubes tend to snap back to a parallel configuration and realign themselves with each other, and above this critical angle of rotation, the interaction rapidly fades out; i.e. the returning force diminishes and cannot overcome the local frictional resistance.

5.2.5. The effect of intra-tubular van der Waals interaction

As Brenner pointed out, [10] the set of parameters in the interatomic potential was determined by fitting the binding energy and lattice constants of graphite, diamond and other possible atomic structures of pure carbon. The binding energy is directly related to the interatomic potential, while the lattice constants are the bond lengths at equilibrium,

which are directly linked to the first-order derivatives of the interatomic potential. Little or essentially no attention was paid to the second-order derivatives of the interatomic potential, but these second-order derivatives are crucial in the force-type behaviors of graphite and diamond such as their elastic moduli. In fact, for instance, the force-distance behavior of the Tersoff-Brenner potential [9-10] under equi-biaxial strain state is not smooth at the first cutoff radius as can be seen in Figure 5.9(b), even though the direct potential itself is smooth as shown in Figure 5.9(a). This issue can lead to unpredictable and/or erroneous results for stiffness-dependant properties of CNTs. Any way, one can also incorporate this elastic deformation with an intra-tubular vdW interaction to come up with improved force-deformation characteristics. However, this has little to do with the dispersibility characteristics of CNTs, and therefore this issue is no further pursued at this section. We just would like to mention here that these potentials are used in other chapters to study the collapse and stability characteristics of CNTs on SiO_2 substrates.

5.2.6. The effect of chirality

The effect of chirality on the dispersibility of nanotube bundles is minimal. We repeated some of the calculations for (n,0) zigzag and (2n,n) chiral CNTs as well as (n,n) armchair ones and did not observe any meaningful difference in the obtained results. Meanwhile it is worth mentioning that recently some registry-dependent interlayer potentials have been proposed that can more accurately take into account the effects of different chiralities and stacking orders. [115]

5.3. Discussion

Dispersing carbon nanotubes has proven to be quite challenging experimentally. For example, one way to do that is to use dispersive surfactants such as sodium dodecyl sulphate (SDS) which is an anionic detergent that denatures CNTs by *wrapping around* them. In so doing, SDS confers a negative charge to the CNT in proportion to its length which eventually may be able to exfoliate clots and clusters. Usually these chemicals are used together with some kind of physical dispersing process as a catalyst such as sonication and centrifuge. Sonication is the act of applying ultrasound energy to agitate particles in a sample to achieve a more uniform distribution and homogeneity. Based on our binding energy calculations, these dispersants in order to be able to successfully disperse carbon nanotubes, have to be able to confer charges on the order of and greater than 0.5 nC/m to overcome van der Waals forces between the parallel tubes. These forces can be calculated from summing up all the individual force components between the adjacent tubes – which of course the maximum attractive force happens at a distance larger than the equilibrium distance of the tubes where by definition the overall force goes to zero. Additionally, such a procedure gives maximum forces on the order of 1 N/m.

5.4. Conclusions

In this chapter, an attempt has been made (via analytical as well as computational simulations) to study the effect of different physical parameters on the dispersibility of

carbon nanotube bundles and to investigate some aspects of the dispersion process of the carbon nanotubes and the roots of the difficulty of achieving this goal. We found out that comparatively very strong long-range structural van der Waals interactions exist among carbon nanotubes in bundles as has also been reported both experimentally and theoretically by other researchers which clearly explains the natural propensity of the carbon nanotubes to agglomerate in the form of clots and clusters. We have come up with some quantitative suggestions to be able to resolve the dispersion problem of the CNT solutions.

5.5. Figures and tables

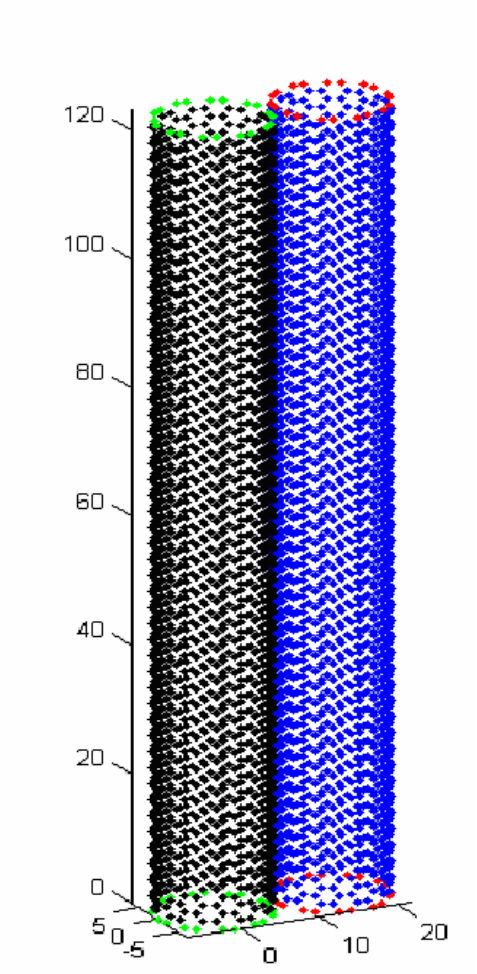


Figure 5.1. Geometric configuration of the carbon nanotubes studied in the perfect cylinder and deformed cases. The distance units of the figure are in angstroms.

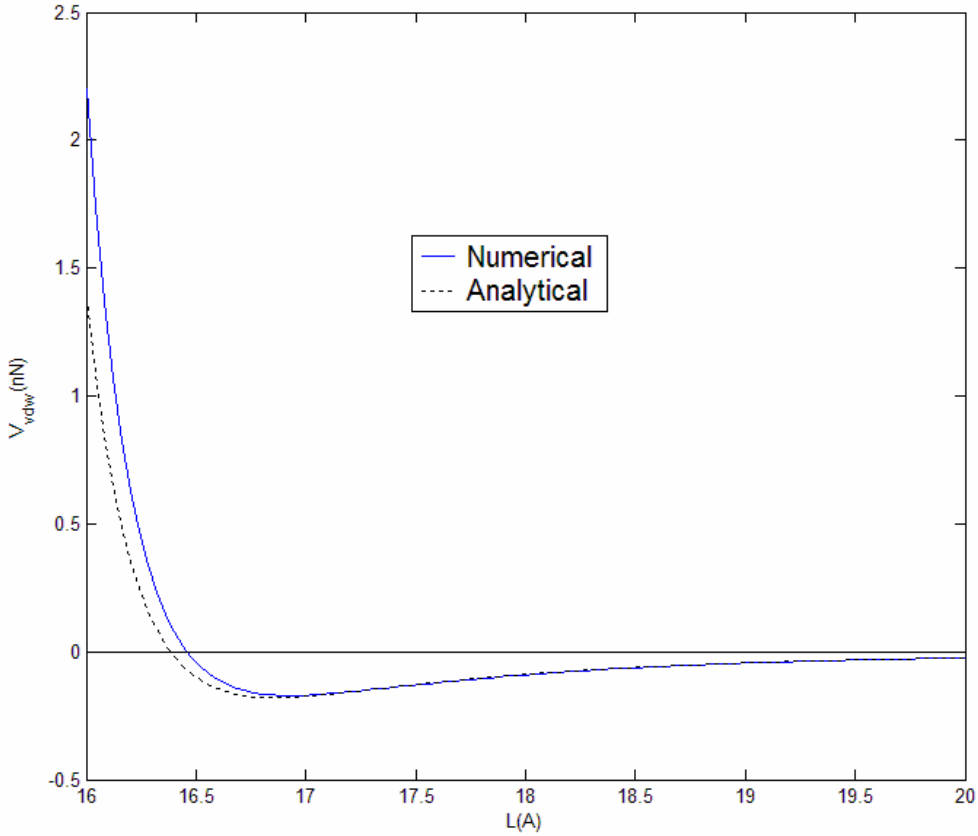
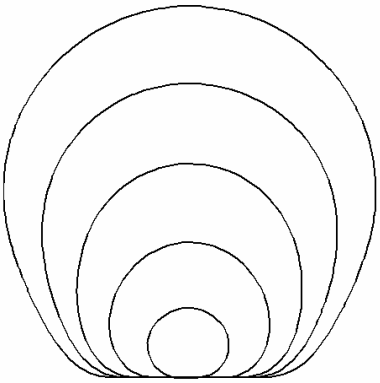
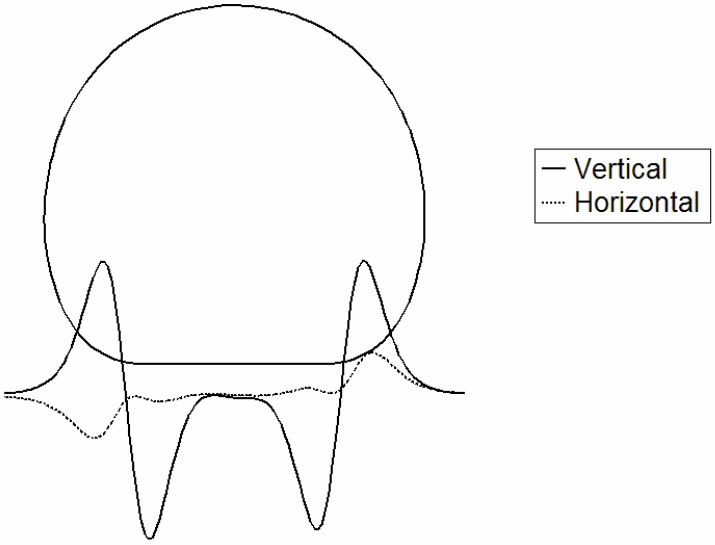


Figure 5.2. Van der Walls interaction of two infinitely long (10,10) armchair tubes in the perfect cylinder case from the analytical analysis (dashed line) and from the numerical analysis (solid line.)



a)



b)

Figure 5.3. (a) The equilibrium shapes of CNTs on an infinite graphene sheet with varying diameters from (10,10) to (50,50) superimposed on each other.

(b) The horizontal and vertical components of the force that a (30,30) CNT exerts on a graphene sheet beneath it.

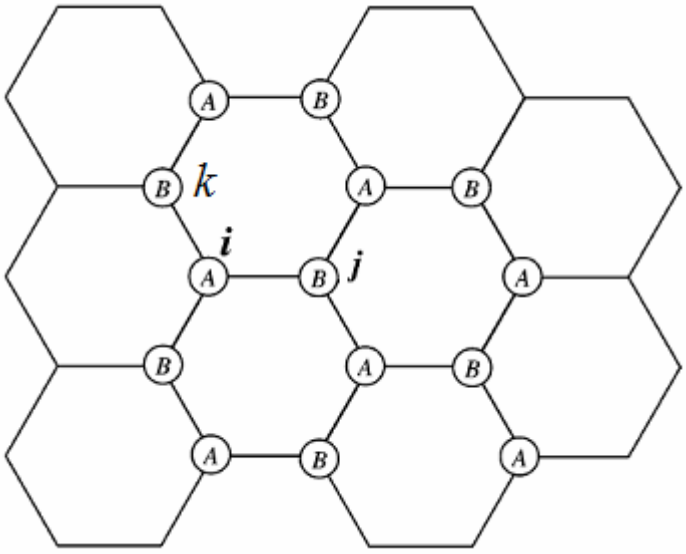


Figure 5.4. The definition of atoms *i*, *j* and *k* in the Brenner potential.

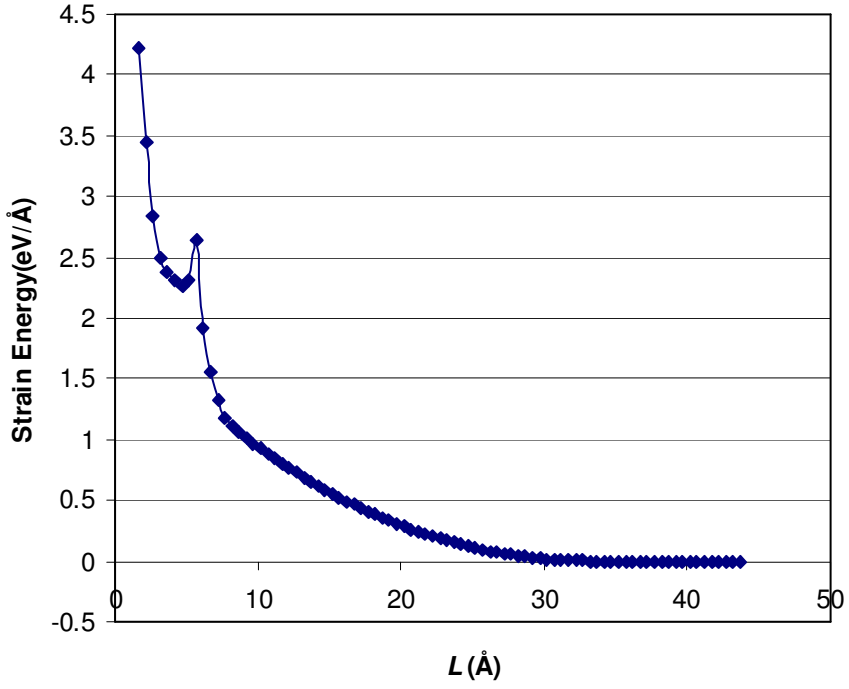


Figure 5.5. The strain energy as a function of distance between two (30,30) CNTs.

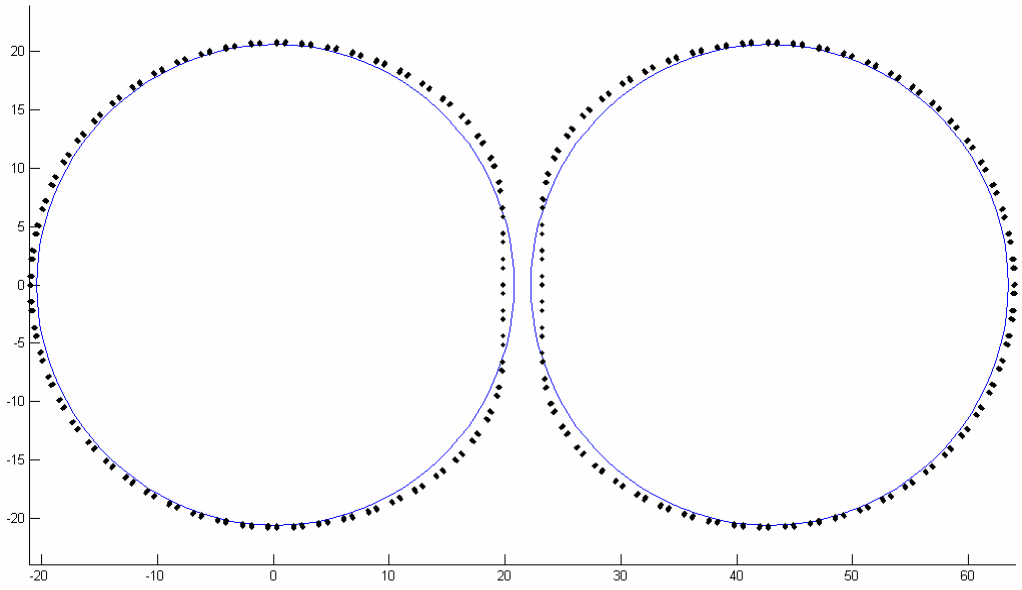


Figure 5.6. The equilibrium configuration of two parallel (50,50) CNTs. The circles denote the undeformed shape of each CNT before contact.

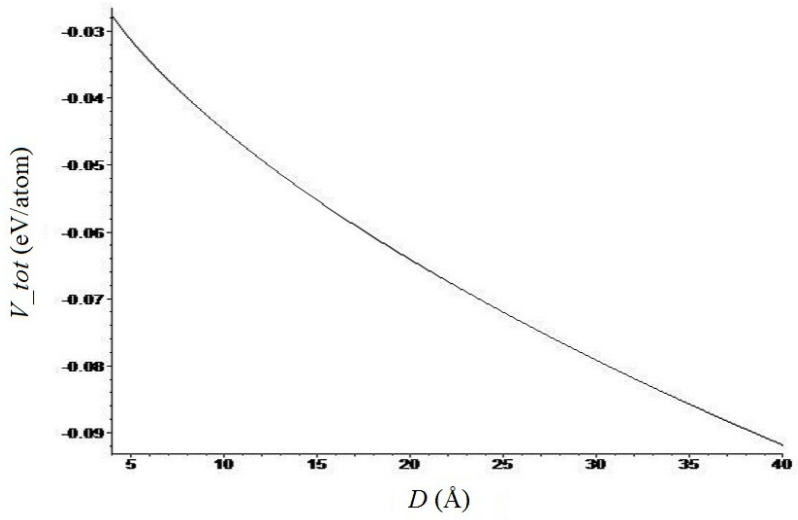
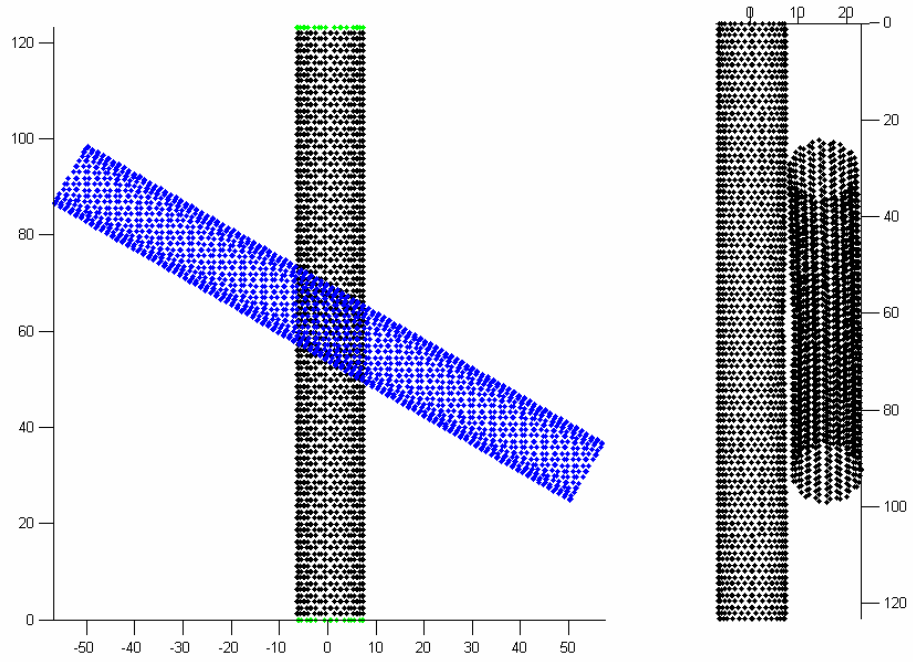
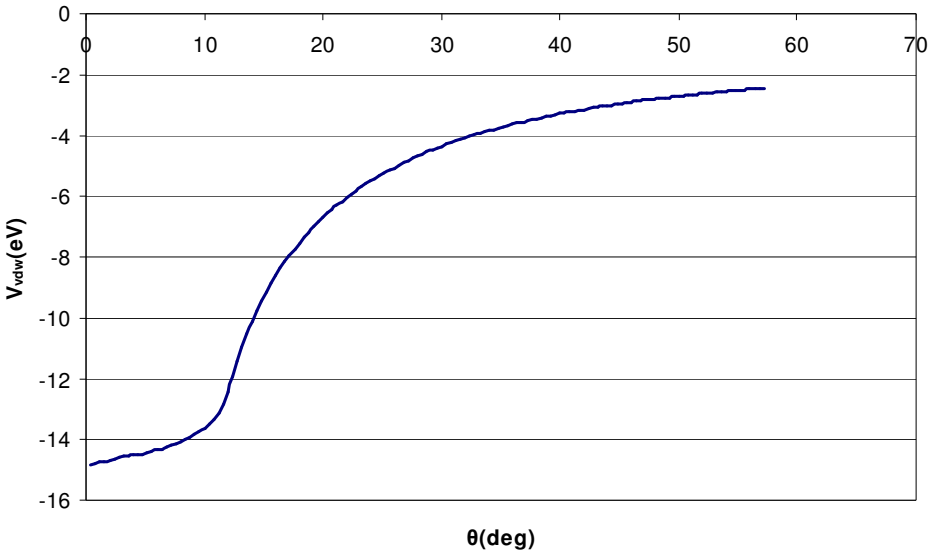


Figure 5.7. The total potential energy per atom of CNTs on a graphene sheet at the equilibrium distance as a function of the diameter of CNTs.

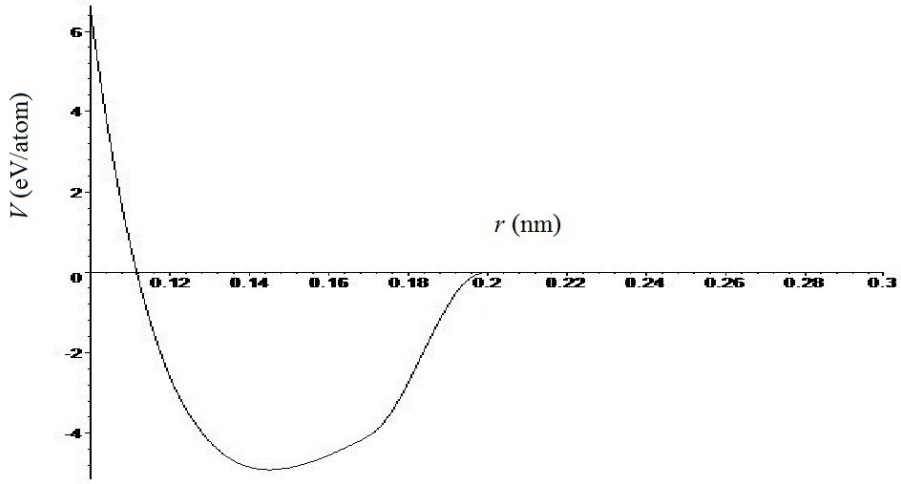


a)

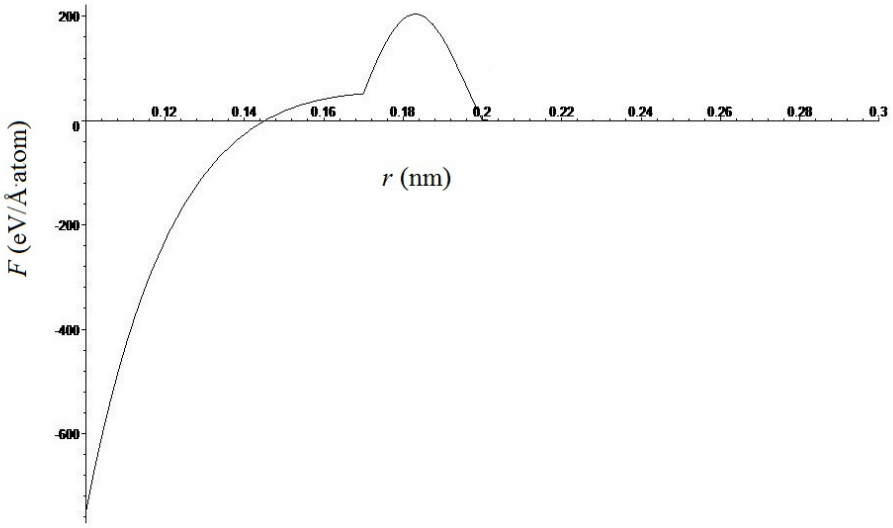


b)

Figure 5.8. (a) Different views of the rotation of two CNTs w.r.t. each other. (b) The variation of vdW potential with the angle of rotation for a ~13nm long (10,10) SWCNT.



a)



b)

Figure 5.9. (a) Dependence of the popular Tersoff-Brenner (TB) potential for carbon on interatomic distance. (b) While the potential curve is smooth for this potential, its force curve suffers from a slope discontinuity at some cutoff distance.

σ (nm)	ε (meV)	γ (nN) @ L(Å)			
		(10,10)	(15,15)	(20,20)	(25,25)
0.339	3.02	0.171 @ 16.9	0.216 @ 23.8	0.253 @ 30.6	0.286 @ 37.5
0.341	2.39	0.138 @ 17.0	0.174 @ 23.8	0.203 @ 30.7	0.229 @ 37.5

Table 5.1. The binding energy and the equilibrium distance in the infinitely-long cylindrical tubes model for typical armchair CNTs (compare with Table 5.2)

σ (nm)	ε (meV)	γ (nN) @ L(Å)			
		(10,10)	(15,15)	(20,20)	(25,25)
0.339	3.02	0.195 @ 16.5	0.271 @ 22.8	0.358 @ 29.5	0.440 @ 36.4
0.341	2.39	0.153 @ 16.6	0.208 @ 23.0	0.269 @ 29.7	0.331 @ 36.6

Table 5.2. The binding energy and the equilibrium distance in the infinitely-long deformed tubes model for typical armchair CNTs (compare with Table 5.1)

σ (nm)	ε (meV)	γ (nN) @ L(Å)			
		(10,10)	(20,20)	(30,30)	(40,40)
0.339	3.02	0.326 @ 9.4	0.642 @ 14.6	0.913 @ 20.8	0.979 @ 28.8
0.341	2.39	0.138 @ 17.0	0.174 @ 23.8	0.203 @ 30.7	0.229 @ 37.5

σ (nm)	ε (meV)	γ (nN) @ L(Å)
		(50,50)
0.339	3.02	1.120 @ 35.9
0.341	2.39	0.255 @ 44.0

Table 5.3. The binding energy and the equilibrium distance of infinitely-long deformed tubes on an infinite sheet of graphite for typical armchair CNTs

CHAPTER 6: THERMAL DISSIPATION AND VARIABILITY IN ELECTRICAL BREAKDOWN OF CARBON NANOTUBE DEVICES

6.1. Introduction*

Carbon nanotubes (CNTs) have excellent intrinsic electrical and thermal properties, and thus are being considered potential candidates for nanoscale circuits, [131] heat sinks [132] or thermal composites. [133] However, their physical properties depend on temperature, and thus are directly affected by power dissipation during electrical operation. [134-136] Joule heating in CNTs goes beyond degrading electrical performance, posing reliability concerns as in other electronics. Electrical Joule breakdown has also been used to remove metallic CNTs in integrated circuits; [137-139] however the technique is not precise, owing to the lack of fine control over CNT heat dissipation. It is presently understood that the thermal boundary conductance (TBC) at CNT interfaces with the environment, substrate, or contacts plays the limiting role in thermal dissipation. [140-142] In addition, the interaction of CNTs with the environment may also change their effective thermal conductivity. [143-144] However, little is currently known about the details of the thermal interaction between CNTs and common

* This chapter is based on Ref. [130]. The second author (R. A.) has had no contribution in the acquisition of the experimental results nor in the derivation of the DMM model.

dielectrics, including the roles of dielectric surface roughness or of CNT diameter and chirality (e.g. metallic vs. semiconducting).

In this study, we examine electrical breakdown and thermal dissipation of CNT devices with the most common interface used in integrated circuit experiments, that of SiO_2 as shown in Figure 6.1(a). We employ electrical breakdown thermometry [141, 145] to extract the TBC between CNTs and SiO_2 for metallic (m-CNT) and semiconducting nanotubes (s-CNT) of diameters $1 < D < 4$ nm. We find the TBC per unit length scales proportionally with CNT diameter, confirming recent simulation work. [146] We also find that m-CNTs appear to have better and more consistent thermal coupling with SiO_2 than s-CNTs, indicating a fundamental challenge for complete m-CNT removal in circuits via electrical breakdowns. We compare our results to both a diffuse mismatch model (DMM) and to molecular dynamics (MD) simulations. The latter reveal the role played by the thermal "footprint" of a deformable CNT on such dielectric substrates. Finally, we uncover the significant role of variability in threshold voltage (for s-CNTs) and of SiO_2 surface roughness (for both m- and s-CNTs) in heat dissipation and electrical breakdown.

6.2. Experiments and data extraction

We fabricated and conducted experiments on carbon nanotube devices in the same back-gated configuration as our previous work, using semi-circular electrodes for better CNT length control [147-149] (here, $2 \leq L \leq 5.6$ μm) as shown in Figures 6.1 and 6.2. The SiO_2 is thermally grown dry oxide, approximately 90 nm thick. We focused on

nanotubes that showed high-bias current near ~25 μA [150] and had diameters $D < 4$ nm as measured by atomic force microscopy (AFM), to ensure devices were single-walled. In addition, only electrical breakdowns with a single, clean drop to zero current were selected, which are typical of single-wall single-connection devices, as shown in Figure 6.1(b); by contrast, multi-wall CNTs and CNT bundles exhibit higher currents and break down with multiple current steps. [137] Joule heating was achieved by increasing the source-drain voltage ($V_{SD} > 0$) while maintaining a negative gate bias ($V_{GD} \approx$ -15 V). In semiconducting CNTs this leads to hole-only conduction, [147] deliberately avoiding ambipolar behavior [151] which would complicate the analysis. Metallic CNTs show no gate voltage dependence in room temperature, ambient conditions. Increasing V_{SD} leads to increasing the power input, which causes the CNT temperature to rise through Joule heating and leads to physical breakdown. We note that in this work the drain is always grounded and the source is the positive terminal, referring to the source of carriers and current flow.

The breakdown voltage, $V_{SD} = V_{BD}$ is taken to be the voltage at which the drain current (I_D) irreversibly drops to zero, as shown in Figure 6.1(b). We assume that during the breakdown process that the CNT stays on the surface and that no buckling or delamination from the surface occurs as a result of the small thermal expansion coefficient of CNTs. [152] Typical broken devices under AFM imaging are shown in Figures 1(c) and 2(a). The power dissipated within the CNT at breakdown is $P_{BD} = I_D(V_{BD}–I_DR_C)$. The combined resistance of the source and drain contacts, R_C, is estimated from the inverse slope of the low-bias I_D-V_{DS} plot, [141, 153] $R_C \approx (dI_D/dV_{DS})^{-1}$, which

includes the quantum contact resistance (R_0 = 6.5 kΩ). The experiments in this study were performed in air where nanotubes are known to break from self-heating and oxidation at a relatively well-known temperature, [154] $T_{BD} \approx 600°C$. By comparison, device breakdowns performed in ~10^{-5} Torr vacuum showed CNTs of similar lengths and diameters breaking at higher power and thus higher temperatures, as in Figure 6.1(b). This suggests that CNT device breakdowns in vacuum occur by a mechanism other than oxidation, e.g. at nanotube defects [155] or by failure of the underlying SiO_2. The latter is supported by the observation of damage to the SiO_2 substrate in some samples, as seen in Figure 6.1(c), which is never seen for breakdowns in air.

We now return to discuss the temperature profile of CNTs during Joule heating, and restrict ourselves to in-air breakdowns for the rest of the manuscript. Figure 6.2(a) displays the breakdown location (L_{BD}) along a CNT, as extracted from scanning electron microscope (SEM) imaging. Figure 6.2(b) shows a histogram of the normalized breakdown locations for ~40 CNTs in this study, distinguishing between m-CNT and s-CNTs. The majority of m-CNTs break at their hottest point near the middle while most s-CNTs break closer to the grounded drain, where the field is higher and the carrier density is lower. Both of these observations are indicative of diffusive heat [141] and charge [156] transport, and of relatively negligible contact resistance. At high field the electron or hole scattering mean free path (MFP) with optical phonons (OP) approaches the minimum value $\lambda_{OP,ems}$ ~15D where D is the diameter in nm. [147, 153] This MFP is significantly shorter than the CNT lengths used in this work (several microns).

To understand the temperature profiles of m-CNTs and s-CNTs, and to extract the interfacial thermal conductance per unit length (g) between CNT and SiO_2 from the breakdown data, we solve the heat diffusion equation along the CNT. [141] The heat generation per unit length can be captured both as uniform (for m-CNTs) and asymmetric (for s-CNTs), by expressing it as:

$$p(x) = p_0\left(C_1 + \frac{C_2 x}{L}\right), \tag{6.1}$$

where $-L/2 \le x \le L/2$, L is the length of the CNT, C_1 and C_2 are unitless parameters and p_0 is a constant term. We note that to a good approximation the heat generation in CNTs is independent of temperature, as the optical phonon emission length (the strongest inelastic scattering mechanism responsible for Joule heating) has very weak temperature dependence. [141, 147]

For m-CNTs the heat generation is uniform due to constant electric field and charge density (barring significant and asymmetric contact resistance [157-158]) we simply set $C_1 = 1$ and $C_2 = 0$. This implies $p_0 = P_{BD}/L$ at breakdown in m-CNTs. For s-CNTs, a linear heat generation profile captures the asymmetry caused by non-uniform electric field and charge density. [159] The general expression for the temperature along the CNT at breakdown is:

$$T(x) = T_0 + \frac{P_{BD}}{g_{tot} L}\left[C_1 + \frac{C_2 x}{L} - \frac{C_1 \cosh\left(\frac{x}{L_H}\right)}{\cosh\left(\frac{L}{2L_H}\right)} - \frac{C_2 \sinh\left(\frac{x}{L_H}\right)}{2\sinh\left(\frac{L}{2L_H}\right)}\right], \tag{6.2}$$

where $L_H = (kA/g)^{1/2}$ is the thermal healing length (of the order ~0.2 μm), [141, 160] k is the thermal conductivity of the CNT, [156] g_{tot} is the thermal conductance per unit length from CNT to ambient (see Section III below), and $A = \pi aD$ is the cross-sectional area assuming a CNT wall thickness a = 0.34 nm.

The typical "inverted U" shape of the temperature profile under uniform heat generation in m-CNTs is shown in Figure 6.2(c) with $C_2 = 0$. This has previously been observed experimentally in nanotubes under high bias operation, both by scanning thermal microscopy (SThM) [161] and by coating the CNTs with a phase-change material which changes volume as it heats up. [160]

On the other hand, s-CNTs have non-uniform electric field and charge density along their length, leading to off-center heat dissipation. [159] This is captured by changing the value of the parameter $C_2 > 0$ above, as shown in Figure 6.2(c). We take this simple approach because uncertainties in threshold voltage, contact resistance, and contributions made by infrequent defects make it difficult to provide a more exact solution of the temperature profile in every s-CNT measured. (by contrast, m-CNTs are immune to threshold voltage variations). More specifically, in our analysis below we choose $C_1 = 1$ and $C_2 = 0.65$ for s-CNTs, such that the hot spot location corresponds to $L_{BD}/L \sim 0.7$ as noted in the breakdown histogram, Figure 6.2(b).

6.3. Modeling

To understand the dependence of thermal coupling g on CNT diameter and substrate properties we use a diffuse mismatch model (DMM) [162] in a similar manner

previously applied to multiwall carbon nanotubes [163] and graphene. [164] The DMM is used to establish an upper bound for heat transport across an interface, as limited by the phonon density of states (PDOS). This approach also presents an advantage of speed and flexibility over full MD methods. [146] The model calculates the transmission probability, τ, for heat transfer across an interface while assuming all phonons scatter diffusely at the interface. By equating the phonon energy flux from the CNT to the SiO_2 with that from the SiO_2 to the CNT and using a detailed balance argument for all frequencies, [164-165] τ is given as:

$$\tau = \frac{\frac{1}{4}\int \omega N_{OX} f_{OX} G_{OX} v_{OX} d\omega}{\frac{\int \omega N_{CNT} f_{CNT} G_{CNT} \langle v_{CNT} \rangle d\omega}{2\pi a D} + \frac{1}{4}\int \omega N_{OX} f_{OX} G_{OX} v_{OX} d\omega}, \tag{6.3}$$

where N refers to the atomic density (in atoms/cm^3 for SiO_2 and atoms/cm for nanotubes), v is the phonon velocity, ω is the phonon frequency, f is the Bose-Einstein (BE) distribution, and G is the PDOS per atom as calculated by MD simulations. [146] We use the realistic PDOS rather than a Debye approximation because the latter has been previously found to cause large discrepancies with experimental data at high temperature. [166] In addition, the linear Debye approximation would not account for the quadratic CNT flexure modes. [167] The PDOS for a (10,10) nanotube with 1.37 nm diameter is calculated and shown in Figure 6.3. Using the PDOS from CNTs of other diameters did not change our results significantly (presumably because the proportion of

phonon modes remains approximately the same [146]), hence we used the PDOS shown in Figure 6.3 as the phonon weighing function throughout the remainder of this work.

The phonon velocity in the amorphous SiO_2 is assumed to be isotropic and fitted with a single value, [168] as shown in Table 6.1. However, the CNT phonon velocity includes contributions from both the transverse and longitudinal polarizations along the c-axis (out-of-plane direction) of graphite. [164, 169] Not included are the a-axis (in-plane) modes which contribute minimally to thermal coupling in the geometry of interest, and are more relevant to vertical CNTs on a surface. [144] We note, however, that even for vertical CNTs some degree of tip bending must always exist, thus the geometry examined here and in Ref. [146] is likely to be most relevant. Furthermore, to treat the nanotube as a hollow cylinder rather than a bulk material, we use the average value of the velocity components in the direction perpendicular to the plane of the substrate. In this context, the atoms at the top and bottom of the CNT (see Figure 6.1(a) or Figure 6.5(b) inset) have the maximum phonon velocity as they oscillate in the direction perpendicular to the substrate plane. The atoms on the left and right sides of the CNT oscillate parallel to the plane of the substrate. We take the average velocity of all CNT atoms in the direction perpendicular to the substrate, and the list of parameters is shown in Table 6.1.

Knowing the transmission probability, we can now calculate the flux of phonons through the interface. This gives the thermal conductance per unit length from the CNT to SiO_2 as:

$$g = \frac{N_{CNT} \cdot b_t}{2\pi aD} \int h\omega \frac{\partial f_{BE}}{\partial T} G_{CNT} \left\langle v_{CNT} \right\rangle \tau d\omega \,, \tag{6.4}$$

where b_t is the effective thermal contact width or footprint between the CNT and the substrate, to be determined by MD simulations (Figure 6.1 and Figure 6.5). This footprint is the effective width between CNT and the substrate over which heat is being transferred. Finally, to calculate a thermal boundary conductance that is comparable to experimental data, we must also include the effect of heat spreading into the oxide, given as: [170]

$$g_{ox} = \frac{2\pi\kappa_{ox}}{\ln\left(\frac{8t_{ox}}{\pi b_t}\right)}, \tag{6.5}$$

where $\kappa_{ox} \approx 1.4$ $Wm^{-1}K^{-1}$ is the SiO_2 thermal conductivity and $t_{ox} \approx 90$ nm is the SiO_2 layer thickness. This simple expression is appropriate when $t_{ox} \gg b_t$ as in our work, and the thermal spreading resistance contribution of the SiO_2 accounts for approximately 10-30% of the total thermal resistance. The total thermal conductance per unit length from CNT to ambient, as used in Eq. (6.2), is given by the simple thermal series network shown in Figure 6.1:

$$g_{tot} = \left(\frac{1}{g} + \frac{1}{g_{ox}}\right)^{-1}. \tag{6.6}$$

We note that any additional thermal spreading resistance into the Si wafer is negligible, and thus the Si wafer is assumed to be isothermal at $T_{Si} = 293$ K. Similarly, heat loss to ambient air can be neglected, where $g_{air} \sim 4\times10^{-4}$ $WK^{-1}m^{-1}$ has been

previously estimated as an upper limit at one atmosphere, [171] three orders of magnitude lower than the heat loss to substrate.

6.4. Derivation of CNT shape and footprint

6.4.1. Equilibrium shape of a CNT

Nanotubes interact with the SiO_2 substrate through van der Waals (vdW) forces. In addition, our previous MD simulations [146] have shown that such CNTs do not remain rigid cylinders, but instead deform to minimize their overall vdW and curvature energy. Beyond a certain diameter CNTs relax to a compressed shape, [127, 172] which changes both their geometrical and equivalent thermal footprint on the substrate. To accurately calculate the shape and thermal footprint of the CNT we employ MD simulations with a simplified Lennard-Jones (LJ) 6-12 potential:

$$V = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right]. \tag{6.7}$$

Here, we simplify the SiO_2 substrate as a continuum plane. Therefore the collective vdW interaction per carbon atom situated at a height h above an infinite half-space of SiO_2 can be approximated by the triple integral

$$V_{vdW} = \sum_{i=Si,O} \int_0^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} 4\varepsilon_i n_i \left[\left(\frac{\sigma_i}{r}\right)^{12} - \left(\frac{\sigma_i}{r}\right)^{6}\right] dx dy dz, \tag{6.8}$$

in which

$$r = \sqrt{x^2 + y^2 + (z+h)^2}\ , \tag{6.9}$$

and

$$\begin{aligned} \varepsilon_{Si} &= 8.909\,\text{meV}, \sigma_{Si} = 3.326\,\text{Å}, n_{Si} = 0.0227\,\text{Å}^{-3} \\ \varepsilon_{O} &= 3.442\,\text{meV}, \sigma_{O} = 3.001\,\text{Å}, n_{O} = 0.0454\,\text{Å}^{-3} \end{aligned}. \tag{6.10}$$

The values here are based on the Universal Force Field (UFF) model by Rappe *et al.* [173] and were used in our previous MD simulations as well. [146] The integral in Eq. (6.7) can be evaluated analytically. It should be noted that this integral tends to give a lower bound estimate of the total interaction potential because it ignores the effects of local spikes of closely positioned atoms. The estimation error is reduced by assuming a relaxed configuration for the nearby silica molecules. Such an analysis gives

$$V_{vdW} = \sum_{i=Si,O} \frac{2\pi\varepsilon_i\sigma_i^3 n_i}{45}[2(\sigma_i/h)^9 - 15(\sigma_i/h)^3], \tag{6.11}$$

which has a similar form as the original LJ potential – except with different exponents and pre-factors. This effectively alludes to an h^{-3} dependence of the vdW interaction potential. A plot of both the calculated potential and its second derivative (which is proportional to the interaction spring constant) is shown in Figure 6.4.

For the covalent C-C interaction we used the empirical bond order Tersoff-Brenner potential. [10] In addition to this potential, we used an intra-molecular LJ vdW potential with the following parameters for graphite: [129]

$$\varepsilon_C = 3.02\ \text{meV}\ , \sigma_C = 3.39\ \text{Å}\ . \tag{6.12}$$

All MD simulations were carried out until the transient motions died off and a final steady-state solution was reached.

6.4.2. Thermal footprint of a CNT

To determine the thermal footprint of the CNT on SiO_2 we consider the square root of the second derivative of the vdW potential with respect to h as heat transfer depends on this effective "spring constant" between the substrate and CNT. For example, the phonon velocities are expected to be proportional to the square root of this spring constant. Thus, to find the effective thermal footprint we used the square root of the spring constant to weigh the horizontal change in position, Δx.

The thermal footprint (b_t) should not to be confused with the geometric footprint, b_g, the physical contact region between the CNT and substrate. In the case of small diameter CNTs, the effective thermal footprint can even be greater than the lateral width of the CNTs, i.e. their diameter. This occurs because in addition to the bottom half of the CNT conducting heat to the substrate, there is also thermal coupling from the top half of the CNT. The results of these simulations are shown in Figure 6.5. Because MD simulations can be carried out for only one CNT of a particular diameter at a time, several were conducted for CNTs over a range of diameters 5–49 Å. We found the following quadratic function fit the simulation results of the thermal footprint for any diameter within the simulated range (Figure 6.5(a)):

$$b_t = 0.0037\,D^2 + 1.1D\,, \tag{6.13}$$

where both b_t and D are in units of Ångstroms.

Our simulations further suggest that there are two different regimes represented by different equilibrium shapes of CNTs, as shown in Figure 6.5. In the first regime (labeled "I"), the diameter of the CNT is $D < 2.1$ nm and the curvature energy of the CNT is stronger than the vdW energy with the substrate. Thus in the first regime the cross-section of the nanotube more closely resembles a perfect circle, as shown in the left inset of Figure 5(b) for a (7,7) CNT. In addition, the geometrical footprint (calculated by finding the furthest distance between the lowest points on the CNT) in this regime remains nearly constant at ~1.4 Å, the chemical bond length, as can be seen in Figure 6.5(a).

In the second regime (labeled "II") the diameter $D > 2.1$ nm, and the vdW energy with the substrate is stronger than the curvature energy of the CNT. Hence the final minimum energy shape for the CNT will be that of a deformed circle, as shown for a (22,22) CNT in the right inset of Figure 6.5(b). It is in this regime where the geometrical footprint starts to increase approximately linearly with diameter, as shown in Figure 6.5(a). Another interesting observation is noted due to the repulsive nature of the vdW forces at very close distances, whose relative magnitudes are illustrated with arrows in Figure 6.5(b). In this case, the bottom of the CNT is not perfectly flat. Instead the middle of the bottom region buckles up slightly, such that the force at the center is nearly zero. All these effects are captured in the thermal footprint calculation (b_t) fitted by Eq. (6.12) above, and used in the DMM thermal coupling simulations.

6.5. Discussion

Figure 6.6(a) shows the directly measured power at breakdown (P_{BD}), and Figure 6.6(b) displays the extracted TBC (g) vs. diameter D for 29 metallic and semiconducting CNT devices. Fig 6.6(b) also includes modeling using the DMM described above (solid line) and the dashed lines fitted to MD simulations with vdW coupling strengths $\chi = 1$ and $\chi = 2$, as described in Ref. [146]. Both data and modeling trends in Figure 6.6(b) suggest that the TBC increases with diameter. The range of extracted g corresponds to approximately the same order of magnitude previously extracted from thermal breakdowns. [141, 174] A representative set of vertical error bars on one of the m-CNT's corresponds to a ±50 °C uncertainty in breakdown temperature. Horizontal error bars represent ±0.4 nm uncertainty in diameter from AFM measurements. Vertical error bars on the s-CNTs are derived as follows. The upper limit is set by assuming $L_{BD}/L = 0.75$ and the lower limit is set in the limit of uniform heat generation. It is interesting to note that that non-uniform heat generation plays a larger role in large diameter s-CNTs than in small diameter s-CNTs.

6.5.1. Dependence of thermal coupling on diameter

We observe that g increases with diameter up to ~0.6 $WK^{-1}m^{-1}$ per unit length for the largest single-wall CNTs considered (D ~ 4 nm). The diameter dependence of g is primarily a result of the increase in thermal footprint, as shown in Fig 6.5(a). Also plotted in Figure 6.6(b) are our previous MD simulations results. [146] The results from the MD simulations do give lower values of g because the DMM assumes, by definition,

that all phonons are scattered diffusely at the interface [162] whereas this does not necessarily happen in MD simulations.

We also obtain the thermal contact conductance per unit area, $h = g/b_t$, as plotted in Figure 6.6(c) and showing almost no dependence on diameter. From the breakdown experiments this value is in the range $h \approx 20$–200 $MWK^{-1}m^{-2}$ which is slightly larger than that recently obtained for graphene on SiO_2. [175] The DMM simulation predicts an upper limit for $h \approx 220$ $MWK^{-1}m^{-2}$ with almost no diameter dependence. This appears to suggest the upper range of the h values obtained experimentally. We note that the extracted and simulated TBCs in this study thus far are at an elevated temperature, given approximately by the CNT breakdown condition (~873 K). To understand the effects of temperature on TBC, we plot our DMM model in Figure 6.6(d) vs. temperature. This shows an expected increase in TBC with temperature, consistent both with graphene-SiO_2 experiments [175] and with CNT-SiO_2 MD simulations. [146] The thermal coupling per unit area at room temperature is ~130 $MWK^{-1}m^{-2}$, or approximately 40 percent lower than the thermal coupling near the CNT breakdown temperature.

6.5.2. Dependence of TBC on phonon DOS and velocity

In addition to the thermal footprint, the PDOS of the SiO_2 as well as the distribution function (f_{BE}) also play a role in heat transport across the interface. We recall that the inset of Figure 6.3 showed the calculated PDOS for both a (10,10) CNT and the SiO_2 substrate. While the nanotube contains a large PDOS peak at 53 THz, this does not come into play because there are no equivalent high-frequency modes in the SiO_2.

Figure 6.3 also shows the Bose-Einstein distribution function at the CNT breakdown temperature (T_{BD} ~ 600 °C). The distribution suggests very low occupation for all high frequency CNT modes. Since the Debye temperature for CNTs is very high, we expect that most substrates will serve as a low-pass filter for CNT phonons.

Aside from changing the thermal footprint, the deformed shape of the CNT also affects the average phonon velocity. This is a more subtle effect than that of diameter or surface roughness, but it is included here for completeness. For instance, in the second regime (D > 2.1 nm) the CNT becomes flattened, leading to more atoms vibrating perpendicular to the substrate interface. After numerical MD calculations of the CNT shape (Figure 6.5(a)) we fit analytic expressions to the angle-averaged phonon velocity:

$$\begin{aligned} \langle v_{CNT} \rangle &= 2.926D + 574 \qquad \text{for } D < 2.1 \text{ nm} \\ &= -2541D^{-0.71} + 932 \qquad \text{for } D > 2.1 \text{ nm} \end{aligned}, \tag{6.14}$$

and these are used in the calculation of the thermal boundary conductance in Eq. (6.4). The units of the equation above are given for D in Angstroms, yielding average velocity in m/s.

6.5.3. Dependence of TBC on surface roughness

There are several variables contributing to the spread of the experimental data shown in Figures 6.6 and 6.7. The primary contributor is surface roughness. Since the value of g is directly related to the contact area at the interface, an imperfect surface is roughly equivalent to a decreased thermal contact area. Figure 6.7(a) replots the calculated TBC vs. diameter for a perfectly smooth surface (100%), for 75% of the

maximum contact area, and for 50% of the maximum contact area. To analyze how surface roughness affects the spread directly, we experimentally find the average surface step height, Δ adjacent to the nanotube via AFM. However, intuitively we except the *ratio* of diameter to roughness (D/Δ) to be more important. Thus, we expect larger diameter CNTs to be less affected by surface roughness than smaller diameter CNTs. Plotting g versus D/Δ in Figure 6.7(b), we see that the spread is smaller in these plots than in Figure 6.6(a) and 6.7(a), confirming our hypothesis.

6.5.4. Role of s-CNT vs. m-CNTs

We note that the spread in m-CNTs breakdown data is smaller than in s-CNTs in Figures 6.6 and 6.7. We believe this is due to threshold voltage (V_{TH}) shifting in s-CNTs during the high-field measurement process, which m-CNTs are essentially immune to. As the devices are swept to high drain bias for breakdown, along with the applied gate bias (-15 V) this can lead to dynamic charge injection into the oxide, as studied in depth by Ref. [148]. To understand the effect of threshold voltage on breakdowns, we plot the extracted P_{BD} vs. initial V_{TH} in Figure 6.7(c), and find a slight but positive relationship. This suggests that in s-CNTs the variation in electronic behavior leads to the larger data spread, in addition to the variation due to surface roughness. Moreover, this also indicates a root cause which renders precisely selective breakdown of m-CNTs (e.g. in CNT networks) as a challenging and imperfect approach: on one hand, the change in threshold voltage of s-CNTs can turn them “on” allowing them to break down, on the

other hand the variation in surface roughness itself cannot guarantee that all m-CNTs will break down at the same input power, or voltage.

6.5.5. Comments on the modeling approach

It is important to note that both the DMM and MD simulations employed in this work only capture the lattice vibration (phonon) contribution to thermal coupling. Nevertheless, the DMM in general appears to represent an upper limit to the spread of the experimental data which is otherwise lowered by effects like surface roughness. However, recent theoretical work has also suggested a possible electronic contribution to heat transport through coupling with surface phonon polaritons (SPPs) from the oxide. [176-177] The SPP interaction drops off exponentially with the CNT-substrate distance, perhaps leading to a larger electronic contribution to heat transport in regime II of the CNT shape (D > 2.1 nm), where more CNT atoms are closer to the SiO_2 surface. However, since the SPP potential is strongly dependent on the interaction distance, it will also be affected by substrate surface roughness. Given these circumstances it is difficult to rule out energy relaxation through SPP scattering in practice, although this appears to be significantly lower than the phonon coupling and any SPP contribution (however small) may become more significant in larger diameter CNTs (D > 2.1 nm).

Another mechanism for CNT-SiO_2 energy dissipation is inelastic phonon scattering at the interface, which is not captured by the DMM. Previously Chen *et al.* [175] had compared an elastic DMM calculated by Duda *et al.* [164] to the TBC between graphene and SiO_2 and found that the elastic DMM under-predicted the TBC by

approximately an order magnitude. Hopkins [165] made a similar argument for inelastic scattering between acoustically mismatched materials. However our simulations do not differ from the data significantly, thus our calculations suggest that the contribution of inelastic scattering here is small (perhaps a result of the 1-D nature of CNTs).

6.6. Conclusions

In summary, we have examined electrical breakdown and thermal dissipation between CNT devices and their SiO_2 substrate, the most common configuration found in CNT electronics. The breakdown location is invariably found in the middle of the CNT, consistent with the CNT temperature profile. In this context, thermal dissipation from CNT to SiO_2 dominates over dissipation at the CNT contacts. We found evidence of a direct relationship between the CNT-SiO_2 thermal boundary conductance (TBC) and the CNT diameter, in accord with previous MD simulations. To provide a more flexible means of analysis we developed a diffuse mismatch model (DMM) of the TBC using the full phonon density of states (PDOS). This approach appears to predict the upper limit of thermal transmission at the CNT-SiO_2 interface, and could be similarly applied to calculate the TBC of other dimensionally mismatched systems. Our experiments and modeling suggests a maximum TBC of ~0.6 $WK^{-1}m^{-1}$ per unit length for the largest diameter CNTs considered (D = 3–4 nm). The maximum thermal conductance per unit area corresponds to approximately 130 $MWK^{-1}m^{-2}$ at room temperature and 220 $MWK^{-1}m^{-2}$ at 600 ^{o}C.

We have also studied the thermal footprint of a CNT through MD simulations which find the atomic configuration of lowest energy. These reveal two interaction regimes, the first one at smaller diameters (D < 2.1 nm) where the CNT shape is dominated by its curvature energy, the other at larger diameters (D > 2.1 nm) where the CNT shape is dominated by van der Waals (vdW) coupling with the substrate. Finally, we found that SiO_2 surface roughness strongly affects the TBC of such nanometer-sized interfaces. To improve CNT heat sinking applications, our results suggest the need to engineer ultra-flat surfaces, use large diameter CNTs, and find substrates with larger vdW coupling. To improve selective electrical breakdown of CNTs (e.g. metallic vs. semiconducting) it will also be essential to control the surface roughness of the substrate, as well as the threshold voltage of the semiconducting CNTs.

6.7. Figures and tables

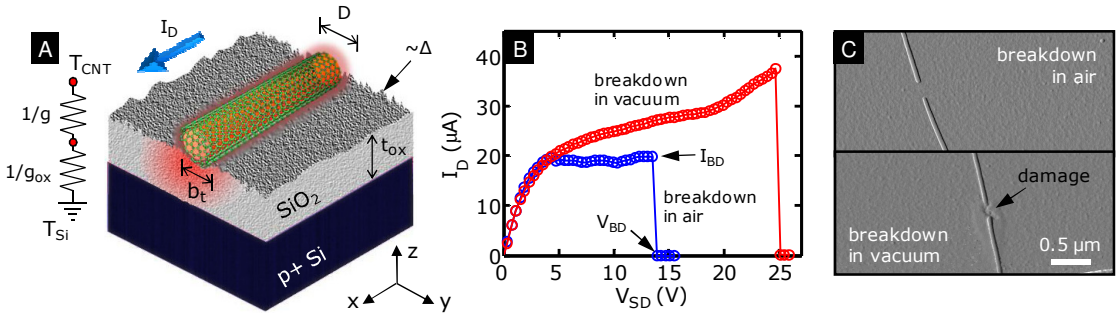


Figure 6.1. (a) Schematic cross-section of typical CNT device with diameter D and thermal footprint b_t (also see Figure 5) on SiO_2 substrate with thickness t_{ox} and surface roughness Δ. The p+ silicon is used as a back-gate. The device layout with source and drain terminals is shown in Figure 2(a). As current (I_D) passes in the CNT, the generated Joule heat dissipates through the substrate. The equivalent thermal circuit includes CNT-SiO_2 interface thermal resistance ($1/g$) and spreading resistance in the SiO_2 ($1/g_{ox}$). (b) Typical electrical breakdown of similar CNTs shows higher breakdown power in vacuum (~10^{-5} torr) than in ambient air. This illustrates the role of oxygen for CNT breakdown in air. (c) Atomic force microscopy (AFM) images of CNTs broken in air (top) and vacuum (bottom). Breakdowns in vacuum can lead to SiO_2 surface damage, which is not observed for air breakdowns.

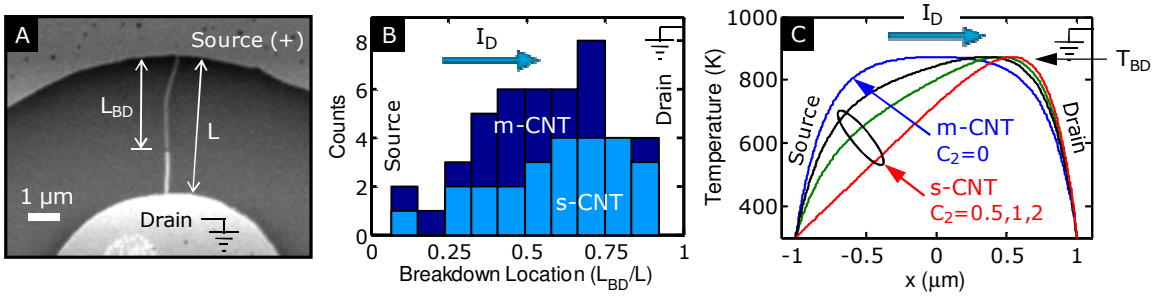


Figure 6.2. (a) Scanning electron microscope (SEM) image of CNT device showing breakdown location (L_{BD}). (b) Histogram of breakdown location normalized by CNT length (L_{BD}/L) indicating the majority of m-CNTs break near the middle and s-CNTs break closer to the drain. Break point is always observed along the CNT, suggesting good contacts with negligible heating. (c) Computed temperature distribution along a 2 µm long CNT (typical in our study) with Eq. (6.1) using C_1 = 1 and varying C_2. The maximum temperature is shown at the breakdown condition (T_{BD}). C_2 = 0 corresponds to m-CNTs (uniform heat dissipation) and C_2 > 0 corresponds to s-CNTs. For s-CNTs biased under hole conduction the heat generation and temperature profile are skewed towards the ground (drain) terminal. Block arrows in (b) and (c) show direction of hole flow.

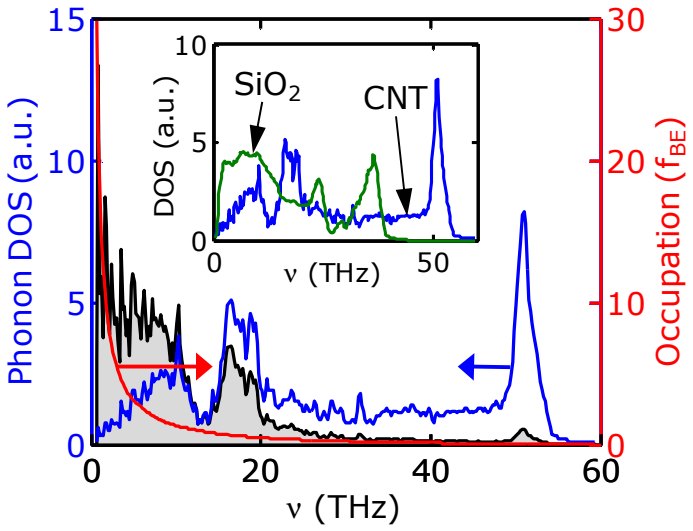


Figure 6.3. The phonon density of states (PDOS) for a (10,10) nanotube from MD simulations. The Bose-Einstein occupation (f_{BE}) at room temperature is plotted in red against the right axis. Shaded in gray is the product of the PDOS with f_{BE}, showing diminished contribution from higher frequency phonon modes. The inset shows the PDOS of the CNT and that of SiO_2, the latter displaying a lower cutoff near 40 THz.

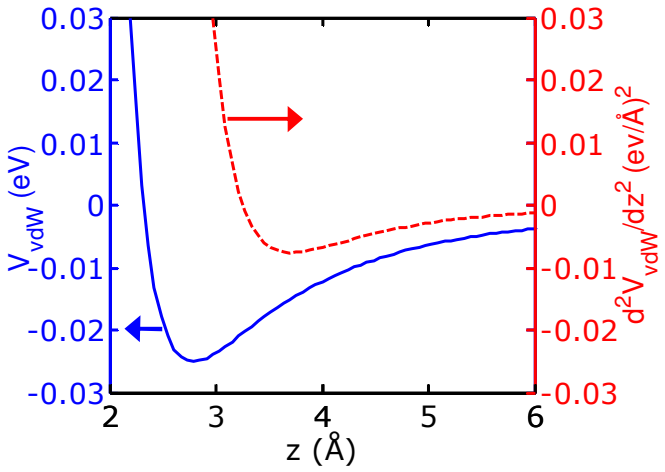


Figure 6.4. Van der Waals potential (blue solid line) interaction between CNT and SiO_2, as used in calculations to derive the thermal footprint (Figure 6.5). The second derivative of the potential (red dashed line) with respect to distance from the surface (z) is used to weigh the contribution of each atom to the effective thermal footprint (b_t) of the CNT.

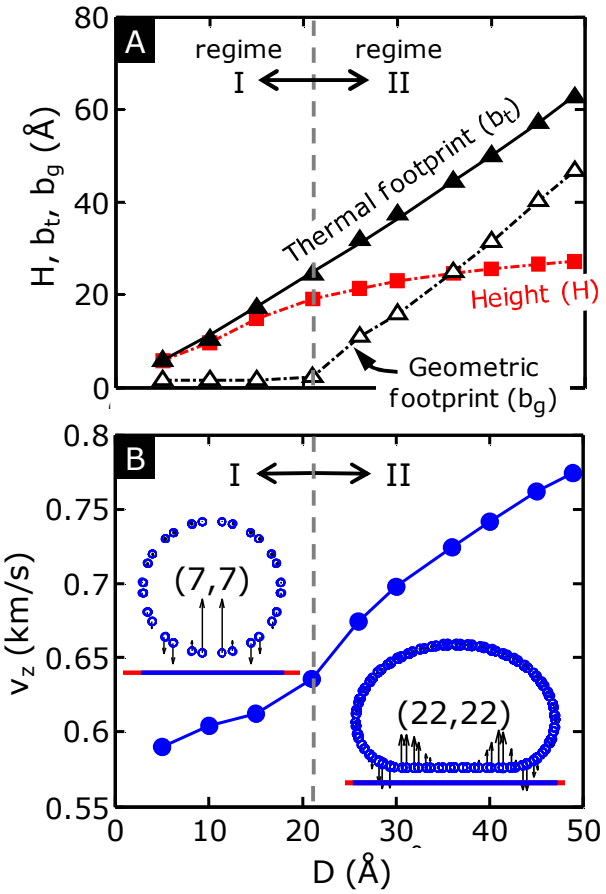


Figure 6.5. (a) Nanotube height (■), geometrical footprint (Δ), and thermal footprint (▲) on the SiO_2 substrate as a function of CNT diameter, obtained from MD simulations. A fit to the thermal footprint is shown as a solid line from Eq. (6.12). (b) Average CNT phonon velocity in the direction perpendicular to the surface. Calculations reveal two distinct regimes: in regime I (left inset, $D < 2.1$ nm) the CNT shape is nearly circular, dominated by the curvature energy; in regime II (right inset, $D >$ 2.1 nm) the CNT shape becomes flattened, with a stronger influence of the surface vdW interaction. Small vertical arrows indicate the relative magnitude of vdW forces with the substrate at each atomic position.

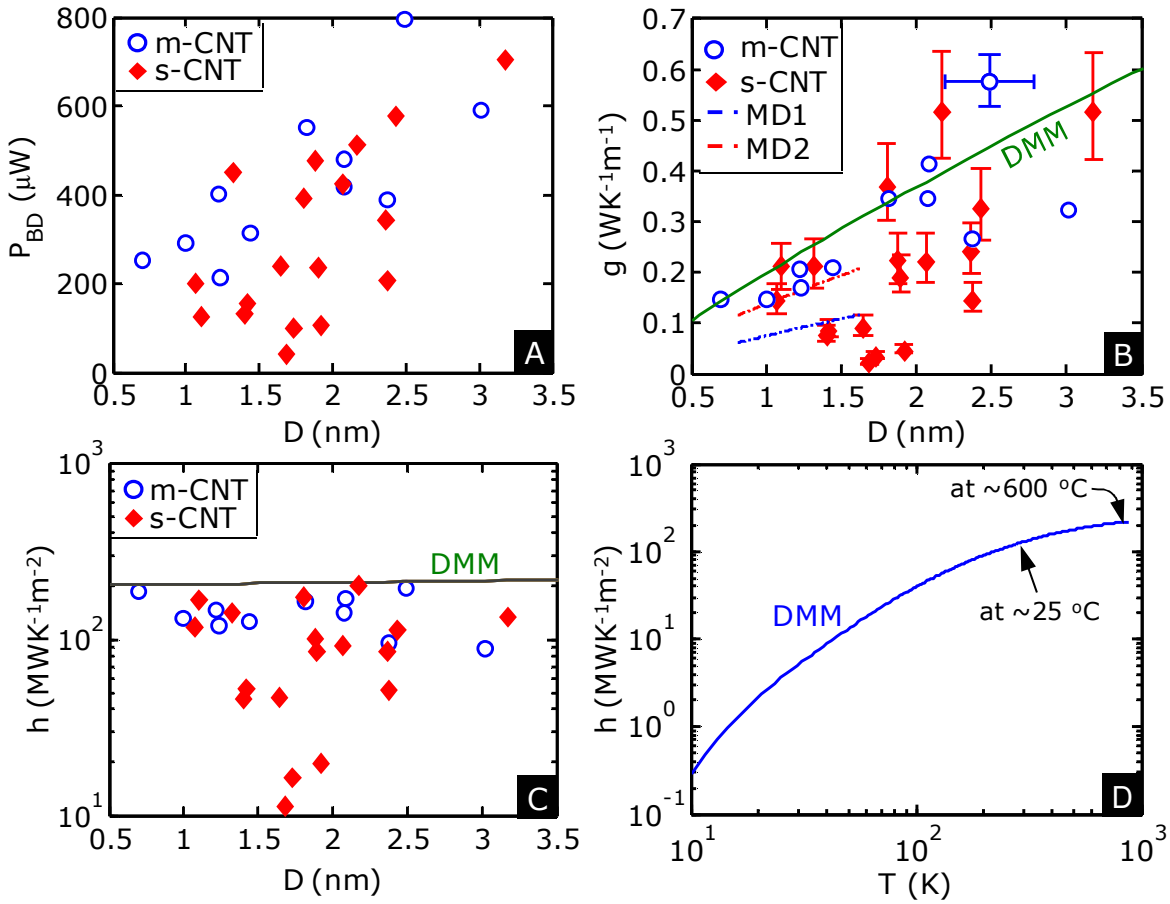


Figure 6.6. (a) Electrical breakdown power (in air) of CNTs vs. diameter *D*, showing proportional scaling. (b) Extracted CNT-SiO_2 thermal coupling *g* vs. *D* (see text) for both metallic (m) and semiconducting (s) CNTs. Solid line is the DMM calculation and dash-dotted lines are fitted to MD simulations with different vdW coupling strengths (χ=1 and χ=2 respectively). (c) CNT-SiO_2 thermal coupling per unit area *h* vs. *D*, showing the DMM represents an upper-limit scenario of heat dissipation. The spread in the data and lower apparent thermal coupling in practice is attributed to SiO_2 surface

(cont. on next page)

roughness, and charge trapping near semiconducting CNTs (see text). (d) Calculated temperature dependence of the upper limit thermal coupling per unit area. Thermal coupling at room temperature (~130 $MWK^{-1}m^{-2}$) is ~40% lower than at the breakdown temperature (~220 $MWK^{-1}m^{-2}$).

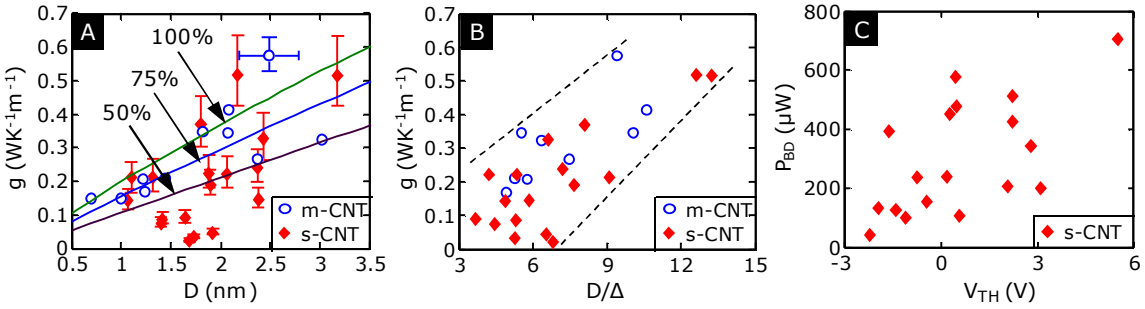


Figure 6.7. (a) CNT-SiO_2 thermal coupling g vs. diameter D (symbols = data) and DMM simulations (lines) for perfect substrate contact (100%), and for 75% and 50% effective contact area due to SiO_2 surface roughness (also see Figure 6.1). (b) Replot of same experimental data vs. diameter scaled by RMS surface roughness (D/Δ) measured by AFM near each CNT. This indicates the role of SiO_2 surface roughness for thermal dissipation from CNTs. Dashed lines are added to guide the eye. (c) Breakdown power P_{BD} for semiconducting CNTs (s-CNTs) alone plotted with respect to threshold voltage (V_{TH}). The variance in V_{TH} is also a contributing factor to the spread in extracted thermal coupling data for s-CNTs.

Parameter	Value
v_{CNT}	932 m/s
v_{ox}	4.1 km/s
N_{CNT}	16.3 atoms/Å
N_{ox}	0.0227 molecules/Å^3
T_{BD}	873 K

Table 6.1. Parameters used in the DMM model.

CHAPTER 7: INTERPRETATION OF TOPOGRAPHICAL INFORMATION FROM ATOMIC FORCE MICROSCOPE (AFM) IMAGES ON INDIVIDUAL CARBON NANOTUBES

7.1. Introduction

Carbon nanotubes (CNTs) have generated much interest over the past two decades due to their extraordinary mechanical and electrical properties. Numerous efforts have used CNTs as building blocks of nanoelectronic devices such as transistors, switches and wires. Unlike semiconductor nanowires, CNTs do not have sharp edges or dangling bonds, thus presenting much higher high carrier mobility and thermal conductivity. [178-180] Carbon nanotubes could also be used as building blocks in nanosensors, nanocomposites and other nano-electro-mechanical (NEMS) devices and nanostructures. [131-132, 181] Among these applications, the CNT diameter strongly affects many CNT properties such as band gap, [182] carrier mobility, [153] contact resistance, [183] and stiffness [121]. The most common method of measuring the diameter of a CNT on a substrate is by atomic force microscopy (AFM), and the most common substrate used in practice is amorphous silicon dioxide, thermally grown on Si wafers. [80, 184-188]

The AFM measurement determines the CNT diameter by sampling the "height" of the CNT on the substrate; however when placed on silica or other solid substrates, the CNT cross-section can deform due to van der Waals (vdW) interactions between the carbon atoms and the substrate. [127, 130, 180, 184, 189] Thus, a central question in such measurements is how to interpret the AFM topography [190-193] of individual CNTs, while taking into account the deformation of CNTs induced by vdW interactions both with the substrate and with the AFM tip. This paper focuses on answering this important issue through detailed molecular dynamics simulations and comparison with experimental measurements. The results obtained herein should be helpful in interpreting AFM measurements of CNTs and in designing and properly evaluating future CNT-based devices.

7.2. Simulation method

To study the deformation of CNTs due to various interactions we first conducted molecular dynamics (MD) simulations. Several different types of interactions were considered: the covalent bonding between the carbon atoms in CNTs, the vdW interactions between the CNT and the substrate, and the vdW interactions between the CNT and the AFM tip.

The neighboring carbon atoms within CNTs form covalent sp^2 C-C bonds, modeled with the empirical bond order Tersoff-Brenner potential [10] which has been well characterized and extensively used by previous researchers [81, 116, 172, 194]. The

C-C interactions occur only when the atoms are within the cut-off radius of each other, taken as r_c ~ 2.5 Å in our simulations.

When CNTs are placed on a substrate, they deform due to the vdW interactions between the carbon atoms and the substrate material. As a result, their cross-sectional profile no longer remains circular, as found in several previous studies, [127, 130, 180, 184, 189]. The CNT-substrate vdW potential used in our calculations is derived elsewhere [130] but repeated here for completeness:

$$V_{vdW} = \sum_{i=Si,O} \frac{2\pi\varepsilon_i\sigma_i^3 n_i}{45}\left[2\left(\frac{\sigma_i}{h}\right)^9 - 15\left(\frac{\sigma_i}{h}\right)^3\right], \tag{7.1}$$

where [173]

$$\begin{aligned} &\varepsilon_{Si} = 8.909 \text{ meV}, \sigma_{Si} = 3.326 \text{ Å}, n_{Si} = 0.0227 \text{ Å}^{-3} \\ &\varepsilon_O = 3.442 \text{ meV}, \sigma_O = 3.001 \text{ Å}, n_O = 0.0454 \text{ Å}^{-3} \end{aligned} \tag{7.2}$$

Furthermore, the non-neighboring carbon atoms in the CNTs that are not directly connected by covalent bond may still interact with each other by vdW forces when the interatomic distance is sufficiently close. We model such interactions via a standard Lennard-Jones (LJ) potential with the following parameters: [128]

$$\varepsilon_C = 2.39 \text{ meV}, \ \sigma_C = 3.415 \text{ Å}, \ n_C = 0.176 \text{ Å}^{-3}. \tag{7.3}$$

The equilibrium cross-section of several single-walled CNTs is shown in Figure 1(a), where each ring represents the cross-section of a particular diameter. The inner (blue) rings represent the equilibrium shapes of CNTs with diameters D < 2.2 nm. In this regime, we find that the elastic energy due to relatively large intrinsic curvature of the

CNTs is dominant. The outer (red) rings represent the equilibrium shapes of CNTs with diameters $D > 2.2$ nm. In this regime, the CNTs deform more and their profile shows a non-circular cross-section. For such large diameters the maximum curvature (which occurs at the left and right ends of the nanotubes) is nearly constant, and the CNT shape is governed by the vdW interactions with the substrate.

The two distinct regions are also noted in Figure 7.1(b), which reveals a bilinear relationship between the maximum curvature in a deformed CNT and inverse of its diameter ($1/D$). For $D > 2.2$ nm ($1/D < 0.45$ nm^{-1}) the maximum curvature of a CNT on SiO_2 changes very slowly with the inverse diameter as shown by the red line. However, for CNTs with $D < 2.2$ nm the maximum curvature begins to vary directly proportional with the inverse diameter (blue line), indicating that CNTs on SiO_2 maintain a nearly circular cross-section in this regime. This bilinear dependence in turn gives rise to bilinear relationship between the height (and to a lesser degree width) vs. diameter of the CNTs as shown in Figure 7.1(c). The width is defined here as the longest possible inscribed horizontal line within a CNT. The width changes almost linearly with the CNT diameter, whereas the height is distinctively bilinear with respect to D. Similar behaviors are predicted by our simulation results for double-walled CNTs on SiO_2 substrates as shown in Figure 7.1(d), with the "deflection point" occurring at a larger CNT diameter.

In order to compare the calculated CNT profiles with results measured by AFM, we then calculate the CNT cross-sections as an AFM tip of given radius is moved across the CNT. Since the radius of curvature of AFM tips (typical range $R_T \approx$ 2–20 nm) is usually comparable to or larger than the CNT diameter ($D \approx$ 0.7–4 nm for single-wall

CNTs), the trace profile of the AFM tip will not conform to the shape of CNTs. Instead, the trace profile of the AFM tip would in general be of roughly the same height (H) as that of the CNT but much wider (W_{AFM}). If one does not take into account the CNT-AFM tip interaction and its resulting deformation of the CNT (i.e., for perfectly rigid CNTs), such AFM tip trace profiles can be obtained analytically based on geometrical considerations, [195] e.g. $W_{AFM} \approx 2(2R_T h)^{1/2}$.

In experiments the AFM tip can interact with the CNT as it moves across it through vdW forces. Such interactions, in turn, cause deformation in the CNTs being imaged, in both contact and tapping modes [196]. To determine the relationship between the AFM tip trace profile and the true dimensions of a CNT, we conduct MD simulations by taking into account the full interactions of the AFM tip and the CNT. We model the tip-CNT interaction with a simplified vdW potential, i.e. the standard Lennard-Jones (LJ) 6-12 potential as:

$$V = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right], \tag{7.4}$$

where ε is an energy parameter, σ is a length parameter (given in Eqs. (7.2) and (7.3)) and r is the interatomic distance. We consider parabolic-shaped diamond AFM tips, but similar conclusions can be reached for AFM tips of different shapes and materials. If the parabolic AFM tip is defined by [195] $f(x,y) = (x^2 + y^2)/2R_T$, the collective vdW interaction per C atom can be approximated as:

$$V_{vdW} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{f(x,y)}^{\infty} 4\varepsilon n\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] dzdxdy\,, \tag{7.5}$$

where

$$r = \sqrt{(x-k)^2 + y^2 + (z+h)^2}\,, \tag{7.6}$$

and (k, h) are the horizontal distance and depth of C atoms under the AFM tip, respectively. (h and k are thus zero at the apex of the AFM tip.) We note that this integral tends to give lower estimates of the total interaction potential due to ignoring the effects of local spikes of closely positioned atoms. This integral cannot be evaluated analytically and should be treated numerically. However, if the lower bound of the z integral is changed to zero, i.e. a flat AFM tip, an analytical formula exists: [130]

$$V_{vdW} = \frac{2\pi\varepsilon\sigma^3 n}{45}\left[2\left(\frac{\sigma}{h}\right)^{9} - 15\left(\frac{\sigma}{h}\right)^{3}\right]. \tag{7.7}$$

This result indicates that in general the potential is highly localized and non-zero only in the close vicinity of the surface. The integral in Eq. (7.5) may be evaluated numerically via partitioning the integration domain into volumetric elements and carrying out a standard Gauss integration as shown in the inset of Figure 7.2. This integral is non-zero only in a narrow band around the tip surface. This observation allows us to use the approximate localized potential between AFM tips and CNTs for the simulations in this work, with the form:

$$V_{vdW} \approx 4\varepsilon n\left[A\left(\frac{\sigma}{h}\right)^{\alpha} - B\left(\frac{\sigma}{h}\right)^{\beta}\right], \tag{7.8}$$

where A = 1.645 Å^3, B = 14 Å^3, $\alpha = 9$ and $\beta = 3.288$ are fitting parameters chosen to match the numerical data, as shown by the fit in Figure 7.2 for diamond AFM tips. In this model, h is the closest perpendicular distance to the surface of the AFM tip from any point in space and n is the atomic density of diamond. If instead of diamond we consider silicon AFM tips, the parameters are: [197]

$$\varepsilon_{Si} = 6.46\ \text{meV},\ \sigma_{Si} = 3.621\ \text{Å},\ n_{Si} = 0.05\ \text{Å}^{-3}, \tag{7.9}$$

with the parameters of the approximation becoming A = 1.943 Å^3, B = 16.49 Å^3, $\alpha = 9$ and β = 3.174. These *effective potentials* for silicon and diamond tips are employed throughout this work, dramatically increasing the efficiency of the simulations.

Just as in experiments, to ensure the stability of AFM scanning, we need to apply a small downward compressive force on the tip at all times. For AFM scanning of a solid surface the exact magnitude of this force is less important. However, for deformable surfaces such as CNTs, the magnitude of the compressive force will affect the degree of deformation, thus affecting the scanning results. In the current simulations, we apply a compressive force ranging from 0.1 nN (extremely small) to 10 nN. It is noted that, without the compressive force, we observe considerable jittering in our numerical results. Another parameter is the AFM tip scanning speed, which is 5 m/s in the numerical simulations. Experimental scan speeds are on the order of μm/s, too time consuming for the numerical simulation. However, we have verified that no significant

spurious effects exist in our simulations with the current scanning speed. We also performed a few sensitivity analyses where we repeated some of our simulations with lower scanning speeds and we didn't observe any meaningful systematic change in our results. Since our computational AFM tip has no mass and therefore it has an infinite gain, its measurements are not very sensitive to the speed of scan as it can instantaneously adjust itself with the various surface features. In this regard, our computational AFM tip behaves more similar to a contact mode actual AFM tip.

7.3. Simulation results

Figure 7.3 shows typical 3-dimensional (3D) simulations of single-walled CNT shapes at different stages as the AFM tip moves across their surface. Figure 7.3(a) shows a CNT on SiO_2 when the AFM tip is sufficiently far away such that any interactions are negligible. Figures 7.3(b)-(d) depict the CNT shape as the AFM tip moves across it (also see figures in Appendix B). These pictures clearly show that, because of the vdW interactions between the AFM tip and CNT as well as deformability of the CNT, an indentation forms in the CNT near the location of the AFM tip. The degree of the indentation depends on the magnitude of the compressive force on the AFM tip and the diameter of the CNT. However, it is noted that, even with zero compressive force on the AFM tip, an indentation still forms if the CNT is sufficiently flexible (large enough diameter). This indentation occurs because the attractive vdW interaction between carbon atoms and the AFM tip can draw the tip into the CNT until the elastic resistance of the tube balances the attractive force.

Figure 7.4(a) shows the trace profiles of AFM tips of several different radii over a CNT of diameter D = 2.7 nm when the tip-CNT interaction is not accounted for (i.e., perfectly rigid CNT). Here we assume that the AFM tip is in perfect contact with the CNT as well as with the substrate, i.e., the distance between them is zero. It is clear that the "measured" height by the AFM tip trace profile can accurately represent the true height of the CNT. However the "measured" width by the AFM tip trace profile W_{AFM} is much larger than the true width of the CNT W. Although highly simplified, this result nevertheless serves to illustrate that the measured width of CNTs from AFM images cannot be used to characterize the dimension of CNTs.

Figure 7.4(b) shows the trace profile of a diamond AFM tip with tip radius R_T = 70 Å as it moves across the same CNT of diameter D = 2.7 nm, perpendicular to its longitudinal axis with and without the vdW interactions. These simulation results indicate that the measured height of a CNT by AFM is slightly larger or smaller than its true height, and the measured width of a CNT is significantly larger than its true width. Using the results in Figures 7.4(b) and 1(c), we can in principle determine the true diameter of the CNT from the AFM measurement data, at least for the idealized case of a perfectly flat substrate and well-defined AFM tip shape.

Figure 7.4(c) shows the effect of the compressive force on the magnitude of CNT deformation along with some of the deformed intermediate configurations of these nanotubes. Since CNTs are extremely compliant in the radial direction, small compressive forces on them can cause fairly large deformations in them and may even

cause them to fully collapse provided that the diameter of the CNTs is adequately large. More discussion on this phenomenon is in the forthcoming paragraphs.

Using the numerical simulations, we can now quantitatively examine the relationship between the CNT height measured by AFM and the true CNT diameter. Figure 7.5(a) compares the simulated results for AFM measurements of single-walled CNTs (SWCNTs) heights on a perfectly flat silica substrate. In these simulations, the equilibrium distance between the AFM tip and the substrate is assumed to be 2.7 Å which is the contact distance of carbon atoms on silica substrate from our model. The actual height of a CNT on substrate is a monotonically increasing function of the nanotube diameter. However, when the CNT-AFM tip vdW interaction is taken into account, the apparent height of the CNT measured from the AFM tip trace is no longer the true CNT height. In particular, for CNTs with D < ~2.0 nm, the AFM measurements should be comparable to the actual CNT height; however for CNTs with D > ~2.0 nm, the measured height data could be notably lower than the true values. Furthermore, there does not seem to be a one-to-one relationship between these two values and one cannot conclusively predict the actual diameter of the CNT from AFM measurements alone, particularly for CNTs with diameter > 2 nm.

Our simulations also indicate that the AFM tip material (i.e. using a different interatomic potential) and temperature do not play a major role in the behavior discussed above. (More detailed information on this issue is provided in Appendix B.) On the other hand, the number of walls of the CNT has an important role in the AFM height measurements. Figure 7.5(b) compares the true height of double-walled CNTs

(DWCNTs) that would be measured by diamond AFM tips. The diameter of a DWCNT is defined as the diameter of the outermost shell in the circular state. The inner shell is a CNT whose radius is ~3.4 Å (the graphite interlayer distance) smaller than the outer shell. As shown in Figure 7.5(b), DWCNTs are less compliant in the transverse direction than SWCNTs, and therefore their AFM height reading is expected to be closer to the true diameter of these nanotubes, for $D < \sim 3.0$ nm. Consequently, we expect the accuracy of diameter measurements from AFM "height" to be further improved for multi-walled CNTs (MWCNTs) with three or more shells, which suffer less deformation when placed on solid substrates. Indeed in a novel experiment, DeBorde, *et al.* devised a technique to identify individual SWCNTs and DWCNTs based on AFM measurements of height vs. downward compression force. [196]

Figure 7.5(c) compares the results of AFM tip measurements of the width of different SWCNTs over a perfectly flat silica substrate. The true width of a CNT is almost a linear function of the nanotube diameter. The apparent width of the CNT is highly dependent on the radius and shape of the AFM tip, but in general, it is much larger than the true CNT width. However, because of the one-to-one relationship between these two quantities, such measurements may be used to estimate the CNT dimensions from the AFM width measurements. To obtain truly accurate information of CNTs from AFM measurements, it is necessary to have the height and width measured by AFM as well as knowledge of the AFM tip radius and the number of walls of the CNT. The number of walls of carbon nanotubes may be obtained, for instance, by the experimental nano-indentation method employed in Ref. [198]. However, in practice,

such complete *a priori* knowledge is generally unavailable, and thus the AFM measurements of CNTs must be carefully interpreted.

Finally, Figure 7.5(d) illustrates the quantitative variation of the measured height of CNTs as a function of compressive force on the CNT. It is evident from these curves that compressive force has an important role on the height profile and it can further complicate the interpretation of AFM height measurements. The height vs. diameter relationship is highly non-linear in the presence of downward loads and therefore, the height profile cannot be effectively utilized for diameter characterizations under these conditions. According to our simulations, transverse forces as small as 5 nN can cause the CNTs to locally collapse under the AFM tip and if the CNT diameter is larger than a certain value, the collapse will propagate in the entire length of the nanotube. Even smaller forces can considerably change the height profile of the CNTs. It should be noted that the force that is referred to here is a direct load on the CNT. In AFM experiments, on the other hand, often downward loads as large as 60 nN are applied on the AFM tips. However, this downward load may not be fully exerted on the CNT alone. There are dynamical effects of the tapping mode as well as the effect of air, photoresist and substrate roughness that also need to be taken into account. These effects are not readily accessible for theoretical quantification, but as discussed in the subsequent sections, they seem to play an important role in AFM experiments.

7.4. Comparison with experiments

In order to validate our numerical results experimentally, we compared data from AFM measurements with resonant Raman spectroscopy, another method used to estimate CNT diameters in practice. [199-201] Single-wall CNTs were grown by chemical vapor deposition (CVD) from Fe catalyst on an SiO_2(90 nm)/Si substrate and electrically connected with Ti/Pd contacts. [147] The CNTs measured in this study are considered to be single-walled CNTs because they exhibit current saturation of ~ 25 μm. [150] The height of the CNTs on the substrate was measured using an Asylum AFM with a silicon tip of radius ~10 nm. For good tracking, the tapping mode scan was performed with a scan speed of 2 μm/s and a scan size of 1 μm x 1 μm. [197] A typical resulting image for one CNT is shown in Figure 7.6(a). There is some error to this measurement coming from the surface roughness of the oxide substrate (Δ ~ 3 Å). The CNT height values were obtained by taking averages from seven readings at different regions of the CNT.

For comparison, the actual diameters of several single-wall CNTs were also measured using resonant Raman spectroscopy with a He-Ne laser wavelength of 633 nm. A typical scan setting is 16 accumulations of 5 min long scans using ~2 mW power. Using the radial breathing mode (RBM) peak (Figure 7.6(b)), the diameter is given by the relation $\omega_{RBM} = (227/D)\sqrt{1+C_e D^2}$ where $C_e = 0.065$ accounts for the environmental effect SiO_2 has on ω_{RBM} [200]. The G peak (discussed below) was also used as another rough measurement of the CNT diameter. The diameters measured by Raman are the un-

collapsed diameters of the CNT. Comparing our AFM measured heights with the Raman extracted diameters in Table 7.1 along with values from literature, [202-203] we see that the AFM measurement tends to under-predict the actual diameter of the CNT, as suggested by the simulations in the previous sections.

7.5. Discussion and recommendations

The results of this work suggest that, while AFM measurements remain the most convenient method for estimating the diameters of substrate-supported CNTs, the numerical data must be carefully interpreted. For single-wall CNTs supported by a smooth substrate the CNT diameter information obtained from AFM is generally expected to be reliable up to diameters $D \sim 2$ nm, whereas for double-wall CNTs (which are stiffer) this limit is extended up to $D \sim 3$ nm. An uncertainty of 0.3–0.8 nm exists around these ranges due to SiO_2 roughness and surface contamination as described below.

In practice, AFM measurements on CNTs are most commonly performed in air, where water, photoresist, and other surface contaminants can alter the measured results. Even if these were eliminated, the effect of substrate surface roughness ($\Delta \sim 3$ Å for the amorphous SiO_2 substrate tested above) lends some uncertainty to the measured diameters. For instance, Ishigami *et. al.* [204] found that graphene height variation on SiO_2 as measured by AFM is ~8 Å before and ~3 Å after cleaning of photoresist residue in an argon/hydrogen atmosphere at 400 ^{o}C. Moreover, the imaged height of graphene

on SiO_2 was ~9 Å in air and ~4.2 Å in vacuum, the latter more closely comparable to the interlayer spacing in graphite of 3.4 Å.

While graphene tends to adhere [204] to the SiO_2 substrate and follow its corrugations, CNTs can span between microscopic valleys and hills, causing them to become "more circular" than what has been calculated in this paper with a perfectly smooth substrate. There are also experimental reports [205] on the effect of a water meniscus on the behavior of AFMs and nanotubes which can affect the diameter measurements and interpretation. Such effects are difficult to model in a simple, generalized approach as that presented here, but ought to be kept in mind as additional challenges in the interpretation of experimental data. However, we believe the computational results of this work provide insight into the mechanisms at play in AFM microscopy of CNTs on solid substrates and should lead to a better understanding for the design and analysis of carbon-based nanodevices and nanostructures.

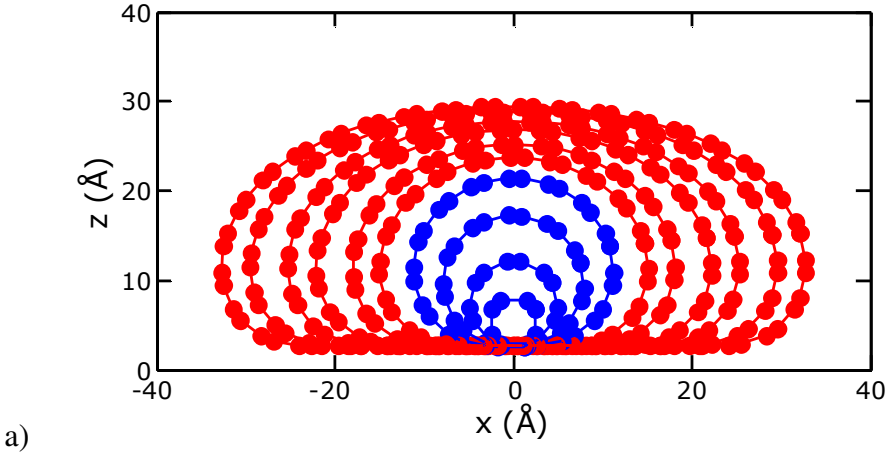
Before concluding, we wish to briefly comment on approaches that are alternative or complementary to AFM for obtaining the true diameter of CNT devices. Among these, transmission electron microscopy (TEM) has been used to measure the diameter and number of CNT walls beginning with the initial studies of Iijima. [118] The challenge of TEM is that CNTs must be placed on ultra-thin, electronically transparent membranes, and cannot be directly applied to typical CNT devices on common SiO_2/Si substrates. Nevertheless, recent work has succeeded in combining *in situ* TEM and device studies [206] for some specialized test structures. Scanning tunneling microscopy (STM) [207] can also be used to obtain the diameter of individual

CNTs. However STM requires samples to be placed on a conducting substrate, thus preventing direct application to CNT devices which are typically on SiO_2, sapphire or quartz. The STM measurements may also suffer from the same tip–sample interaction studied in this paper as well as additional geometrical effects on the tunneling probabilities. [208]

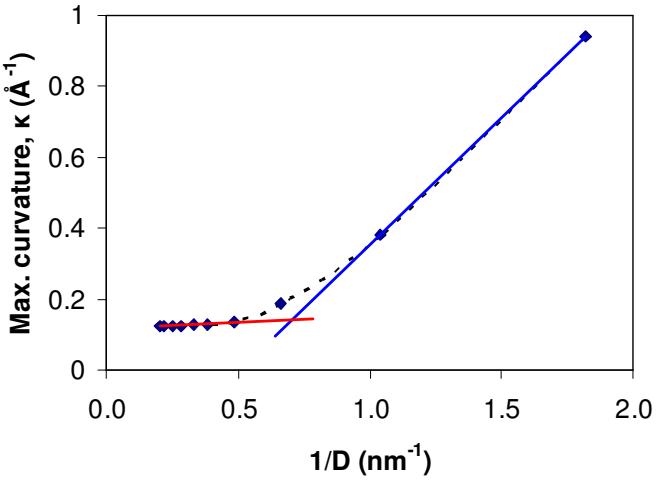
Optical measurements offer another alternative, non-invasive approach for obtaining CNT diameters. These include three methods, photoluminescence (PL) [209-210], Rayleigh scattering [211], and resonant Raman spectroscopy [200]. The restriction of PL is that it can only be used to measure semiconducting and not metallic CNTs. Rayleigh scattering and resonant Raman have the added benefit that they can be used to detect both metallic and semiconducting CNTs. The drawback of Rayleigh scattering is that for individual CNT detection, environmental perturbations significantly lower the signal to noise ratio. While removing the substrate is a solution to increasing the signal-to-noise ratio, this is not a preferred option for device studies. Finally, Raman spectroscopy requires the CNT to be resonant with the laser excitation energy, but does not require the removal or special preparation of the substrate. In our studies described above, only approximately 1 in 10 CNTs displayed sufficient resonance with the laser energy to reveal a radial breathing mode (RBM). Such Raman measurements of RBMs are also limited by the cut-off filter present in Raman systems, allowing only detection of $\omega_{RBM} > 100$ cm^{-1} and thus restricting the measurable CNT diameter range to < 2.78 nm. In addition to the RBM, the G peak also exhibits diameter dependence. The G peak is split into the main G^+ peak and the smaller G^- peak and the diameter

dependence is given by the relationship $\omega_G = 1591\text{-}C/D^2$. Here C is determined by the peak (G^+ or G^-) and the chirality (semiconducting or metallic) of the CNT. We used the following values from literature: $C_{G^+} = 0$, $C_{G^-}^{S} = 47.7\,\text{cm}^{-1}$, and $C_{G^-}^{M} = 79.5\,\text{cm}^{-1}$. [201] However, it should be stated that while this relationship works well for semiconducting CNTs, it does not work that well for metallic CNTs because of increased sensitivity to doping, [212] as can be seen in Table 7.1.

7.6. Figures and tables

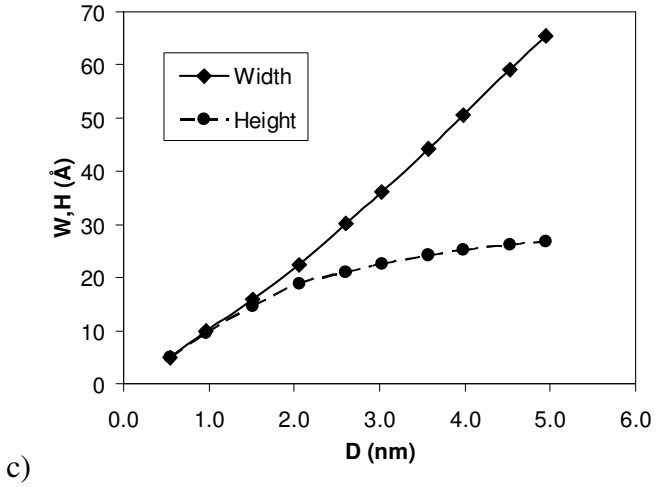


a)

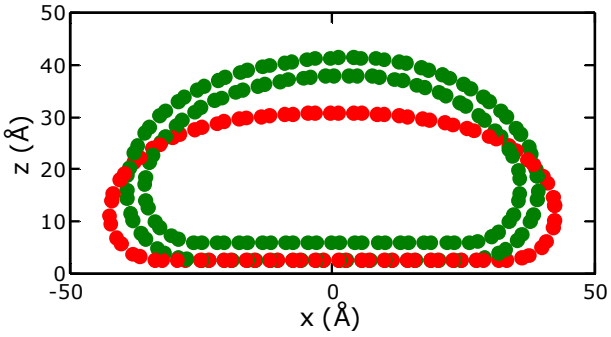


b)

Figure 7.1. (cont. on next page)



c)



d)

Figure 7.1. CNTs on SiO_2 substrate: a) Some of the typical equilibrium state single walled CNT geometries. The inner (blue) CNTs represent the nanotubes with diameters less than 2.2 nm. The outer (red) CNTs represent the nanotubes with diameters larger than 2.2 nm. b) Maximum curvature as a function of inverse of the tube diameter. Maximum curvature always happens at the side edges of the tubes. c) Variation of height and width of SWCNTs on SiO_2 with the diameter of the nanotubes. d) A typical (45,45) armchair SWCNT at equilibrium state (red) compared to a double-walled CNT with the same outer shell CNT (green).

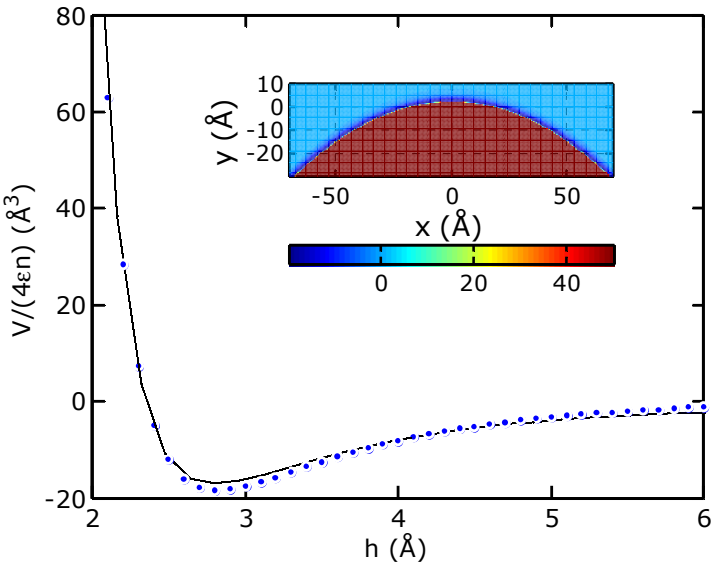


Figure 7.2. The van der Waals potential map around an AFM tip (inset). The light blue region is identically zero (after the incorporation of a cut-off radius, $r_c = 2.5\sigma_c$) and the dark magenta region is the AFM tip. The comparison of the numerical van der Waals potential (Eq. 7.5, the blue dots) and the analytical fit to it (Eq. 7.8, the black curve.)

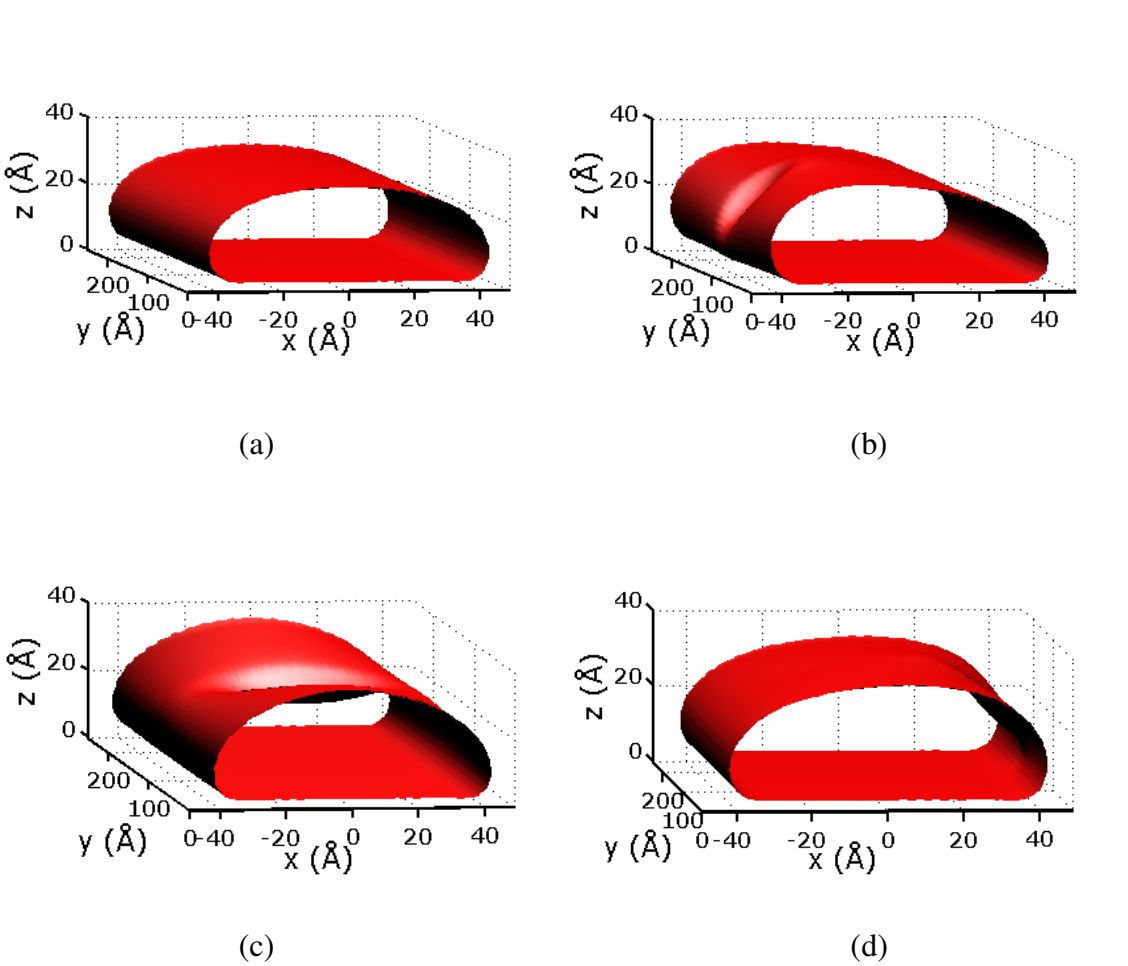
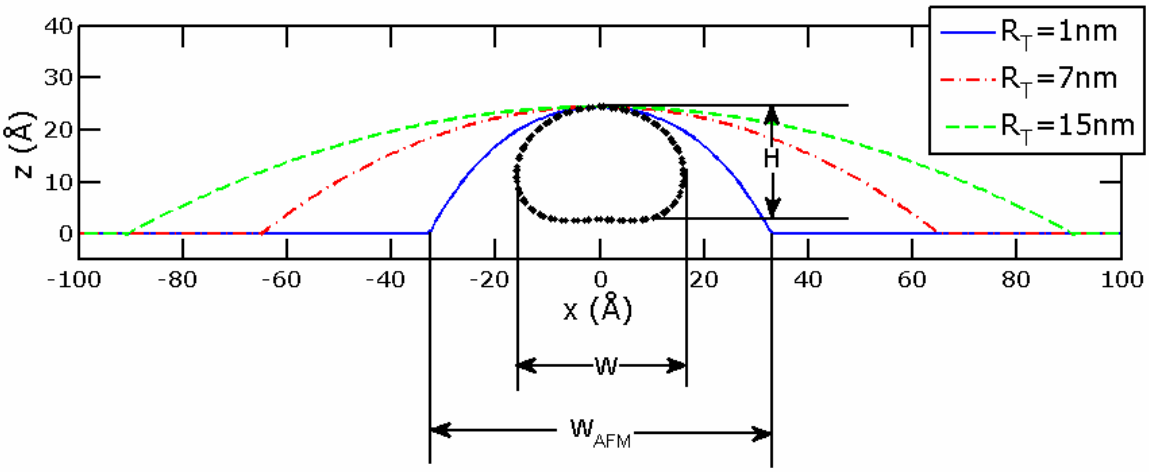
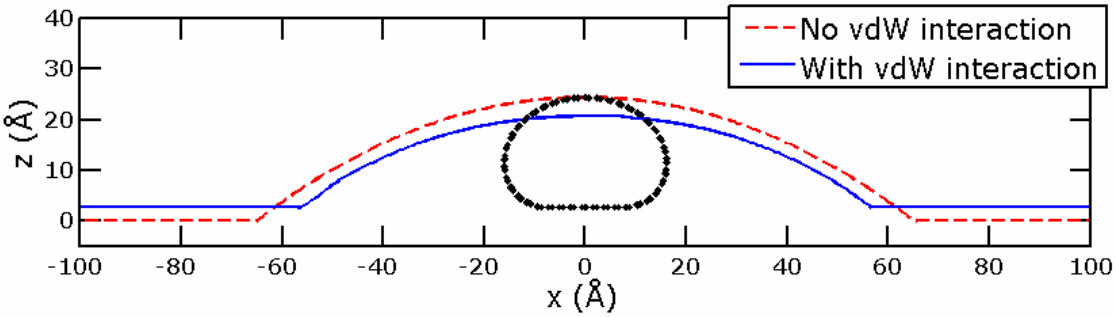


Figure 7.3. The different stages of the AFM scan. A relaxed armchair (45,45) CNT (a) snaps towards the tip (b), dips in (c), snaps away from the tip (d) during the motion of the AFM tip over it from left to right.

a)



b)



c)

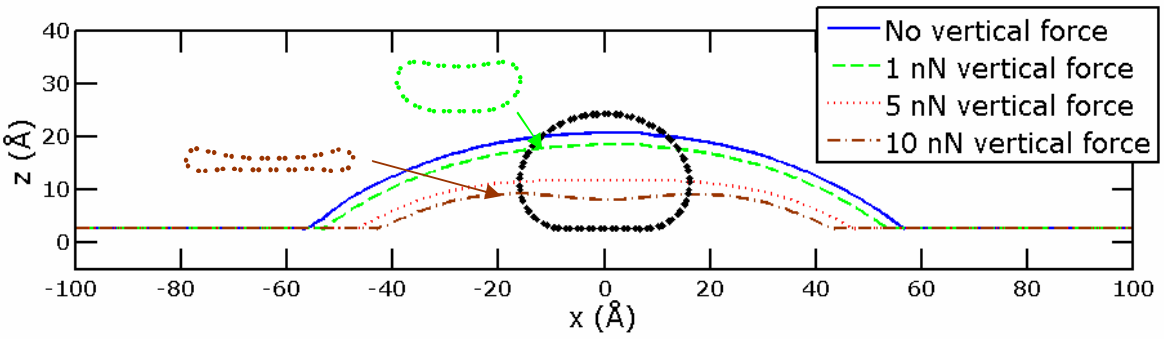
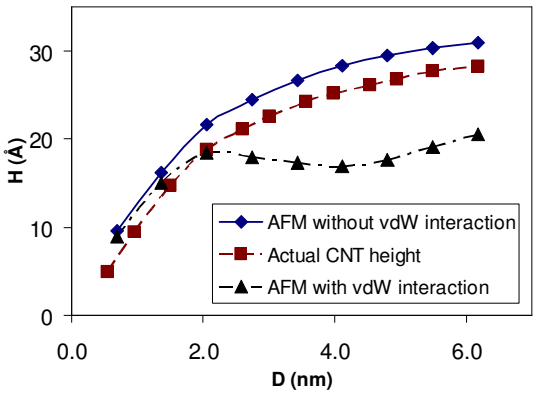
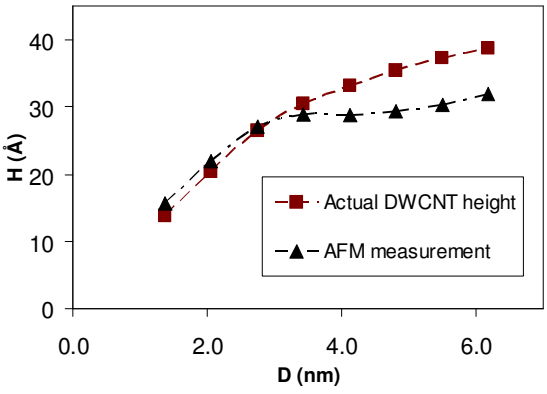


Figure 7.4. (cont. on next page)

Figure 7.4. a) The non-interacting parabolic AFM tip path for some different tip radii (R_T). The CNT is a typical (20,20) armchair nanotube with the actual diameter of ~2.7 nm. b) The AFM tip path for the same nanotube with and without considering the vdW interaction. The tip radius is R_T = 70 Å for both curves and the black dots show the relaxed position of the CNT. c) This figure compares the trajectory of a diamond AFM probe with a tip radius of R_T = 70 Å over a (20,20) CNT with the various values of downward vertical force.

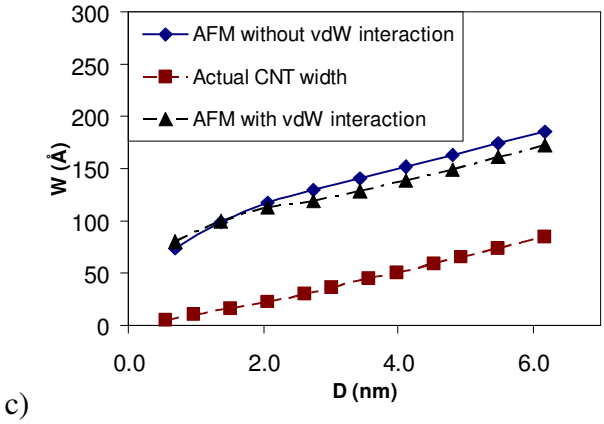


a)

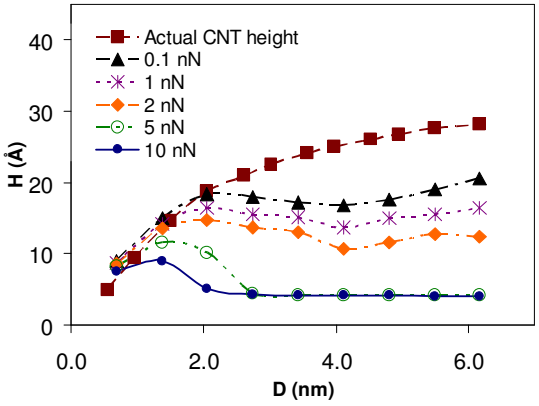


b)

Figure 7.5. (cont. on next page)



c)



d)

Figure 7.5. a) The variation of the height of CNTs on atomically flat quartz with the nanotube diameter. b) The variation of the height of DWCNTs with the nanotube diameter. c) The variation of the width of CNTs with actual nanotube diameter. d) The effect of downward compressive force from AFM tip on the CNT. In all curves, the dark magenta squares show the real properties of the CNT. The blue diamonds show the apparent height or width with a non-interacting AFM tip. The black triangles depict what an interacting diamond tip would "see". The tip radius is R_T= 70 Å for all curves. Lines are just drawn to guide the eyes.

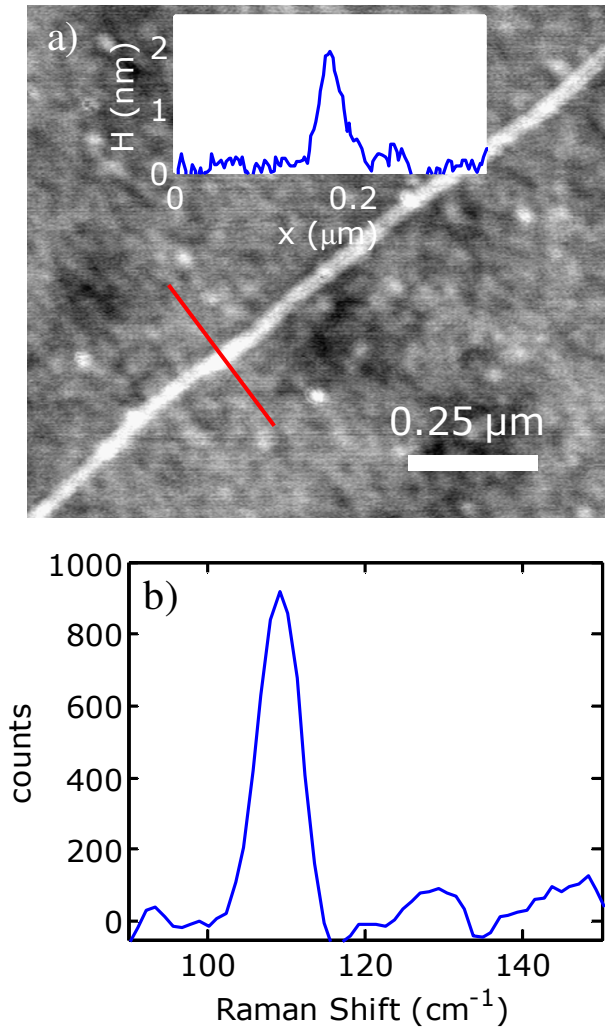


Figure 7.6. a) Topographical image of a CNT generated by an atomic force microscope (AFM). The measured height is 1.6 ± 0.2 nm. The inset shows the height profile along the red line. b) The radial breathing mode (RBM) peak from a resonant Raman spectroscopy measurement of the same CNT. The RBM peak indicates the CNT is a 2.27 nm CNT.

AFM Height (nm)	RBM Diameter (nm)	G^- peak Diameter (nm)	Simulated AFM Height (nm)	Chirality / S or M
1.2 ± 0.3	1.78	1.75	1.74	(19,5) S
1.4 ± 0.3	1.78	1.85	1.74	(23,0) S
1.7± 0.3	2.27	2.47	1.85	(18,18) M
1.2± 0.2[a]	1.12[a]	2.38[a]	1.30	(11,5) M[a]
1.3± 0.2[a]	1.37[a]	1.31[a]	1.50	(13,7) M[a]
1.5[b]	1.33[b]	1.35[b]	1.47	(NA) S[b]

Table 7.1. Comparison between the AFM measured height, Raman measured diameter, and the simulated AFM height. (a) Data from Ref. [197] (b) Data from Ref. [202].

CHAPTER 8: COLLAPSE AND INFLATION PROPAGATION OF SINGLE WALLED CARBON NANOTUBES ON SiO_2 SUBSTRATE

8.1. Introduction

Carbon nanotubes are axially very strong and stiff due to the sp^2 covalent carbon bonds within the hexagonal network. They are, however, extremely compliant in the out-of-plane direction and can undergo significant deformation, including kinks and bends, while remaining in the elastic regime [121, 213-218]. In 1995, Chopra *et al.* reported the observation of multi-walled nanotubes that were completely collapsed along their length [125] (see also [124, 219-221]). They also used a theoretical model to show that a collapsed nanotube may be energetically favored over one with a circular cross section, depending on the radius of the circular tube and the number of walls. Specifically, there exists a transition diameter D_1 above which the collapsed tube is favored and below which the circular tube is favored. Furthermore, a nanotube with a diameter smaller than D_1 may have a metastable collapsed state. Thus, there is a second transition diameter D_2. A tube with diameter below D_2 will not stay collapsed. With a diameter between D_1 and D_2, the tube will prefer the circular shape but will remain in a

metastable collapsed state. For a tube with diameter greater than D_1, the collapsed state will be energetically favored, and the circular state will be metastable.

Gao, *et al.* performed molecular dynamics and molecular mechanics simulations to find the equilibrium states for collapsed and circular free-standing single-walled nanotubes (SWNTs) of varying radii. [222] They also examined the effect of chirality (the orientation of the hexagonal structure of carbon bonds with respect to the centerline) by considering three chiral forms ((*n*,*n*) armchair, (*n*,0) zigzag, and (2*n*,*n*) chiral). For the armchair tubes, they found D_1 to be between 5.92 nm and 6.06 nm, while D_2 was found to be between 2.16 nm and 2.28 nm; the results for the other two chiral forms were similar.

From these equilibrium and metastable states, the behavior of ideal tubes can be predicted. A free standing SWCNT with a diameter greater than D_1 in the metastable circular state would prefer the collapsed state. If a section of the tube collapsed, the collapse would propagate through the length of the tube (Figure 8.1.) Likewise, a nanotube with a diameter between D_1 and D_2 in the metastable collapsed state prefers the circular shape. If a portion of it were inflated to a circular cross section, this inflation would propagate through the entire tube. In the current study, we examine the energetics of the CNT collapse (or inflation) process on SiO_2 substrate. Presence of a substrate, changes the collapsed/uncollapsed equilibrium shapes [127] as well as the critical diameters. AFM experiments are performed to verify these findings.

8.2. Energetic analysis of nanotube collapse

The CNT collapse process is controlled by the balance of strain energy and the intramolecular van der Waals (vdW) attraction. For a free standing carbon nanotube, the strain energy is obviously minimum for the undeformed (circular) tube. For CNTs on a surface, on the other hand, the uncollapsed state is not circular. Some of the typical in equilibrium uncollapsed shapes are shown in Figure 8.2(a). Maximum curvature has a linear relationship with the inverse of the tube diameter for circular tubes, but Figure 8.2(b) reveals a bilinear relationship for the tubes on a substrate. The initial linear regime in this figure corresponds to CNTs with large diameters. In this regime, the maximum curvature (which occurs at the left and right edges of the nanotubes) is constant and the CNT shape is predominated by the surface vdW interactions. The trailing region in this figure corresponds to the smaller CNTs. In this regime, the elastic curvature energy is dominant and the nanotubes are roughly circular in shape. As a tube collapses, the increase in strain energy may be overcome by the vdW attraction of the opposing walls. The strain energy increases monotonically as the area decreases. The vdW energy, which is very close to zero until the opposing walls come within close proximity, creates a double well form for the total energy profile. (See Figure 8.3.)

To illustrate the energetics governing the single-walled carbon nanotube (SWCNT) collapse, we examined the energy of nanotube cross sections as a function of degree of collapse. For a quantitative measure of the degree of collapse, we used the area A of the interior of the tube divided by the area A_o of the corresponding tube in its uncollapsed configuration. Thus, A/A_o ranges from 0 to 1. The total energy versus A/A_o

is dependent on the specific collapse path taken and is not universal (Figure 8.3.) The optimal collapse path is the one in which, for each value of A/A_o, the tube assumes the shape with the minimum energy. [172] Generally speaking, the uncollapsed state is always a minimum energy state (either global or local) and getting to the other minimum states from the uncollapsed state requires overcoming an energy barrier in the form of activation energy or force.

Tersoff-Brenner's empirical multi-body bond order interatomic potential for carbon [9-10] has been widely used in the study of CNTs and has also been utilized here for modeling the covalent C-C interactions. For modeling the interaction of CNT-silica, the analytical potential given in [130] has been used:

$$V_{vdW}(z) = \sum_{i=Si,O} \frac{2\pi\varepsilon_i\sigma_i^3 n_i}{45}[2(\sigma_i / z)^9 - 15(\sigma_i / z)^3], \tag{8.1}$$

where

$$\begin{aligned} &\varepsilon_{Si} = 8.909\ \text{meV}, \sigma_{Si} = 3.326\ \text{Å}, n_{Si} = 0.0227\ \text{Å}^{-3} \\ &\varepsilon_O = 3.442\ \text{meV}, \sigma_O = 3.001\ \text{Å}, n_O = 0.0454\ \text{Å}^{-3} \end{aligned} \tag{8.2}$$

In addition to this potential, an intra-molecular Lennard-Jones (LJ) vdW potential has been used with the following parameters for carbon [129]

$$\varepsilon_C = 3.02\ \text{meV}, \sigma_C = 3.39\ \text{Å}. \tag{8.3}$$

Based on these potentials, we obtained the total energy versus A/A_o from quasi-static molecular mechanics (MM) calculations by an energy minimization technique called Polak-Ribiere Conjugate Gradient method (CGPR). In dynamical cases, the

equations of motion were integrated using a fifth-order predictor-corrector algorithm. These optimization computations yielded the results in Figure 8.4. In this figure, width is the longest possible inscribed horizontal line. For uncollapsed tubes, width changes approximately linearly with respect to D and height is bi-linear. For collapsed tubes, width is linear and height is constant.

In Figure 8.5 the total internal potential energy per atom (which in turn is composed of many-body interaction and pair potential parts) for different diameter CNTs are shown. The critical diameter D_1 occurs when the energy of the collapsed and uncollapsed states are equal. With our model, this occurred when D was approximately 3.7 nm. For tubes with $D > 3.7$ nm, the energy in the collapsed state is lower than that of the uncollapsed state. The second critical diameter D_2 occurs when the left well becomes flat, and the collapsed state is neutrally stable; for our model, this occurred when $D = 2.1$ nm. For tubes with diameters less than D_2, the second well disappears, and only the uncollapsed state is stable.

8.3. Speed of the collapse/inflation propagation

The question of whether a collapse will propagate is governed by the change in energy from the uncollapsed to the collapsed state. If ΔE is defined as the energy per atom of the inflated tube minus the energy per atom of the collapsed tube (see Figure 8.5), then the collapse of a nanotube will propagate if $\Delta E > 0$. Conversely, an inflation will propagate if $\Delta E < 0$.

The energy of the collapsed tube has both bending (curvature) and vdW components. The bending energy of the collapsed tube is not a function of D; the bulbs at the ends of the cross section are the same shape and size for tubes of various radii. Gao, *et al*. [222] found that the radius of these bulbs to be approximately 0.525 nm. The vdW energy of the collapsed tube increases approximately linearly with D. In order to obtain an estimate of the speed of collapse/inflation propagating we proceed as follows. The propagation is assumed to have reached a steady state (Figure 8.6). It is viewed from a coordinate system that follows the propagation, such that the shape of the transition region remains the same, and the atoms move from left to right. The tube transitions from uncollapsed to collapsed over a length L. We call L the length of active collapse region and it is shown in Figure 8.6b. To the left of cross section A-A, the tube is uncollapsed (metastable); to the right of B-B, the tube is collapsed (stable). If a row of atoms in the nanotube travels from A-A to B-B over a time Δt, then the speed of the collapse propagation is

$$v_{cp} = \frac{L}{\Delta t}. \tag{8.4}$$

In order to have a quantitative measure of L, here it's defined as the distance between two cross sections in which the inscribed area is equal to 105% of the fully collapsed inscribed area and 95% of the inflated inscribed area. Figure 8.7 shows the variation of L with the diameter of the CNT. The average speed of a particular atom i over this interval of time is

$$\bar{v}_i = \frac{l_i}{\Delta t}, \tag{8.5}$$

where l_i is the distance through which the i^{th} atom travels (see Figure 8.8). Each row of atoms undergoes a change in potential energy as the CNT collapses. The gained kinetic energy for one row of atoms after completion of the collapse is:

$$KE = \frac{1}{2} m \sum_{i=1}^{N} v_i^{\,2}, \tag{8.6}$$

where m is the mass of one atom, N is the number of atoms in one row, and v_i is the speed of atom i at the end of the collapse process. Assuming a linear profile for velocity, $v_i = 2\bar{v}_i$ and assuming a perfect conversion of potential energy to kinetic nuclear energy, i.e. $|\Delta E| = KE$, an upper bound for the speed of collapse propagation is obtained:

$$v_{cp} = \left(\frac{|\Delta E| L^2}{2m \sum_{i=1}^{N} l_i^{\,2}} \right)^{\frac{1}{2}}. \tag{8.7}$$

Figure 8.9 shows the form of this equation for the different CNT diameters. As it can been seen from this graph, speed of propagation is a monotonically increasing function of D, but it saturates for large nanotubes, attaining a maximum value of ~2000

m/s. This can be compared to the speed of sound wave propagation in CNTs which is typically on the order of ~5000-10000 m/s. [223-225]

In the case of chiral CNTs, sometimes Moiré patterns similar to the one shown in Figure 8.10 may form which feature a gradual and periodic change of the interlayer stacking pattern. In order to investigate the effect of this phenomenon on the collapse behavior of CNTs, we compared Kolmogorov, *et al.*'s registry dependent potential [115] with the conventional LJ potential. This registry-dependent graphitic potential has an r^{-6} two-body vdW attraction, an exponential atomic-core repulsion, and a short-ranged term describing the energy gain due to interlayer delocalization of π orbitals; this term dominates corrugation against interlayer sliding. We used this potential in some of our simulations of the collapse of varying radii CNTs and no meaningful change in the collapse behavior was observed and thus it is concluded that this phenomenon is of no significant effect in the collapse behavior of CNTs, although it has been reported to be of some importance in rolling and twist formation of carbon nanotubes. [116, 226-227]

8.4. Effect of temperature

A natural question which is of great practical importance particularly in device applications, is that can heat inflate a collapsed CNT? If so, approximately at what temperature does the inflation occur? It is a known fact that Joule heating can sometime generate a great amount of heat in nanowires or can even break them down. In order to answer these questions, we conducted some molecular dynamics simulations on collapsed CNTs on SiO_2. The temperature was gradually raised from absolute zero and

the approximate temperature that the inflation occurred was recorded. This computational experiment was repeated several times for the same diameter CNTs and with different diameter CNTs and the results are shown in Figure 8.11. This figure does not imply that there is a linear relationship between the nanotube diameter and the inflation temperature. It only suggests that there's a positive correlation between the inflation temperature and the CNT diameter, and the actual dependence may be more complicated. Since typically CNTs break down at a temperature about ~2000 ^{o}K, there's no need to consider larger diameter CNTs, as they all burn out before they get re-inflated by heat which is by the way, not desirable in nanoelectronics applications. To further hamper this behavior, a dielectric layer which maintains compressive stress and prevents the CNT from "inflating" again could be utilized. (I.e. embedding the circuit.)

8.5. Experimental observation of collapse/inflation

In this section we basically intend to verify the reversibility of the metastable collapse, provide the experimental setup details with an estimate of the diameter of the tubes tested. Electronic, thermal and transport properties of the collapsed tubes vs. uncollapsed tubes is another objective here. Our preliminary experiments with ultra-sharp diamond AFM tips confirmed the existence of a metastable collapse state which was reversible with furnace heating. (Figure 8.12.) However, silicon AFM tips were found incapable of repeating this result due to a more blunt/softer tip, wear-and-tear or other unknown factors and it was suggested to only use diamond AFM tips. Unfortunately, these tips did not become available to us during the course of the

completion of this thesis and therefore, no more experimental verification of this process was possible.

8.6. Discussion and conclusions

The propensity of a carbon nanotube on silica substrate to collapse or inflate depends on its diameter and the number of its walls. Collapsed carbon nanotubes have been predicted to exist theoretically and have been observed experimentally as well. We have specified the conditions for propagation of collapse and inflation of ideal SWCNTs over silica substrate. If a section of an uncollapsed tube with $D > 3.7$ nm is collapsed, the collapse will propagate. The theoretical limits on the speed of this propagation were established. Conversely, if a section of a collapsed tube with 2.1 nm $< D <$ 3.7 nm is inflated, the inflation will propagate. These critical diameters are generally smaller than those for the corresponding free-standing CNTs. These predictions have been confirmed by our AFM experiments. Similar results could be derived for double and multi-walled nanotubes; for an n-walled tube, the transition radii R_1 and R_2 are increasing functions of n [124]. The chirality of the nanotube was found inconsequential in this behavior.

Carbon nanotubes hold a great promise to be used as sensors, nanowires, nanotransistors and as building blocks in other nano-electro-mechanical (NEMS) devices and nanostructures. Thus such, an understanding of the stability of nanotubes is important for the study of both mechanical and electrical properties, since both of these are quite different for collapsed tubes versus uncollapsed tubes [179, 219, 228-232]. For example, according to Lammert, *et al.*, collapsing a nanotube drastically alters the low-

energy electronic properties and can reverse the metallic versus semiconducting behavior [233] and it can significantly enhance thermal coupling between CNTs and SiO_2 substrates [130]. This information could be helpful in the design of future carbon-based nanoelectronic devices.

8.7. Figures and tables

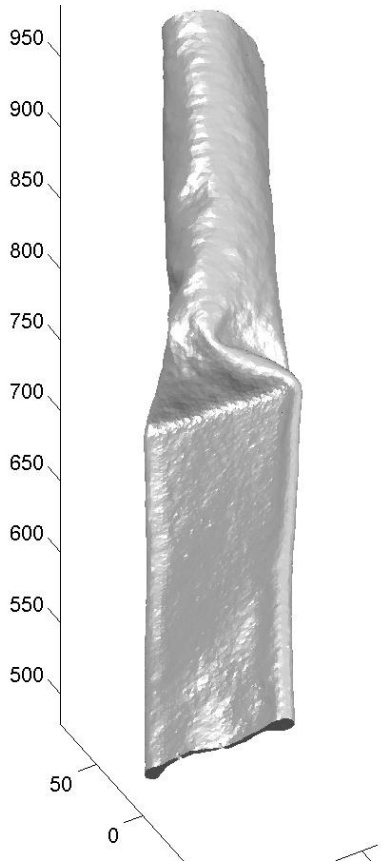
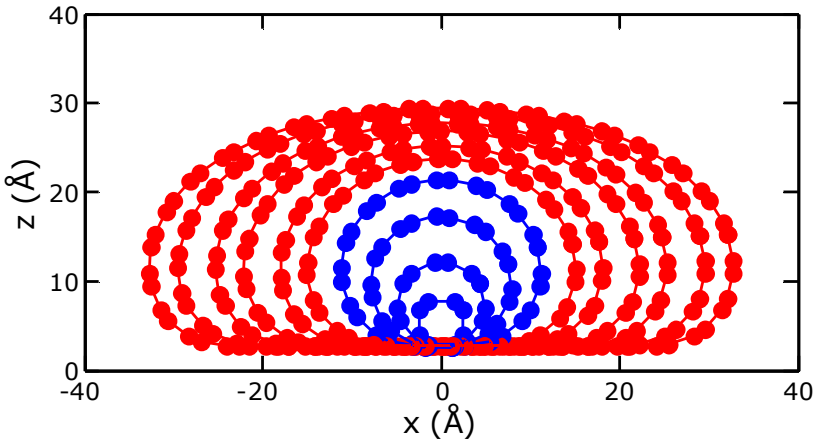
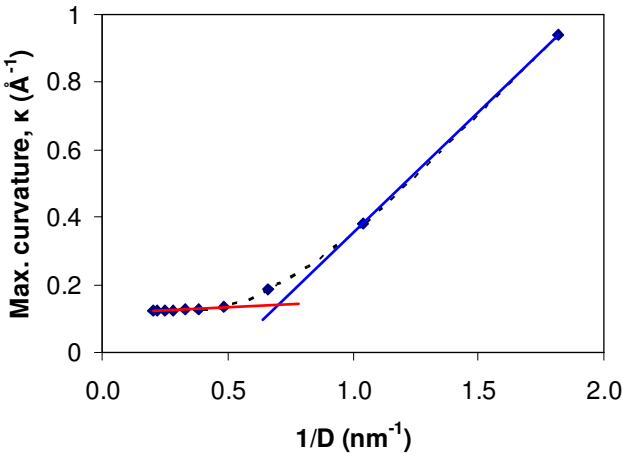


Figure 8.1. When free-standing carbon nanotubes larger than ~6 nm in diameter are pinched or otherwise disturbed at one point, collapse propagates in them. This phenomenon has both experimentally and computationally been observed. The units of the figure are in angstroms.



a)



b)

Figure 8.2. (a) Some of the typical equilibrium state CNT geometries on SiO_2 substrate. The inner (blue) CNTs represent the nanotubes with diameters less than 2.2 nm. In this regime, nanotubes are very close to circles. The outer (red) CNTS represent the nanotubes with diameters larger than 2.2 nm. (b) Maximum curvature always happens at the side edges of the tubes. For nanotubes larger than ~2.5 nm in diameter the maximum curvature is constant and for the smaller nanotubes, maximum curvature varies linearly with the inverse of the diameter of the tube.

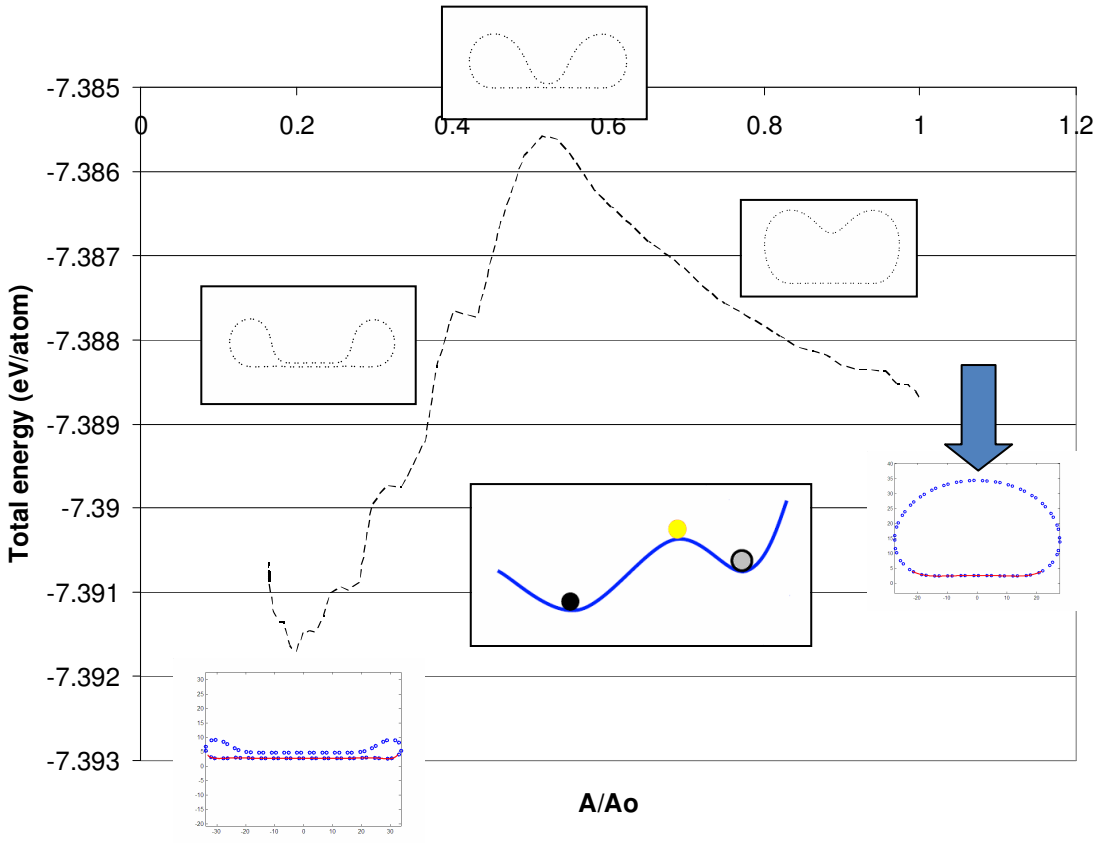
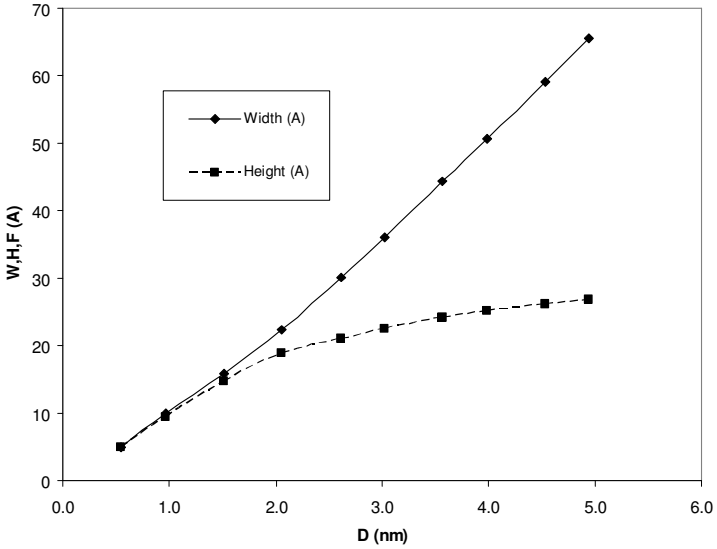
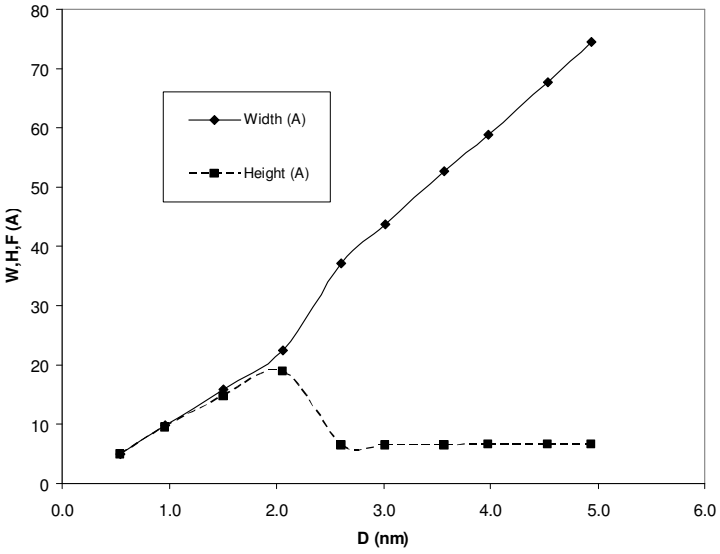


Figure 8.3. The different stages of collapse of a SWCNT on a substrate. The total energy has to overcome a double well potential (the middle inset).



a)



b)

Figure 8.4. Variation of height and width of SWCNTs on SiO_2 with diameter for a) uncollapsed and b) collapsed tubes. Width is defined as the longest possible inscribed horizontal line. For uncollapsed tubes, width changes linearly w.r.t *D* and height is bi-linear. For collapsed tubes, width is linear and height is constant with *D*.

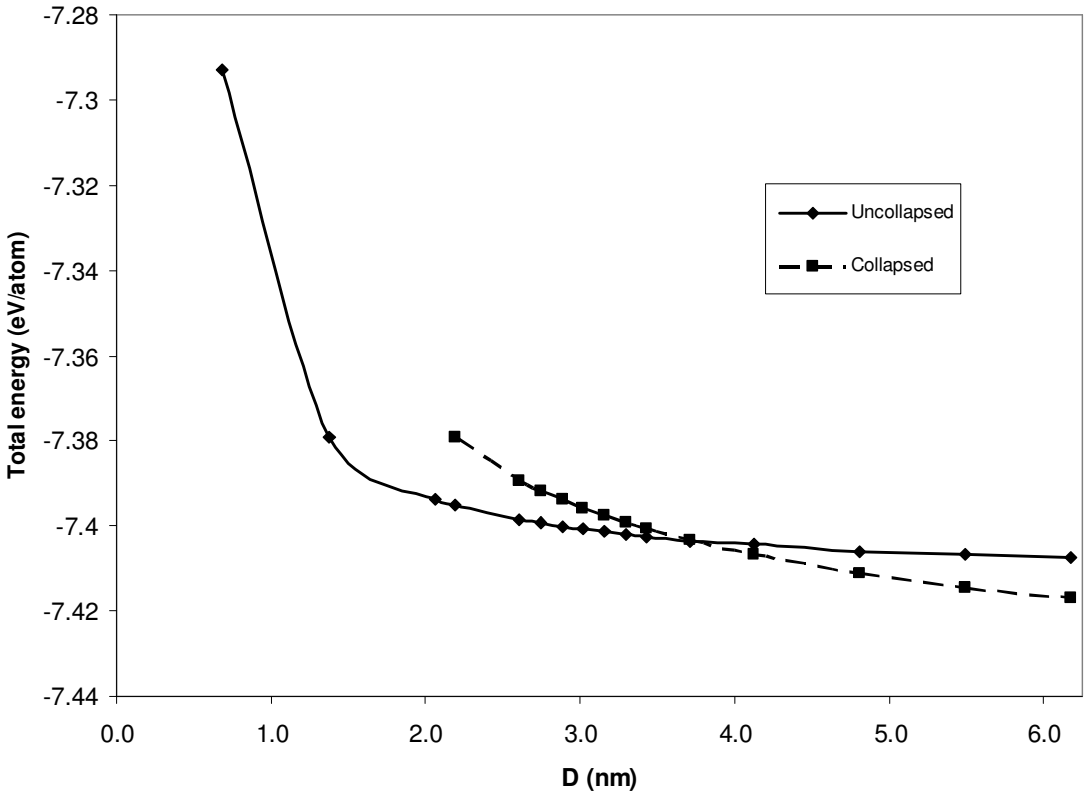
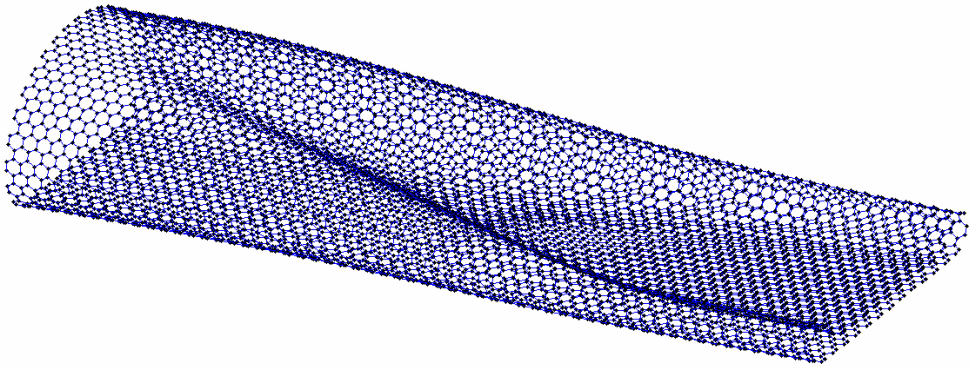
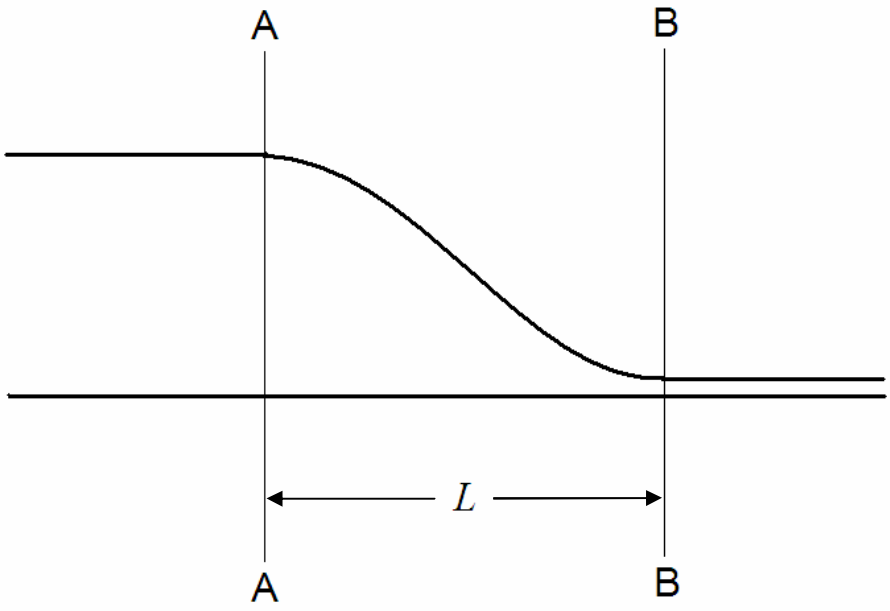


Figure 8.5. Variation of the total energy per carbon atom as a function of diameter for the uncollapsed and collapsed SWCNTs on SiO_2 substrate.



a)



b)

Figure 8.6. (a) Region of active collapse. (b) Schematic variation of the cross sectional area along the axis of the tube. *L* is called the length of the active collapse region.

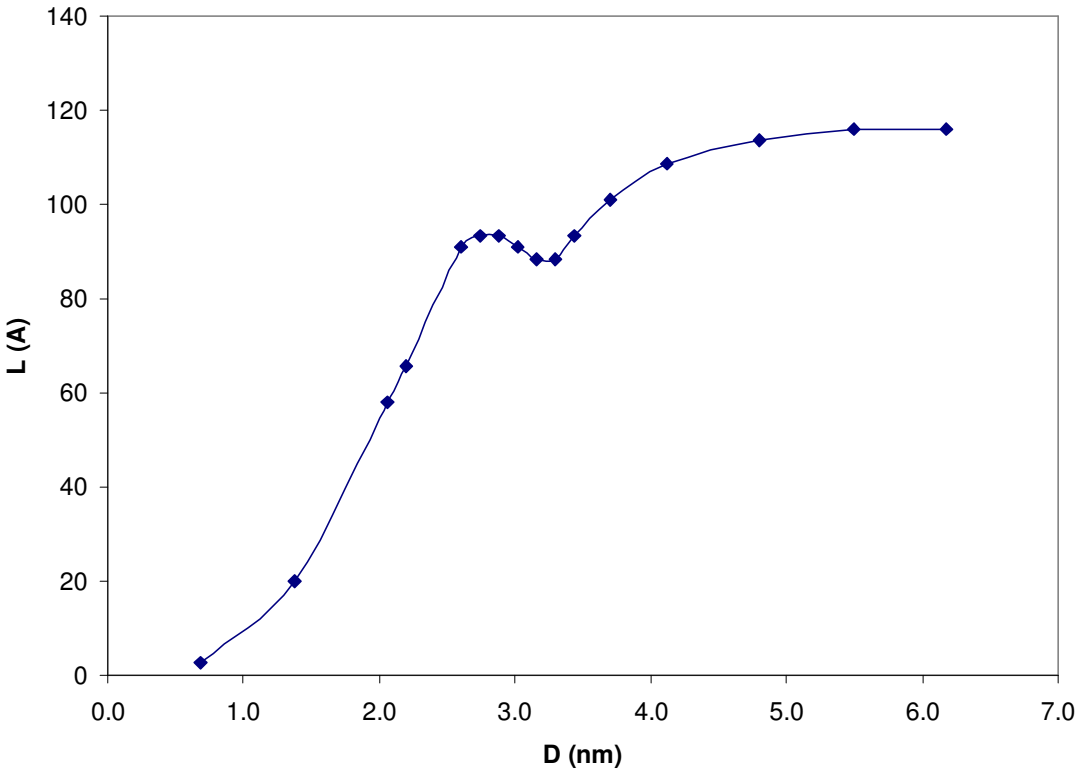


Figure 8.7. Variation of *L* with the diameter of the tube. The highly nonlinear region in the middle corresponds to the metastable nanotube diameter range (2.1 nm < *D* < 3.7 nm) where an increase in the tube diameter doesn't necessarily lead to an increase in the length of the active collapse region.

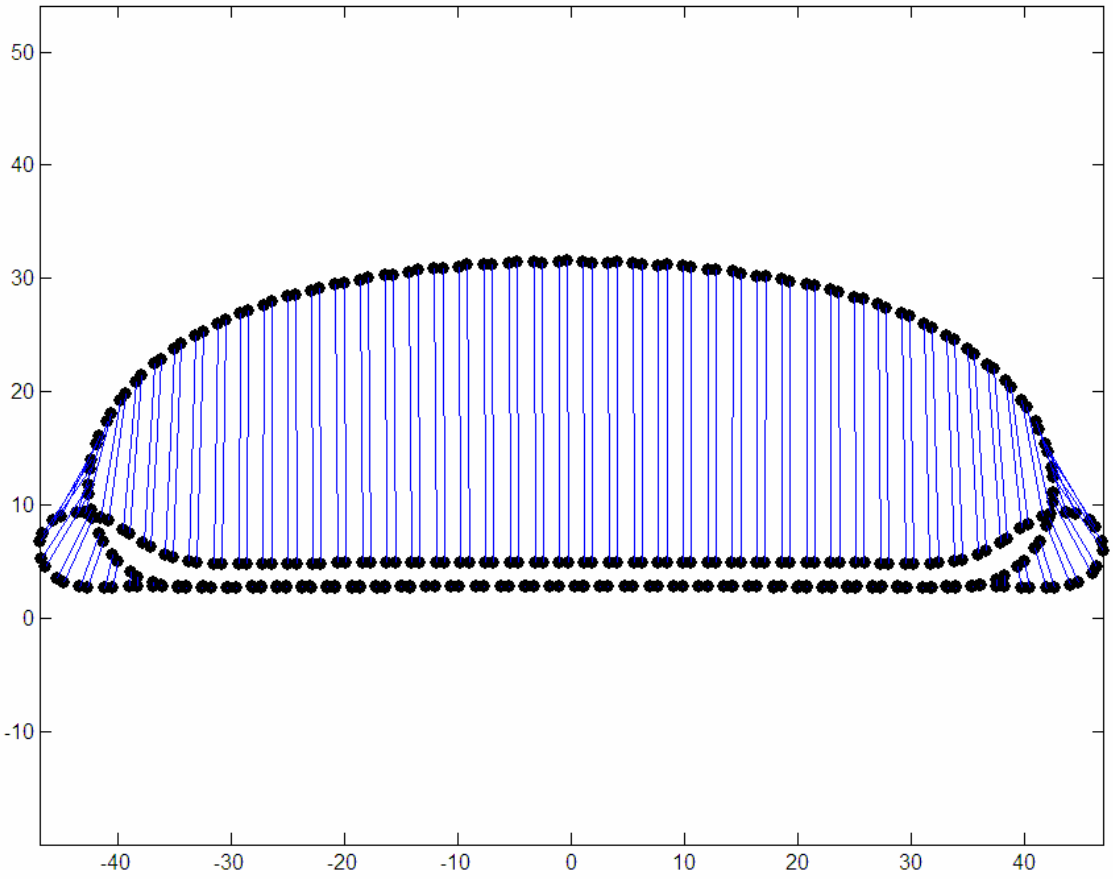


Figure 8.8. The uncollapsed and collapsed states of a (40,40) armchair carbon nanotube with an interplanar spacing of 0.335 nm. The lines designate the linear displacement of each carbon atom during the collapse process.

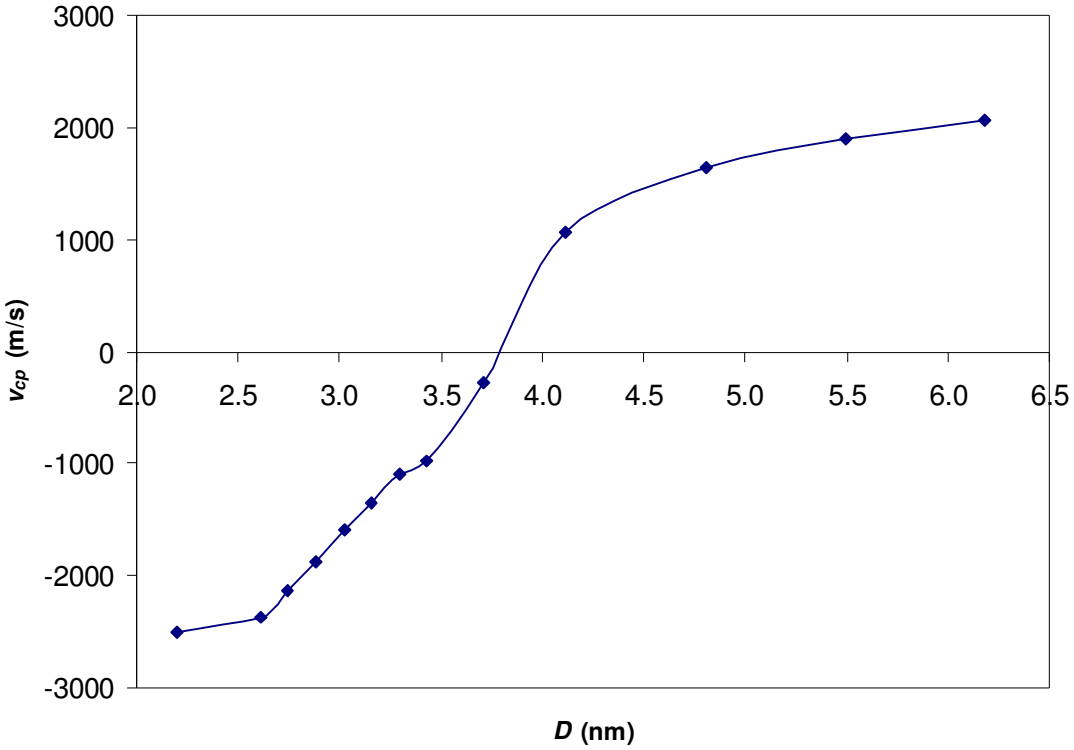


Figure 8.9. Upper bound of the speed of collapse (inflation) propagation as a function of diameter for the SWCNTs on SiO_2 substrate.

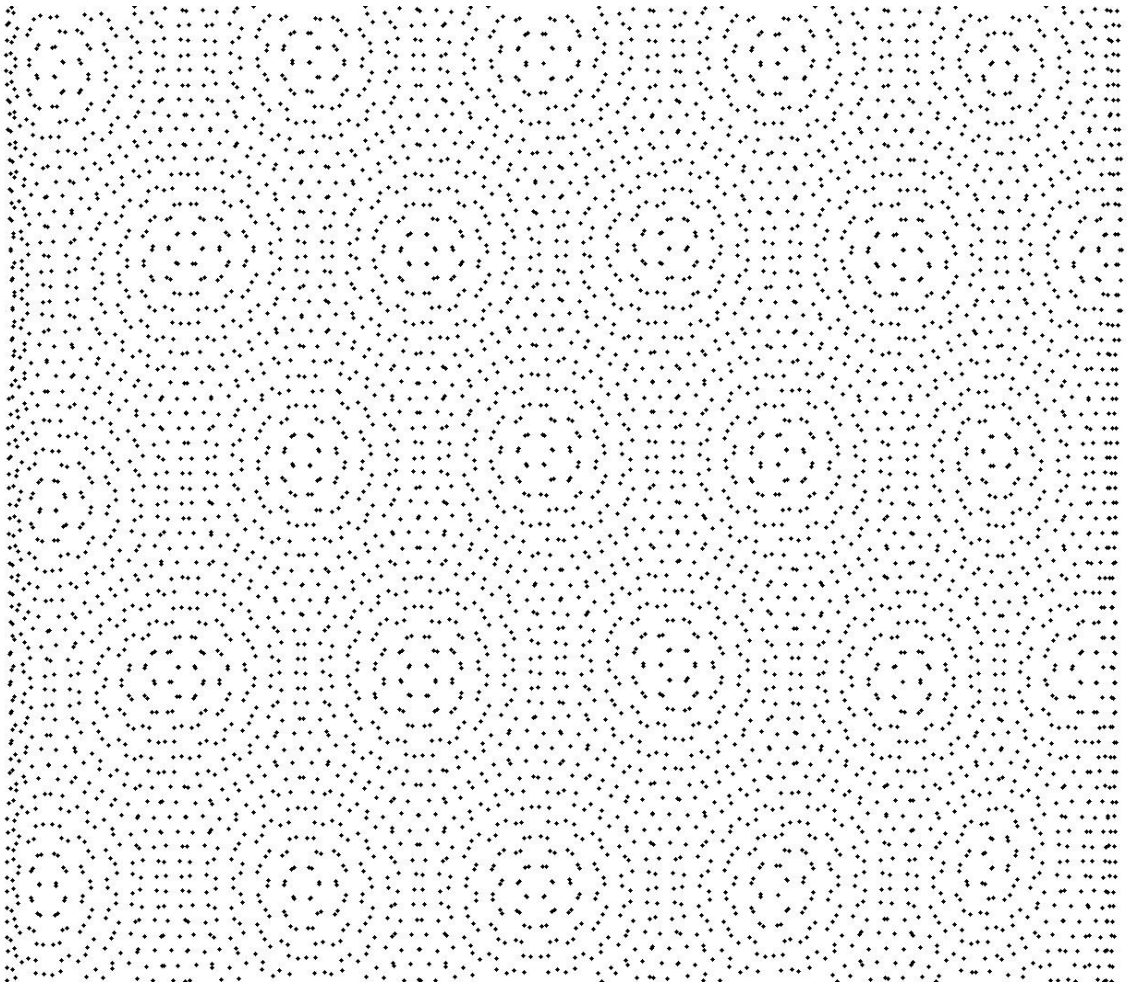


Figure 8.10. Depending on the chirality of the CNTs, Moiré patterns similar to the one shown here may form for the collapsed nanotubes.

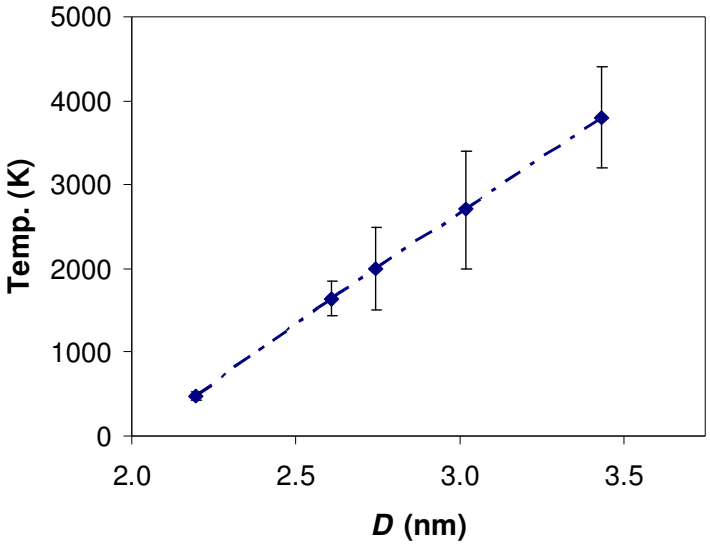
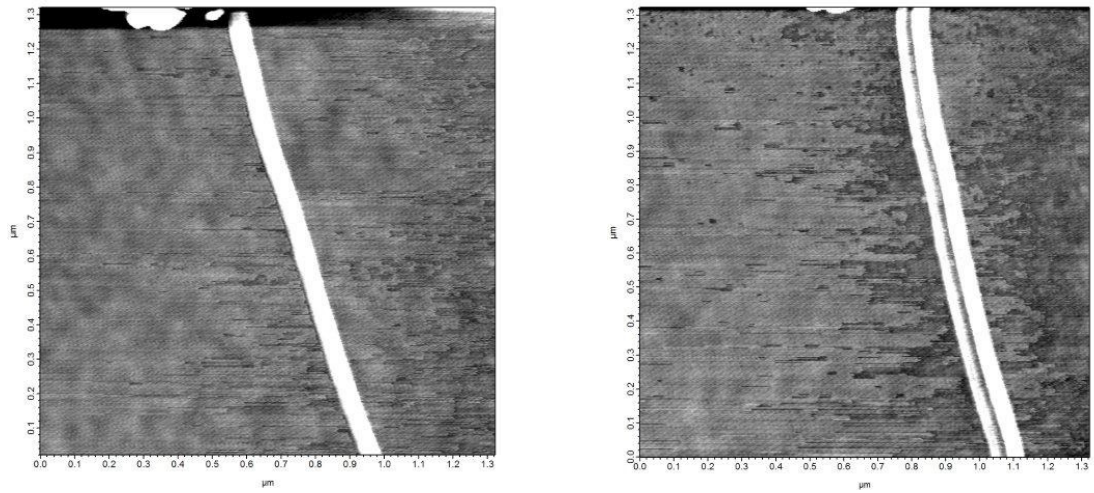


Figure 8.11. The relationship between inflation temperature of collapsed CNTs and their diameter. The dash-dot line is the linear fit to the numerical data and the error bars indicate the maximum scatter in the data for each CNT diameter calculated.



Before collapse After collapse

Figure 8.12. Experimental evidence for nanotube collapse. [By Albert Liao & Feng Xiong.]

APPENDIX A: THE ORIGINS OF THE QUADRATIC TIME SCALING OF HF-FEM WITH THE NUMBER OF DOFS

As mentioned in Chapter 2, the matrix formation and solution in HF-FEM is hard-coded using Matlab. As long as the Hamiltonian matrix is banded, the computational time for the matrix-vector product scales linearly with the number of degrees of freedom. Thus the effort for any iterative diagonalization technique should scale with Nn where N is the number of DOFs and n is the number of occupied orbitals, because one needs at least n matrix-vector products to get all the occupied orbitals. The orthogonalization step should also scale as Nn^2, because for each orbital, one needs to compute overlaps (N operations) with n other orbitals. Therefore, as noted by one of the reviewers, the reported quadratic time scaling is curious. To identify the sources of this quadratic dependence on number of DOF, we repeated the simulations of a helium atom with precise accounting for the computing time in the different parts of our code. The most time consuming part turned out to be the assembly process, i.e., the Gauss integration over the elements. The solution process of the matrix, on the other hand, requires comparatively little time and scales nearly linearly with the total DOF, as shown in Figure A.1. For a moderate system size (e.g., 50000 DOF), the time consumed

during the assembly process is in fact an order of magnitude higher than the solution time.

Subsequently, we performed a thorough timing of the assembly process. Figure A.2 shows the time needed to integrate each element (in milliseconds). Each point in Figure A.2 represents the time needed for Gaussian integration of one particular element. As can be seen, other than for a few elements, it takes almost a constant amount of time (less than 1.5 ms) to complete the Gaussian integration in an element, regardless of its size or position. This should result in linear time dependence on the number of DOF.

What is not scaled linearly with #DOF, however is the augmentation time in Matlab. Matlab's internal implementation of sparse matrices requires a quick sort algorithm for the indices [234]. Thus, the increased Matlab internal communications overhead of handling larger matrices in the memory appears to be the explanation for the non-linear dependence. To verify this, we used a third-party software package [235] for accessing large sparse matrices and Figure A.3 shows the same simulation as that in Figure A.1 using the third-party software, repeated with significantly improved performance. These results indeed demonstrate a linear time scaling with the number of DOF, and show significantly reduced total computing time.

A.1. Figures and tables

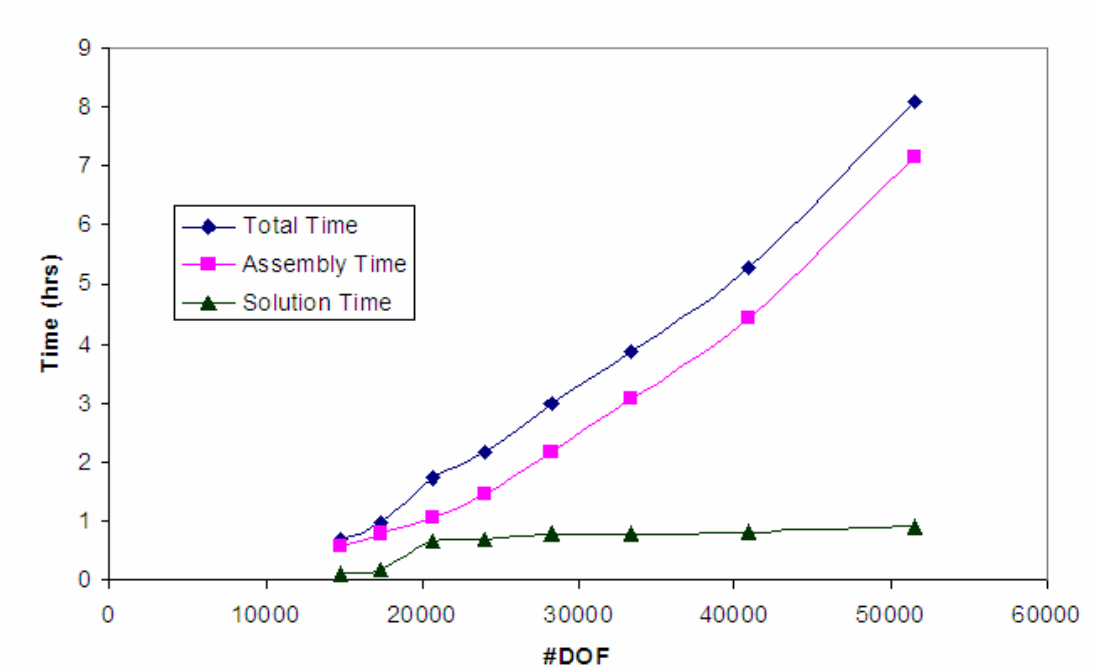


Figure A.1. Time-scaling of the assembly and solution phases in the original Matlab code. While the solution time is almost linear, the assembly time appears to be quadratic.

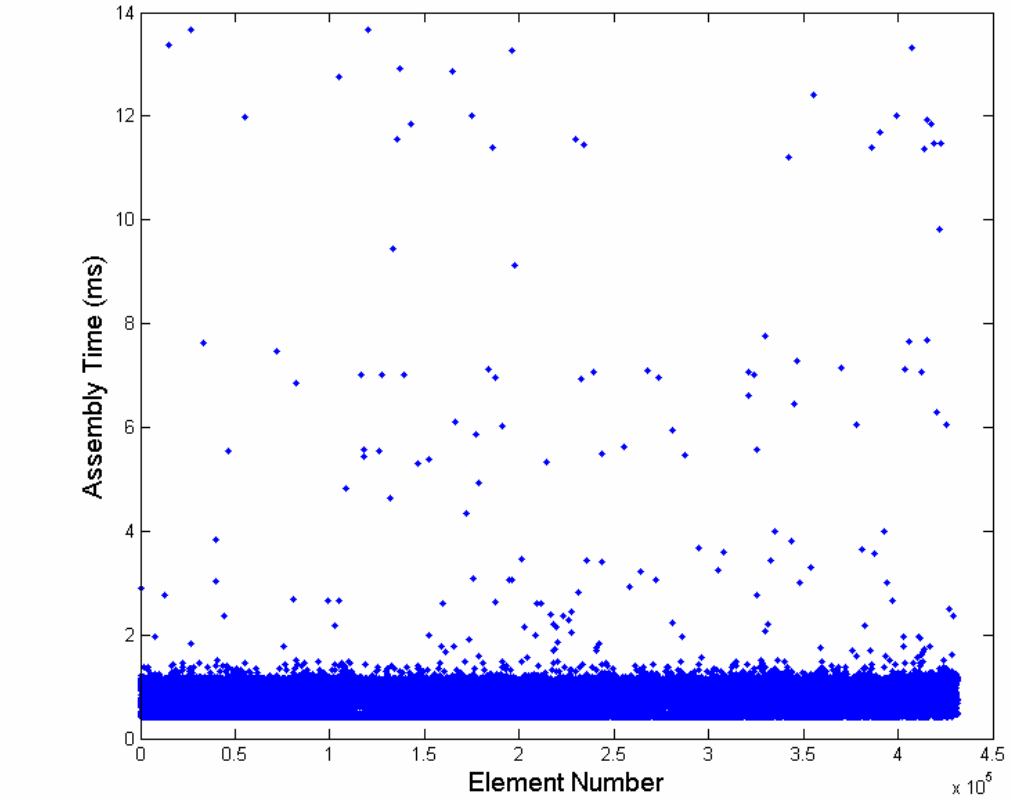


Figure A.2. The assembly time of each individual element in the original Matlab code. This is essentially a constant smaller than 2 milliseconds which is approximately equivalent to 6 million CPU-clock cycles.

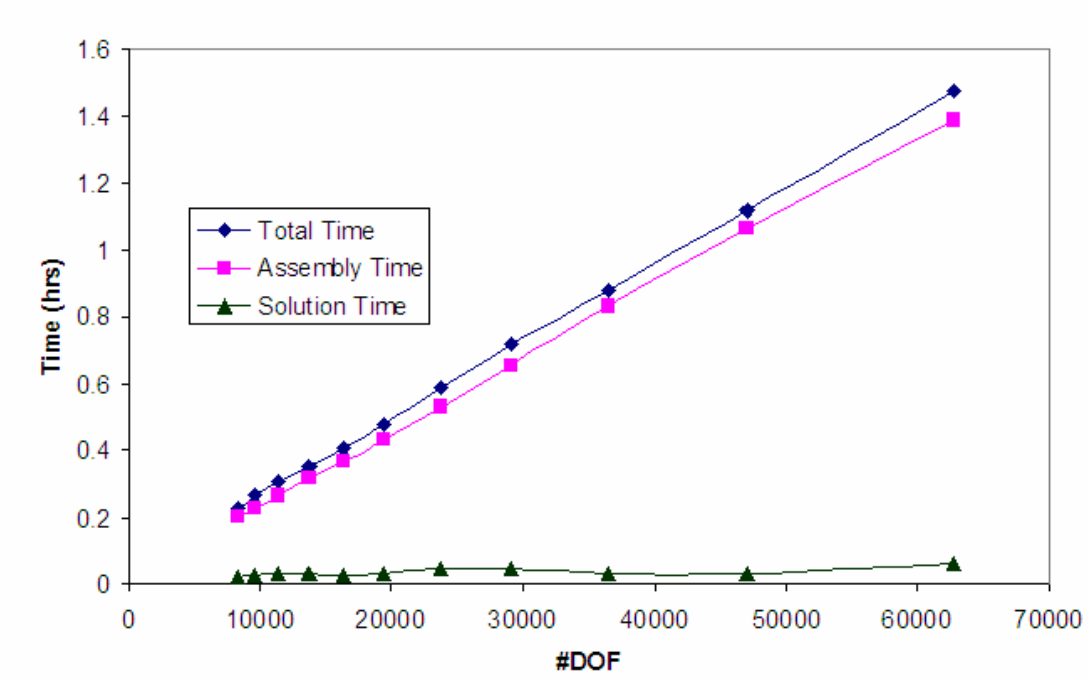


Figure A.3. Time-scaling of the assembly and solution phases in the optimized Matlab code. Linear time-scaling for both assembly and solution phases is attained.

APPENDIX B: THE DIFFERENT STAGES OF THE AFM SCAN AND THE MISCELLANEOUS EFFECTS

In order to better demonstrate how Figure 7.2b and the curves in Figure 7.4 were obtained in Chapter 7, Figure B.1 below is included as supplemental materials. It shows more clearly how a relaxed armchair (20,20) CNT snaps towards the AFM tip, dips in, snaps away from the tip and relaxes during the motion of the AFM tip over it from left to right.

B.1. The less important effects on AFM microscopy

The following effects were found to be less important than the number of walls of the CNT and the vertical force on AFM microscopy of individual CNTs.

B.1.1. The effect of tip material

The material dependence of AFM paths over CNTs is small. Figure B.2 compares the trajectory of diamond and silicon AFM probes with a tip radius of R_T = 70 Å as they move over a typical (20,20) armchair CNT. These paths are less than 5% different at the point of maximum which is insignificant for most practical purposes.

B.1.2. The effect of tip radius on height measurement

Figure B.3 shows that even though the width profile is strongly dependent on the radius of the AFM tip, the height profile is not.

B.1.3. The effect of temperature

The temperature dependence of AFM paths over CNTs is small. Figure B.4 compares the trajectory of a diamond AFM probe with a tip radius of R_T = 70 Å over a (20,20) armchair CNT at 10°K and at room temperature. Regardless of the small perturbations, these paths are effectively identical.

B.2. Figures and tables

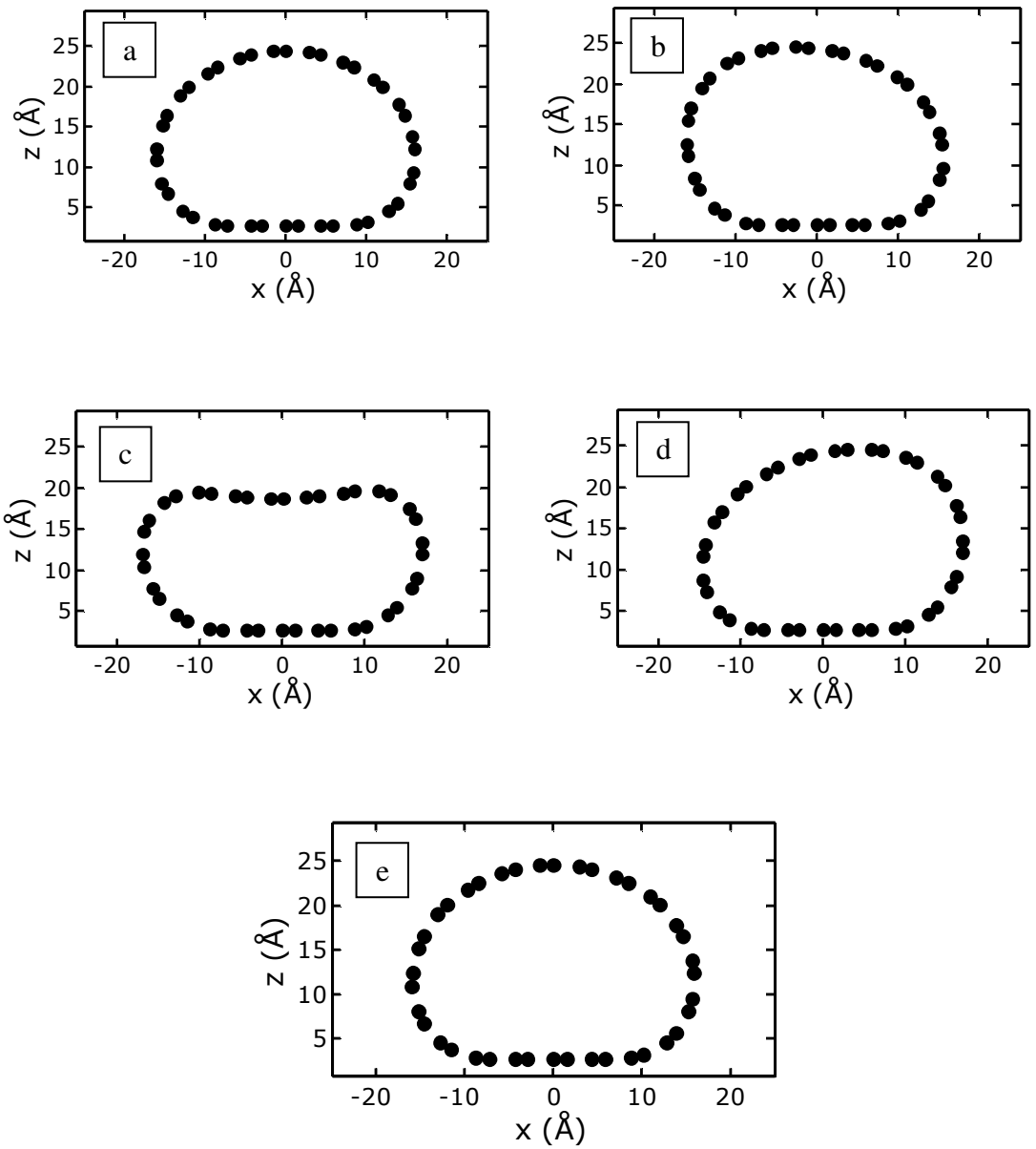


Figure B.1. The different stages of the AFM scan. A relaxed armchair (20,20) CNT (a) snaps towards the tip (b), dips in (c), snaps away from the tip (d) and relaxes (e) during the motion of the AFM tip over it from left to right.

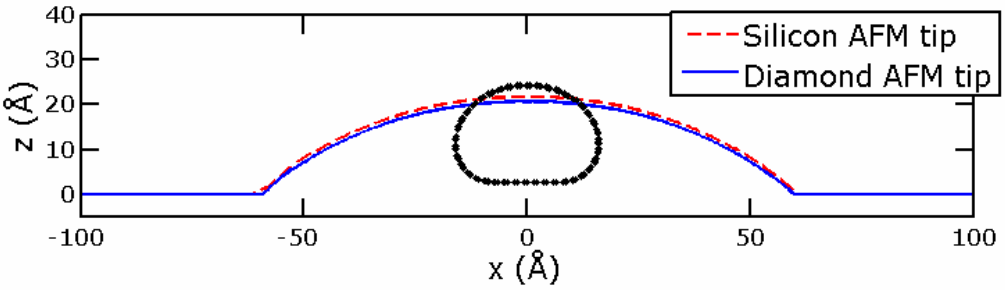


Figure B.2. The material dependence of AFM paths over CNTs is small. This figure compares the trajectory of diamond and silicon AFM probes with a tip radius of R_T= 70 Å over a (20,20) CNT.

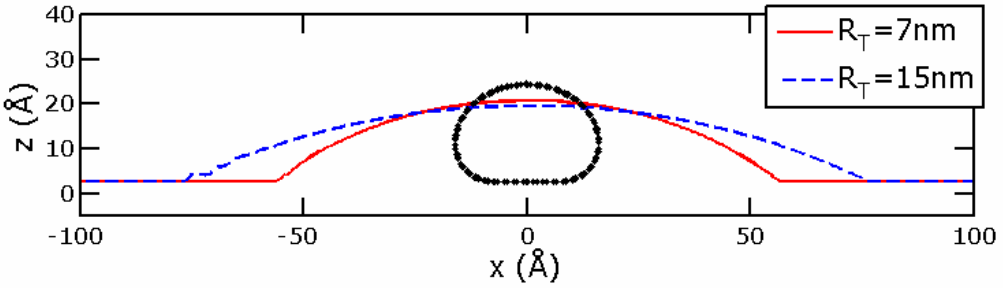


Figure B.3. The radius of the AFM tip has small effect on height profile. This figure compares the trajectory of diamond AFM probes with a tip radius of R_T= 70 Å and R_T= 150 Å over a (20,20) CNT.

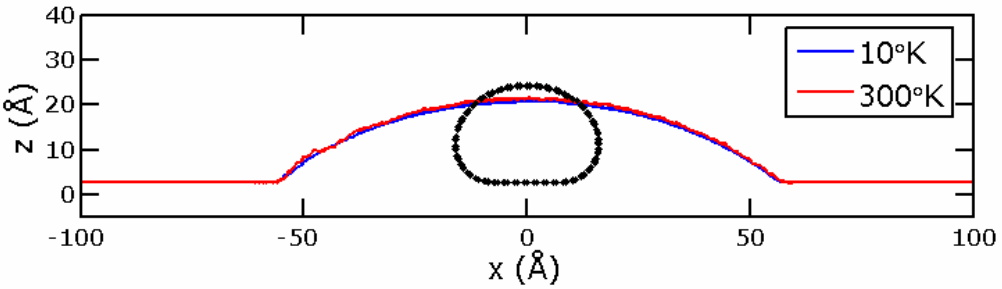


Figure B.4. The temperature dependence of AFM paths over CNTs is small. This figure compares the trajectory of a diamond AFM probe with a tip radius of R_T = 70 Å at 10°K and at room temperature over a (20,20) CNT.

References

[1] P. A. M. Dirac, *Proc. Roy. Soc.,* vol. 123, p. 714, 1929.

[2] R. M. Martin, *Electronic Structure; Basic Theory and Practical Methods*: Camb. Univ. Press, 2004.

[3] J. E. Pask*, et al.*, "Real-space local polynomial basis for solid-state electronic-structure calculations: A finite-element approach," *Phys. Rev. B,* vol. 59, pp. 12352-8, 1999.

[4] J. E. Pask*, et al.*, "Finite-element methods in electronic-structure theory," *Comp. Phys. Comm.,* vol. 135, pp. 1-34, 2001.

[5] J. E. Pask and P. A. Sterne, "Finite element methods in ab initio electronic structure calculations," *Model. Simul. Mat. Sci. Eng.,* vol. 13, pp. R71-R96, 2005.

[6] J. E. Pask and P. A. Sterne, "Real-space formulation of the electrostatic potential and total energy of solids," *Phys. Rev. B,* vol. 71, p. 113101, 2005.

[7] S. Jun, "Meshfree implementation for the real-space electronic-structure calculation of crystalline solids," *Int. J. Numer. Eng.,* vol. 59, pp. 1909-1923, 2004.

[8] K. Ohno*, et al.*, *Computational Materials Science; from Ab initio to Monte Carlo Methods* vol. 129. Berlin, Germany: Springer-Verlag, 1999.

[9] J. Tersoff, "New empirical approach for the structure and energy of covalent systems," *Phys. Rev. B,* vol. 37, pp. 6991–7000, 1988.

[10] D. W. Brenner, "Empirical potential for hydrocarbons for use in simulating the chemical vapor deposition of diamond films," *Physical Review B,* vol. 42, pp. 9458-9471, 1990.

[11] C. M. Goringe*, et al.*, "Tight-binding modelling of materials," *Reports Progress Phys.,* vol. 60, pp. 1447-1512, 1997.

[12] R. G. Y. Parr, W., *Density-Functional Theory of Atoms and Molecules*. New York: Oxford University Press, 1994.

[13] C. Møller and M. S. Plesset, *Phys. Rev. ,* vol. 46 pp. 618-622, 1934.

[14] P. Hohenberg and W. Kohn, "Inhomogeneous electron gas," *Phys. Rev.,* vol. 136 p. B864, 1964.

[15] W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlation effects," *Phys. Rev.,* vol. 140 pp. A1133-A1138, 1965.

[16] A. D. Becke, "Density-functional exchange-energy approximation with correct asymptotic behavior," *Phys. Rev. A,* vol. 38, pp. 3098-3100, 1988.

[17] A. D. Becke, "New mixing of Hartree-Fock and local density-functional theories," *J. Chem. Phys.,* vol. 98, pp. 1372-1377, 1993.

[18] J. P. Perdew and A. Zunger, "Self-interaction correction to density-functional approximations for many-electron systems," *Phys Rev B,* vol. 23, pp. 5048-79, 1981.
[19] C. Adamo and V. Barone, "Toward reliable density functional methods without adjustable parameters: The PBE0 model," *J. Chem. Phys.,* vol. 110, pp. 6158-70, 1999.
[20] J. Simons, *An Introduction to Theoretical Chemistry*. Cambridge: Cambridge University Press, 2003.
[21] C. Froese-Fischer, *The Hartree-Fock Method for Atoms*: John Wiley, 1977.
[22] A. Szabó and N. S. Ostlund, *Modern Quantum Chemistry: introduction to advanced electronic structure theory* McGraw-Hill, 1996
[23] P. Fulde, *Electron Correlations in Molecules and Solids*: Springer-Verlag, 1995.
[24] J. Cizek, "Stability conditions for the solutions of the Hartree-Fock equations for the simple open-shell case," *Adv. Chem. Phys.,* vol. 14 1969
[25] T. Torsti and J. E. T. Eirola, T. Hakala, P. Havu, V. Havu, T. Höynälänmaa, J. Ignatius, M. Lyly, I. Makkonen, T. T. Rantala, J. Ruokolainen, K. Ruotsalainen, E. Räsänen, H. Saarikoski, M. J. Puska, "Three real-space discretization techniques in electronic structure calculations," *Physica Status Solidi (B),* vol. 243, pp. 1016-1053, 2006.
[26] J. R. Flores*, et al.*, "Solution of atomic Hartree-Fock equations with the P version of the finite element method," *J. Chem. Phys.,* vol. 91, pp. 7030-7038, 1989.
[27] J. R. Flores*, et al.*, "A p-version finite-element approach for atomic Hartree-Fock calculations," *Chem. Phys. Lett.,* vol. 163, pp. 198-201, 1989.
[28] D. Heinemann*, et al.*, "Solution of the Hartree-Fock-Slater equations for diatomic molecules by the finite-element method," *Phys. Rev. A,* vol. 38, pp. 4994-5001, 1988.
[29] D. Heinemann*, et al.*, "Accurate Hartree-Fock-Slater calculations on small diatomic molecules with the finite-element method," *Chem. Phys. Lett.,* vol. 145, pp. 125-127, 1988.
[30] D. Heinemann*, et al.*, "Solution of the Hartree-Fock equations for atoms and diatomic molecules with the finite element method," *Phys. Scr.,* vol. 42, pp. 692-696, 1990.
[31] D. Heinemann*, et al.*, "Spin-polarized Hartree-Fock-Slater calculations in atoms and diatomic molecules with the finite element method," *Chem. Phys. Lett.,* vol. 166, pp. 627-629, 1990.
[32] C. Lavor*, et al.*, "Solving Hartree-Fock systems with global optimization methods," *Europhys. Lett.,* vol. 77, pp. 26-30, 2007.
[33] D. Sundholm, "Finite-element multiconfiguration Hartree-Fock calculations of the atomic quadrupole moment of Ar+(P-2(3/2))," *Phys. Rev. A,* vol. 59, pp. 3355-3358, Jan 1 1999.
[34] D. Sundholm and J. Olsen, "Finite-element multiconfiguration Hartree-Fock calculations on carbon, oxygen, and neon - the nuclear-quadrupole moments of C-11, O-17, and NE-21," *J. Phys. Chem.,* vol. 96, pp. 627-630, Jan 1 1992.

[35] D. Sundholm and J. Olsen, "Finite-element multiconfiguration Hartree-Fock determination of the atomic quadrupole-moment of CA (3D4S 1D)," *Chem. Phys. Lett.,* vol. 198, pp. 526-530, Jan 1 1992.
[36] D. Sundholm and J. Olsen, "Finite-element multiconfiguration Hartree-Fock calculations of the atomic quadrupole-moments of excited-states of Beryllium, Al, Indium, Ne, Ar, Kr, and Xe," *Phys. Rev. A,* vol. 47, pp. 2672-2679, Jan 1 1993.
[37] D. Sundholm and J. Olsen, "Finite-element multiconfiguration Hartree-Fock determination of the nuclear-quadrupole moments of chlorine, potassium, and calcium isotopes," *J. Chem. Phys.,* vol. 98, pp. 7152-7158, Jan 1 1993.
[38] D. Sundholm and J. Olsen, "Finite-element multiconfiguration Hartree-Fock calculations of electron-affinities of manganese," *Chem. Phys. Lett.,* vol. 233, pp. 115-122, Jan 1 1995.
[39] D. Sundholm*, et al.*, "Finite-element multiconfiguration Hartree-Fock calculations on the excitation-energies and the ionization-potential of oxygen," *J. Chem. Phys.,* vol. 96, pp. 5229-5232, Jan 1 1992.
[40] J. Ackermann*, et al.*, "A self-adaptive multilevel finite-element method for the stationary Schrodinger equation in 3-space dimensions," *J. Chem. Phys.,* vol. 101, pp. 7643-7650, Jan 1 1994.
[41] J. Ackermann and R. Roitzsch, "A 2-dimensional multilevel adaptive finite-element method for the time-independent Schrodinger equation," *Chem. Phys. Lett.,* vol. 214, pp. 109-117, Jan 1 1993.
[42] H. M. Quiney*, et al.*, "The dirac equation in the algebraic approximation: VIII. Comparison of finite basis set and finite element molecular Dirac-Hartree-Fock calculations for the H-2, LiH, and BH ground states," *Int. J. Quant. Chem.,* vol. 89, pp. 227-236, Jan 1 2002.
[43] S. R. White*, et al.*, "Finite-element method for electronic structure," *Phys. Rev. B,* vol. 39, pp. 5819-5833, Jan 1 1989.
[44] E. Tsuchida and M. Tsukada, "Electronic-structure calculations based on the finite-element method," *Phys. Rev. B,* vol. 52, pp. 5573-5578, Jan 1 1995.
[45] E. Tsuchida and M. Tsukada, "Real-space approach to electronic-structure calculations," *Solid State Comm.,* vol. 94, pp. 5-8, Jan 1 1995.
[46] E. Tsuchida and M. Tsukada, "Adaptive finite-element method for electronic-structure calculations," *Phys. Rev. B,* vol. 54, pp. 7602-7605, Jan 1 1996.
[47] E. Tsuchida and M. Tsukada, "Large-scale electronic-structure calculations based on the adaptive finite-element method," *J. Phys. Soc. Jpn,* vol. 67, pp. 3844-3858, Jan 1 1998.
[48] D. B. Cook, *Handbook of Computational Quantum Chemistry*: Oxford Univ. Press, 1998.
[49] G. t. Velde*, et al.*, "Chemistry with ADF," *J. Comp. Chem.,* vol. 22, pp. 931-967, 2001.

[50] J. R. Chelikowsky, *et al.*, "Finite-difference-pseudopotential method: electronic structure calculations without a basis," *Phys. Rev. Lett.,* vol. 72, pp. 1240-3, 1994.
[51] J. R. Chelikowsky, *et al.*, "Higher-order finite-difference pseudopotential method: an application to diatomic molecules," *Phys. Rev. B,* vol. 50, pp. 11355-64, 1994.
[52] E. L. Briggs, *et al.*, "Large-scale electronic-structure calculations with multigrid acceleration," *Phys. Rev. B,* vol. 52, pp. R5471-4, 1995.
[53] E. L. Briggs, *et al.*, "Real-space multigrid-based approach to large-scale electronic structure calculations," *Phys. Rev. B,* vol. 54, pp. 14362-75, 1996.
[54] G. Zumbach, *et al.*, "Adaptive coordinate, real-space electronic structure calculations on parallel computers," *Solid State Comm.,* vol. 99, pp. 57-61, 1996.
[55] N. A. Modine, *et al.*, "Adaptive-coordinate real-space electronic-structure calculations for atoms, molecules, and solids," *Phys. Rev. B,* vol. 55, pp. 10289-301, 1997.
[56] F. Gygi and G. Galli, "Real-space adaptive-coordinate electronic-structure calculations," *Phys. Rev. B,* vol. 52, pp. R2229-32, 1995.
[57] K. A. Iyer, *et al.*, "Application of a distributed nucleus approximation in grid based minimization of the Kohn-Sham energy functional," *J. Chem. Phys.,* vol. 103, pp. 227-33, 1995.
[58] T. L. Beck, "Real-space multigrid solution of electrostatics problems and the Kohn-Sham equations," *Int. J. Quant. Chem.,* vol. 65, pp. 477-86, 1997.
[59] T. L. Beck, *et al.*, "Multigrid methods in density functional theory," *Int. J. Quant. Chem.,* vol. 61, pp. 341-8, 1997.
[60] T. Hoshi, *et al.*, "Density-functional molecular dynamics with real-space finite difference," *Phys. Rev. B,* vol. 52, pp. R5459-62, 1995.
[61] J. Kobus, *et al.*, "A numerical Hartree-Fock program for diatomic molecules," *Comp. Phys. Comm.,* vol. 98, pp. 346-358, 1996.
[62] B. Hermansson and D. Yevick, "Finite-element approach to band-structure analysis," *Phys. Rev. B,* vol. 33, pp. 7241-2, 1986.
[63] K. Tagami, *et al.*, "First-principles study of vibrational spectra on dihydride-terminated Si(001)/H surfaces," *Surf. Sci.,* vol. 446, pp. L108-12, 2000.
[64] T. A. Arias, "Multiresolution analysis of electronic structure: semicardinal and wavelet bases," *Rev. Mod. Phys.,* vol. 71, pp. 267-311, 1999.
[65] S. Goedecker and O. Ivanov, "Linear scaling solution of the Coulomb problem using wavelets," *Solid State Comm.,* vol. 105, pp. 665-9, 1998.
[66] T. Yanai, *et al.*, "Multiresolution quantum chemistry in multiwavelet bases: Hartree-Fock exchange," *J. Chem. Phys.,* vol. 121, pp. 6680-6688, 2004.
[67] E. J. Bylaska, *et al.*, "Adaptive Finite Element Method for Solving the Exact Kohn-Sham Equation of Density Functional Theory," *J. Chem. Theo. Comp.,* vol. 5, pp. 937-948, 2009.
[68] K. Ho-Le, "Finite element mesh generation methods: a review and classification," *Comput. Aided Des.,* vol. 20, pp. 27-38, 1988.

[69] P. L. George, *Automatic Mesh Generation: Application to Finite Element Methods*: John Wiley and Sons, 1991.
[70] O. C. Zienkiewicz, *The Finite Element Method*: Butterworth-Heinemann, 2000.
[71] J. C. Slater, "Statistical exchange-correlation in the self- consistent field," *Adv. Quantum Chem.,* vol. 6, pp. 1-92, 1972.
[72] J. H. V. Lenthe and P. Pulay, "A space-saving modification of Davidson's eigenvector algorithm," *J. Comp. Chem.,* vol. 11, pp. 1164-1168 1990.
[73] H. A. van der Vorst, *Iterative Krylov Methods for Large Linear Systems*. Cambridge: Cambridge University Press, 2003.
[74] A. George and J. Liu, *Computer Solution of Large Sparse Positive Definite Systems*: Prentice-Hall, 1981.
[75] P. Pulay, "Convergence acceleration of iterative sequences. The case of SCF iteration," *Chem. Phys. Lett.,* vol. 73, pp. 393-398, 1980.
[76] G. E. Scuseria, "Linear Scaling Density Functional Calculations with Gaussian Orbitals," *J. Phys. Chem. A,* vol. 103, pp. 4782-4790, 1999.
[77] W. Yang and T. S. Lee, "A density-matrix divide-and-conquer approach for electronic structure calculations of large molecules," *J. Chem. Phys.,* vol. 103, pp. 5674-5678, 1995.
[78] F. L. Hirshfeld, "Bonded-atom fragments for describing molecular charge densities," *Theo. Chim. Acta,* vol. 44, pp. 129-138, 1977.
[79] P. O. Persson and G. Strang, "A simple mesh generator in MATLAB," *SIAM Rev.,* vol. 46, pp. 329-345, 2004.
[80] R. Alizadegan*, et al.*, "A divide and conquer real space finite-element Hartree–Fock method," *J. Chem. Phys.,* vol. 132, p. 034101, 2010.
[81] S. S. Terdalkar*, et al.*, "Molecular Dynamics Simulations of Ion-irradiation Induced Deflection of 2D Graphene Films," *Intl. J. Solids Struct.,* vol. 45, pp. 3908-3917, 2008.
[82] S. Huzinaga, "Concept of active electrons in chemistry," *Can. I. Chem.,* vol. 73, pp. 619-628 1995.
[83] P. L. Goodfriend and S. I. Tsonchev, "Molecular calculations using space-restricted basis functions," *Intl. J. Quant. Chem.,* vol. 48, pp. 367-373, 1993.
[84] P. R. Heyliger, "Boundary singularities in the Laplace equation by enriched finite elements," *Comp. & Struc.,* vol. 34, pp. 681-683, 1990.
[85] T. Nguyen-Thoi*, et al.*, "A face-based smoothed finite element method (FS-FEM) for 3D linear and geometrically non-linear solid mechanics problems using 4-node tetrahedral elements," *Int. J. Num. Meth. Eng.,* vol. 78, pp. 324-353, 2009.
[86] J. M. Melenk and I. Babuska, "The partition of unity finite element method: Basic theory and applications," *Comp. Meth. Appl. Mech. Eng.,* vol. 139, pp. 289-314, 1996.
[87] I. Babuska and J. M. Melenk, "The partition of unity method," *Int. J. Num. Meth. Eng.,* vol. 40, pp. 727-758, 1997.

[88] T. Belytschko, *et al.*, "GFEM A review of extended/generalized finite element methods for material modeling.," *Model. Sim. Mat. Sci. Eng.,* vol. 17, p. 043001, 2009.
[89] N. Sukumar and J. E. Pask, "Classical and enriched finite element formulations for Bloch-periodic boundary conditions," *Int. J. Num. Meth. Eng.,* vol. 77, pp. 1121-1138, 2009.
[90] J. S. Chen, Wu, C. T., Yoon, S. and You, Y., *Int. J. Num. Meth. Eng.,* vol. 50, pp. 435–466, 2000.
[91] G. R. Liu, *et al.*, "A smoothed finite element method for mechanics problems (SFEM)," *Comput. Mech.,* vol. 39, pp. 859-877, 2007.
[92] G. R. Liu, *et al.*, "Theoretical aspects of the smoothed finite element method (SFEM)," *Int. J. Num. Meth. Eng.,* vol. 71, pp. 902-930, 2007.
[93] G. R. Liu, *et al.*, "A node-based smoothed finite element method (NS-FEM) for upper bound solutions to solid mechanics problems," *Comput. Struct.,* vol. 87, pp. 14-26, 2009.
[94] L. Chen, *et al.*, "Assessment of smoothed point interpolation methods for elastic mechanics," *Commun. Numer. Meth. Eng.,* vol. 26, pp. 1635-1655, 2009.
[95] P. Pulay, "Improved SCF convergence acceleration," *J. Comp. Chem.,* vol. 3, pp. 556-560, 1982.
[96] M. F. Crommie, *et al.*, "Confinement of electrons to quantum corrals on a metal surface," *Science* vol. 262 pp. 218-20, 1993.
[97] T. W. Tombler, *et al.*, "Reversible electromechanical characteristics of carbon nanotubes under local-probe manipulation," *Nature,* vol. 405, pp. 769-772, 2000.
[98] D. R. Hartree, *Proc. Camb. Phil. Soc.,* vol. 24 1927
[99] V. A. Fock, *Z. Phys. ,* vol. 62 1930
[100] J. C. Slater and G. F. Koster, "Simplified LCAO method for the periodic potential problem," *Phys. Rev. ,* vol. 94 pp. 1498-1524, 1954.
[101] C. H. Xu, *et al.*, "A transferable tight-binding potential for carbon," *J. Phys.: Condens. Matter* vol. 4 pp. 6047-6054, 1992
[102] K. Esfarjani and Y. Kawazoe, "Self-consistent tight-binding formalism for charged systems," *J. Phys.: Condens. Matter,* vol. 10 pp. 8257-8267, 1998.
[103] S. Datta, *Quantum Transport: Atom to Transistor*: Camb. Univ. Press, 2005.
[104] S. N. Yaliraki, *et al.*, "The injecting energy at molecule/metal interfaces: Implications for conductance of molecular junctions from an ab initio molecular description," *J. Chem. Phys.,* vol. 111, p. 6997, 1999.
[105] B. Liu, *et al.*, "The influence of mechanical deformation on the electrical properties of single wall carbon nanotubes," *J. Mech. Phys. Solids,* vol. 52, pp. 1-26, 2004.
[106] H. T. Johnson, *et al.*, "Electron transport in deformed carbon nanotubes," *J. Eng. Mater. Tech.,* vol. 126, pp. 222-9, 2004.
[107] A. Rochefort, *et al.*, "Effects of finite length on the electronic structure of carbon nanotubes," *J. Phys. Chem. B,* vol. 103, pp. 641-6, 1999.

[108] S. N. Yaliraki, *et al.*, "Conductance of molecular wires: Influence of molecule-electrode binding," *J. Am. Chem. Soc.,* vol. 121, pp. 3428-3434, 1999.
[109] Y. He, *et al.*, "Effects of strain and defects on the electron conductance of metallic carbon nanotubes," *Phys. Rev. B,* vol. 75, p. 235429, 2007.
[110] K.-t. Lau, *et al.*, "Thermal and mechanical properties of single-walled carbon nanotube bundle-reinforced epoxy nanocomposites: the role of solvent for nanotube dispersion," *Composites Science and Technology,* vol. 65, pp. 719-725, 2005.
[111] J. H. Rouse, "Polymer-Assisted Dispersion of Single-Walled Carbon Nanotubes in Alcohols and Applicability toward Carbon Nanotube/Sol-Gel Composite Formation," *Langmuir,* vol. 21, pp. 1055-1061, 2005.
[112] Y. Sabba and E. L. Thomas, "High-Concentration Dispersion of Single-Wall Carbon Nanotubes," *Macromolecules,* vol. 37, pp. 4815-4820, 2004.
[113] J. Sandlera, *et al.*, "Development of a dispersion process for carbon nanotubes in an epoxy matrix and the resulting electrical properties," *Polymer,* vol. 40, pp. 5967-5971, 1999.
[114] M. Sennett, *et al.*, "Dispersion and alignment of carbon nanotubes in polycarbonate," *Appl. Phys. A,* vol. 76, pp. 111-113, 2003.
[115] A. N. Kolmogorov and V. H. Crespi, "Smoothest Bearings: Interlayer Sliding in Multiwalled Carbon Nanotubes," *Phys. Rev. Lett.,* vol. 85, p. 4727, 2000.
[116] S. Zhang, *et al.*, "Atomistic simulations of double-walled carbon nanotubes (DWCNTs) as rotational bearings," *Nano Lett.,* vol. 4, pp. 293-297, 2004.
[117] K. G. Kornev, *et al.*, "Ribbon-to-Fiber Transformation in the Process of Spinning of Carbon-Nanotube Dispersion," *Phys. Rev. Lett.,* vol. 97, p. 188303, 2006.
[118] S. Iijima, "Helical Microtubules of Graphitic Carbon," *Nature,* vol. 354, pp. 56-58, Nov 7 1991.
[119] B. I. Yakobson and L. S. Couchman, "Carbon Nanotubes: Supramolecular Mechanics," *Dekker Encyclopedia of Nanoscience and Nanotechnology,* p. 587, 2004.
[120] C.-H. Sun, *et al.*, "Van der Waals interactions between two parallel infinitely long single-walled nanotubes," *Chemical Physics Letters,* vol. 403, pp. 343-346, 2005.
[121] J. P. Lu, "Elastic Properties of Carbon Nanotubes and Nanoropes," *Phys. Rev. Lett.,* vol. 79, p. 1297, 1997.
[122] J. Tersoff and R. S. Ruoff, "Structural Properties of a Carbon-Nanotube Crystal," *Phys. Rev. Lett.,* vol. 73, p. 676, 1994.
[123] X. Ma, *et al.*, "Tribological Behavior of Aligned Single-Walled Carbon Nanotubes," *Transact. ASME,* vol. 126, p. 258, 2004.
[124] L. X. Benedict, *et al.*, "Microscopic determination of the interlayer binding energy in graphite," *Chem. Phys. Lett.,* vol. 286, p. 490, 1998.
[125] N. G. Chopra, *et al.*, "Fully Collapsed Carbon Nanotubes," *Nature,* vol. 377, pp. 135-138, Sep 14 1995.

[126] R. S. Ruoff, *et al.*, "Radial deformation of carbon nanotubes by van der Waals forces," *Nature,* vol. 364, pp. 514-516, 1993.
[127] T. Hertel, *et al.*, "Deformation of carbon nanotubes by surface van der Waals forces," *Physical Review B,* vol. 58, pp. 13870-13873, Nov 15 1998.
[128] L. A. Girifalco, *et al.*, "Carbon nanotubes, buckyballs, ropes, and a universal graphitic potential," *Phys. Rev. B,* vol. 62, p. 13104, 2000.
[129] B. Chen, *et al.*, "Binding energy of parallel carbon nanotubes," *Appl. Phys. Lett.,* vol. 83, pp. 3570-3571, Oct 27 2003.
[130] A. Liao, *et al.*, "Thermal dissipation and variability in electrical breakdown of carbon nanotube devices," *Phys. Rev. B,* vol. 82, p. 205406, 2010.
[131] K. Ryu, *et al.*, "CMOS-Analogous Wafer-Scale Nanotube-on-Insulator Approach for Submicrometer Devices and Integrated Circuits Using Aligned Nanotubes," *Nano Letters,* vol. 9, pp. 189-197, Jan 2009.
[132] M. A. Panzer, *et al.*, "Thermal properties of metal-coated vertically aligned single-wall nanotube arrays," *Journal of Heat Transfer,* vol. 130, p. 052401, May 2008.
[133] S. T. Huxtable, *et al.*, "Interfacial heat flow in carbon nanotube suspensions," *Nat Mater,* vol. 2, pp. 731-734, 2003.
[134] E. Pop, *et al.*, "Negative differential conductance and hot phonons in suspended nanotube molecular wires," *Phys. Rev. Lett.,* vol. 95, p. 155505, Oct 7 2005.
[135] M. Lazzeri, *et al.*, "Electron Transport and Hot Phonons in Carbon Nanotubes," *Phys. Rev. Lett.,* vol. 95, p. 236802, 2005.
[136] M. A. Kuroda, *et al.*, "Nonlinear transport and heat dissipation in metallic carbon nanotubes," *Phys. Rev. Lett.,* vol. 95, p. 266803, Dec 31 2005.
[137] P. G. Collins, *et al.*, "Current saturation and electrical breakdown in multiwalled carbon nanotubes," *Phys. Rev. Lett.,* vol. 86, pp. 3128-3131, Apr 2 2001.
[138] S. J. Kang, *et al.*, "High-performance electronics using dense, perfectly aligned arrays of single-walled carbon nanotubes," *Nat Nano,* vol. 2, pp. 230-236, 2007.
[139] N. Patil, *et al.*, "VMR: VLSI-compatible metallic carbon nanotube removal for imperfection-immune cascaded multi-stage digital logic circuits using Carbon Nanotube FETs," in *IEEE Intl. Electron Devices Mtg. (IEDM)*, Baltimore, MD, 2009, pp. 573-576.
[140] E. Pop, "The role of electrical and thermal contact resistance for Joule breakdown of single-wall carbon nanotubes," *Nanotechnology,* vol. 19, p. 295202, July 2008.
[141] E. Pop, *et al.*, "Electrical and thermal transport in metallic single-wall carbon nanotubes on insulating substrates," *J. Appl. Phys.,* vol. 101, p. 093710, May 2007.
[142] M. Steiner, *et al.*, "Phonon populations and electrical power dissipation in carbon nanotube transistors," *Nature Nanotechnology,* vol. 4, pp. 320-324, May 2009.
[143] J. Shiomi and S. Maruyama, "Molecular dynamics of diffusive-ballistic heat conduction in single-walled carbon nanotubes," *Japanese Journal of Applied Physics,* vol. 47, pp. 2005-2009, Apr 2008.

[144] R. Prasher, *et al.*, "An acoustic and dimensional mismatch model for thermal boundary conductance between a vertical mesoscopic nanowire/nanotube and a bulk substrate," *Journal of Applied Physics,* vol. 102, p. 104312, Nov 2007.
[145] H. Y. Chiu, *et al.*, "Ballistic phonon thermal transport in multiwalled carbon nanotubes," *Physical Review Letters,* vol. 95, p. 226101, Nov 25 2005.
[146] Z.-Y. Ong and E. Pop, "Molecular Dynamics Simulation of Thermal Boundary Conductance Between Carbon Nanotubes and SiO_2," *Phys. Rev. B,* vol. 81, p. 155408, 2010.
[147] A. Liao, *et al.*, "Avalanche-Induced Current Enhancement in Semiconducting Carbon Nanotubes," *Physical Review Letters,* vol. 101, p. 256804, Dec 19 2008.
[148] D. Estrada, *et al.*, "Reduction of Hysteresis for Carbon Nanotube Mobility Measurements Using Pulsed Characterization," *Nanotechnology,* vol. 21, p. 085702, 2010.
[149] J. Lee, *et al.*, "Electrical and Thermal Coupling to a Single-Wall Carbon Nanotube Device Using an Electrothermal Nanoprobe," *Nano Letters,* vol. 9, pp. 1356-1361, Apr 2009.
[150] Z. Yao, *et al.*, "High-field electrical transport in single-wall carbon nanotubes," *Physical Review Letters,* vol. 84, pp. 2941-2944, Mar 27 2000.
[151] Y. F. Chen and M. S. Fuhrer, "Electric-field-dependent charge-carrier velocity in semiconducting carbon nanotubes," *Physical Review Letters,* vol. 95, p. 236803, Dec 2 2005.
[152] P. K. Schelling and R. Keblinski, "Thermal expansion of carbon structures," *Physical Review B,* vol. 68, p. 035425, Jul 15 2003.
[153] Y. Zhao, *et al.*, "Multiband Mobility in Semiconducting Carbon Nanotubes," *IEEE Electron Device Letters,* vol. 30, pp. 1078-1080, Oct 2009.
[154] K. Hata, *et al.*, "Water-assisted highly efficient synthesis of impurity-free single-waited carbon nanotubes," *Science,* vol. 306, pp. 1362-1364, Nov 19 2004.
[155] N. Y. Huang, *et al.*, "Mechanism Responsible for Initiating Carbon Nanotube Vacuum Breakdown," *Physical Review Letters,* vol. 93, p. 075501, 2004.
[156] E. Pop, *et al.*, "Thermal conductance of an individual single-wall carbon nanotube above room temperature," *Nano Letters,* vol. 6, pp. 96-100, Jan 2006.
[157] I. K. Hsu, *et al.*, "Optical measurement of thermal transport in suspended carbon nanotubes," *Applied Physics Letters,* vol. 92, p. 063119, Feb 11 2008.
[158] V. V. Deshpande, *et al.*, "Spatially Resolved Temperature Measurements of Electrically Heated Carbon Nanotubes," *Physical Review Letters,* vol. 102, pp. 105501-4, 2009.
[159] Y. Ouyang and J. Guo, "Heat dissipation in carbon nanotube transistors," *Applied Physics Letters,* vol. 89, p. 183122, Oct 30 2006.
[160] F. Xiong, *et al.*, "Inducing chalcogenide phase change with ultra-narrow carbon nanotube heaters," *Applied Physics Letters,* vol. 95, p. 243103, 2009.
[161] L. Shi, *et al.*, "Thermal probing of energy dissipation in current-carrying carbon nanotubes," *J. Appl. Phys.,* vol. 105, p. 104306, May 15 2009.

[162] E. T. Swartz and R. O. Pohl, "Thermal boundary resistance," *Reviews of Modern Physics,* vol. 61, p. 605, 1989.
[163] R. Prasher, "Thermal boundary resistance and thermal conductivity of multiwalled carbon nanotubes," *Physical Review B,* vol. 77, p. 075424, Feb 2008.
[164] J. C. Duda*, et al.*, "Extension of the diffuse mismatch model for thermal boundary conductance between isotropic and anisotropic materials," *Appl. Phys. Lett.,* vol. 95, p. 031912, Jul 20 2009.
[165] P. E. Hopkins, "Multiple phonon processes contributing to inelastic scattering during thermal boundary conductance at solid interfaces," *Journal of Applied Physics,* vol. 106, p. 013528, Jul 1 2009.
[166] P. Reddy*, et al.*, "Diffuse mismatch model of thermal boundary conductance using exact phonon dispersion," *Applied Physics Letters,* vol. 87, p. 211908, Nov 21 2005.
[167] G. D. Mahan and G. S. Jeon, "Flexure modes in carbon nanotubes," *Physical Review B,* vol. 70, p. 075405, Aug 2004.
[168] P. G. Sverdrup*, et al.*, "Sub-continuum simulations of heat conduction in silicon-on-insulator transistors," *Journal of Heat Transfer,* vol. 123, pp. 130-137, Feb 2001.
[169] K. Sun*, et al.*, "Graphite C-axis thermal conductivity," *Superlattices and Microstructures,* vol. 45, pp. 60-64, Feb 2009.
[170] F. P. Incropera and D. P. DeWitt, *Fundamentals of Heat and Mass Transfer*, 5th ed. New York: Wiley, 2001.
[171] D. Mann*, et al.*, "Thermally and Molecularly Stimulated Relaxation of Hot Phonons in Suspended Carbon Nanotubes," *Journal of Physical Chemistry B,* vol. 110, pp. 1502-1505, 2006.
[172] S. Zhang*, et al.*, "Transition states and minimum energy pathways for the collapse of carbon nanotubes," *Physical Review B,* vol. 73, p. 075423, 2006.
[173] A. K. Rappe*, et al.*, "Uff, a Full Periodic-Table Force-Field for Molecular Mechanics and Molecular-Dynamics Simulations," *Journal of the American Chemical Society,* vol. 114, pp. 10024-10035, Dec 2 1992.
[174] H. Maune*, et al.*, "Thermal resistance of the nanoscale constrictions between carbon nanotubes and solid substrates," *Applied Physics Letters,* vol. 89, p. 013109, Jul 3 2006.
[175] Z. Chen*, et al.*, "Thermal contact resistance between graphene and silicon dioxide," *Appl. Phys. Lett.,* vol. 95, p. 161910 Oct 19 2009.
[176] A. G. Petrov and S. V. Rotkin, "Energy relaxation of hot carriers in single-wall carbon nanotubes by surface optical phonons of the substrate," *JETP Letters,* vol. 84, pp. 156-160, Oct 2006.
[177] S. V. Rotkin*, et al.*, "An Essential Mechanism of Heat Dissipation in Carbon Nanotube Electronics," *Nano Letters,* vol. 9, pp. 1850-1855, May 2009.
[178] T. Shimizu and et al., "Electrical conductivity measurements of a multi-walled carbon nanotube," *Surf. Interface Anal.,* vol. 37, pp. 204-207, 2005.

[179] R. Martel, *et al.*, "Single- and multi-wall carbon nanotube field-effect transistors," *Appl. Phys. Lett.,* vol. 73, p. 2447, 1998.
[180] P. Avouris, *et al.*, "Carbon nanotubes: nanomechanics, manipulation, and electronic devices," *Appl. Surface Sci.,* vol. 141 pp. 201-209, 1999.
[181] L. Durrer, *et al.*, "SWNT growth by CVD on Ferritin-based iron catalyst nanoparticles towards CNT sensors," *Sensors and Actuators B (Chemical),* vol. 132, pp. 485-90, 2008.
[182] H. Kataura, *et al.*, "Optical properties of single-wall carbon nanotubes," *Synthetic Metals,* vol. 103, pp. 2555-2558, Jun 1999.
[183] W. Kim, *et al.*, "Electrical contacts to carbon nanotubes down to 1 nm in diameter," *Applied Physics Letters,* vol. 87, p. 173101, Oct 24 2005.
[184] T. Hertel, *et al.*, "Manipulation of Individual Carbon Nanotubes and Their Interaction with Surfaces," *J. Phys. Chem. B,* vol. 102, pp. 910-915, 1998.
[185] M. R. Falvo, *et al.*, "Gearlike rolling motion mediated by commensurate contact: Carbon nanotubes on HOPG," *Phys. Rev. B,* vol. 62, p. R10665, 2000.
[186] K. Miura, *et al.*, "Natural Rolling of Zigzag Multiwalled Carbon Nanotubes on Graphite," *Nano Lett.,* vol. 1, pp. 161-163, 2001.
[187] A. Bulduma and J. P. Lu, "Modeling and simulations of carbon nanotubes and their junctions on surfaces," *Appl. Surface Sci.,* vol. 219 pp. 123-128, 2003.
[188] J. Wu and et al., "Computational design of carbon nanotube electromechanical pressure sensors," *Phys. Rev. B,* vol. 69, p. 153406, 2004.
[189] A. Pantano, *et al.*, "Mechanics of deformation of single-and multi-wall carbon nanotubes," *J. Mech. Phys. Solids,* vol. 52, pp. 789 – 821, 2004.
[190] J. L. Hutter and J. Bechhoefer, "Measurement and manipulation of van der Waals forces in atomic-force microscopy," *J. Vac. Sci. Technol. B,* vol. 12, pp. 2251-3, 1994.
[191] R. P. Jaiswal, *et al.*, "Modeling and validation of the van der Waals force during the adhesion of nanoscale objects to rough surfaces: A detailed description," *Langmuir,* vol. 25, pp. 10612-10623, 2009.
[192] J. L. Hutter and J. Bechhoefer, "Manipulation of van der Waals forces to improve image resolution in atomic-force microscopy," *J. Appl. Phys.,* vol. 73, pp. 4123-9, 1993.
[193] C. Horie and H. Miyazaki, "Atomic-force-microscopy images of graphite due to van der Waals interactions," *Phys. Rev. B,* vol. 42, pp. 11757-61, 1990.
[194] X. Huang, *et al.*, "Coordinated Buckling of Thick Multi-Walled Carbon Nanotubes Under Uniaxial Compression," *Nano Research,* vol. 3, p. 32, 2010.
[195] C. Odin, *et al.*, "Tip's finite size effects on atomic force microscopy in the contact mode: simple geometrical considerations for rapid estimation of apex radius and tip angle based on the study of polystyrene latex balls," *Surface Science,* vol. 317 pp. 321-340, 1994.
[196] T. DeBorde, *et al.*, "Identifying Individual Single-Walled and Double-Walled Carbon Nanotubes by Atomic Force Microscopy," *Nano Lett.,* vol. 8, pp. 3568-3571, 2008.

[197] J. L. Xiao, *et al.*, "Alignment Controlled Growth of Single-Walled Carbon Nanotubes on Quartz Substrates," *Nano Letters,* vol. 9, pp. 4311-4319, Dec 2009.
[198] M. Minary-Jolandan and M.-F. Yu, "Experiment and model analysis of significant elastic radial deformation of carbon nanotubes in nanoindentation," *J. Appl. Phys.,* vol. 103, p. 073516, 2008.
[199] A. M. Rao, *et al.*, "Diameter-selective Raman scattering from vibrational modes in carbon nanotubes," *Science,* vol. 275, pp. 187-191, Jan 10 1997.
[200] A. Jorio, *et al.*, "Structural (n, m) determination of isolated single-wall carbon nanotubes by resonant Raman scattering," *Physical Review Letters,* vol. 86, pp. 1118-1121, Feb 5 2001.
[201] A. Jorio, *et al.*, "G-band resonant Raman study of 62 isolated single-wall carbon nanotubes," *Physical Review B,* vol. 65, p. 155412, Apr 15 2002.
[202] W. Lu, *et al.*, "A Scanning Probe Microscopy Based Assay for Single-Walled Carbon Nanotube Metallicity," *Nano Letters,* vol. 9, pp. 1668-1672, Apr 2009.
[203] S. B. Cronin, *et al.*, "Measuring the uniaxial strain of individual single-wall carbon nanotubes: Resonance Raman spectra of atomic-force-microscope modified single-wall nanotubes," *Physical Review Letters,* vol. 93, p. 167401, Oct 15 2004.
[204] M. Ishigami, *et al.*, "Atomic Structure of Graphene on SiO2," *Nano Letters,* vol. 7, pp. 1643-1648, 2007.
[205] M. P. Rossi, *et al.*, "Deformation of Carbon Nanotubes by Exposure to Water Vapor," *Langmuir,* vol. 25, pp. 2804-2810, 2009.
[206] G. E. Begtrup, *et al.*, "Probing nanoscale solids at thermal extremes," *Physical Review Letters,* vol. 99, p. 155901, Oct 12 2007.
[207] J. W. G. Wildoer, *et al.*, "Electronic structure of atomically resolved carbon nanotubes," *Nature,* vol. 391, pp. 59-62, Jan 1 1998.
[208] L. Tapaszto, *et al.*, "Apparent diameter of carbon nanotubes in scanning tunnelling microscopy measurements," *Journal of Physics-Condensed Matter,* vol. 18, pp. 5793-5805, Jul 5 2006.
[209] S. M. Bachilo, *et al.*, "Structure-assigned optical spectra of single-walled carbon nanotubes," *Science,* vol. 298, pp. 2361-2366, Dec 20 2002.
[210] R. B. Weisman and S. M. Bachilo, "Dependence of optical transition energies on structure for single-walled carbon nanotubes in aqueous suspension: An empirical Kataura plot," *Nano Letters,* vol. 3, pp. 1235-1238, Sep 2003.
[211] M. Y. Sfeir, *et al.*, "Probing electronic transitions in individual carbon nanotubes by Rayleigh scattering," *Science,* vol. 306, pp. 1540-1543, Nov 26 2004.
[212] H. Farhat, *et al.*, "Phonon softening in individual metallic carbon nanotubes due to the Kohn anomaly," *Physical Review Letters,* vol. 99, p. 145506, Oct 5 2007.
[213] M. F. Yu, *et al.*, "Investigation of the radial deformability of individual carbon nanotubes under controlled indentation force," *Phys. Rev. Lett.,* vol. 85, pp. 1456-9, 2000.

[214] M. Huhtala, *et al.*, "Carbon nanotube structures: molecular dynamics simulation at realistic limit," *Comput. Phys. Commun.,* vol. 146 pp. 30-7, 2002.
[215] B. I. Yakobson, *et al.*, "Nanomechanics of carbon tubes: instabilities beyond linear response," *Phys. Rev. Lett.,* vol. 76, pp. 2511-14, 1996.
[216] M. R. Falvo, *et al.*, "Bending and buckling of carbon nanotubes under large strain," *Nature,* vol. 389 pp. 582-4, 1997.
[217] W. Shen, *et al.*, "Investigation of the radial compression of carbon nanotubes with a scanning probe microscope," *Phys. Rev. Lett.,* vol. 84 pp. 3634-7, 2000.
[218] V. Lordi and N. Yao, "Radial compression and controlled cutting of carbon nanotubes," *J. Chem. Phys.,* vol. 109, pp. 2509-12, 1998.
[219] M. F. Yu, *et al.*, "Structural analysis of collapsed, and twisted and collapsed, multiwalled carbon nanotubes by atomic force microscopy," *Phys. Rev. Lett.,* vol. 86, pp. 87-90, 2001.
[220] J. A. Elliott, *et al.*, "Collapse of single-wall carbon nanotubes is diameter dependent," *Phys. Rev. Lett.,* vol. 92, pp. 095501-1, 2004.
[221] T. Tang, *et al.*, "Collapse of single-walled carbon nanotubes," *J. Appl. Phys.,* vol. 97, p. 074310, 2005.
[222] G. Gao, *et al.*, "Energetics, structure, mechanical and vibrational properties of single-walled carbon nanotubes," *Nanotechnology,* vol. 9, pp. 184-91, 1998.
[223] J. Yu, *et al.*, "Phonons in graphitic tubules: a tight-binding molecular dynamics study," *J. Chem. Phys.,* vol. 103, pp. 6697-705, 1995.
[224] V. N. Popov, *et al.*, "Elastic properties of single-walled carbon nanotubes," *Phys. Rev. B,* vol. 61, pp. 3078-84, 2000.
[225] J. Yoon, *et al.*, "Sound wave propagation in multiwall carbon nanotubes," *J. Appl. Phys.,* vol. 93, pp. 4801-6, 2003.
[226] B. Liu, *et al.*, "The role of lattice registry on the full collapse and twist of carbon nanotubes," *Phys. Rev. B,* vol. 70, p. 161402(R), 2004.
[227] J. Xiao, *et al.*, "Collapse and stability of single- and multi-wall carbon nanotubes," *Nanotechnology,* vol. 18 p. 395703 2007.
[228] C. L. Cheung, *et al.*, "Diameter-controlled synthesis of carbon nanotubes," *J. Phys. Chem. B,* vol. 106, pp. 2429-33, 2002.
[229] C. J. Park, *et al.*, "Band-gap modification by radial deformation in carbon nanotubes," *Phys. Rev. B,* vol. 60 p. 10656, 1999.
[230] J. Q. Lu, *et al.*, "Metal-to-semiconductor transition in squashed armchair carbon nanotubes," *Phys. Rev. Lett.,* vol. 90, pp. 156601-1, 2003.
[231] J. Wu, *et al.*, "Computational design of carbon nanotube electromechanical pressure sensors," *Phys. Rev. B,* vol. 69, p. 153406 2004.
[232] J. Q. Lu, *et al.*, "Structural trends interpretation of the metal-to-semiconductor transition in deformed carbon nanotubes," *J. Appl. Phys.,* vol. 97, p. 114314 2005.
[233] P. E. Lammert, *et al.*, "Gapping by squashing: metal-insulator and insulator-metal transitions in collapsed carbon nanotubes," *Phys. Rev. Lett.,* vol. 84 pp. 2453-6, 2000.

[234] J. R. Gilbert, *et al.*, "Sparse Matrices in MATLAB: Design and Implementation," *SIAM J. Matrix Anal. App.,* vol. 13, pp. 333-356, 1992.
[235] *http://www.mathworks.com/matlabcentral/fileexchange/23488*.

AUTHOR'S BIOGRAPHY

Mr. Alizadegan was born in September, 1979 in downtown Tehran, the capital of Iran. He received his high school diploma in physics and mathematics discipline from the Tehran Allameh-Helli National Organization for Development of Exceptional Talents (NODET) in 1997. He received his BS degree in Civil Engineering from Sharif University of Technology in 2001 with the highest honor. He was the first student in the Iranian national MS entrance scientific competition among more than 11,000 Civil Engineering undergraduate students across the country. Subsequently, he received an MS degree in Structural Mechanics at the same university in 2003 with honors. His MS thesis topic was "Computational modeling of elasto-plastic behavior in large deformations using Element-Free Galerkin Method" which infused him with a great enthusiasm in Computational Mechanics and sparked his interest in pursuing a second MS degree in Theoretical and Applied Mechanics (TAM) at the University of Illinois at Urbana-Champaign (UIUC) which he decided to continue with the Ph.D. degree in the same major. However, during this time period, he switched his interests from engineering to science and he has been working on Computational Nanoscience/Nanotechnology projects for the most of his thesis work. He has had several publications with the collaboration of other groups/people. He has been a teaching assistant (TA) for several courses during his graduate work including Heat Transfer, Introductory (and Advanced) Solid Mechanics, Mechanics of Polymers/Composite Materials, and Advanced Engineering Mathematics. He will be joining a software development company in Manhattan, New York City upon the completion of his Ph.D. work.