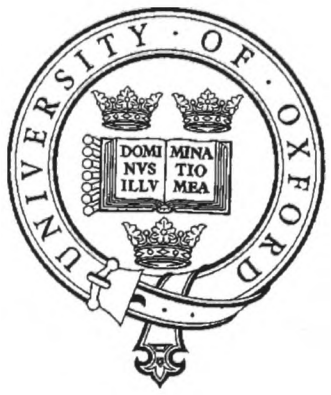
Tight-binding studies of carbon nanotubes

Kirsty Skidmore

Linacre College, Oxford



A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy of the University of Oxford

Michaelmas Term, 2002

Acknowledgements

I would like to thank my supervisor Dr. David Manolopoulos for all his help over the past three (and a bit) years. He has never failed to make useful suggestions and has always encouraged me to think for myself.

The EPRSC provided the financial support for this project.

I am grateful to the past and present members of the Manolopoulos group: Dimitris, Tomás, Adam, Haydn, Ed, Ian, Aya, Olly, Charles, Carrie and Rob. All of them have made my time here more enjoyable than it would otherwise have been, but Adam, Ed and Tomás deserve a special mention. My project arose out of Adam's thesis and he was kind enough to explain his work to me. I am indebted to Tomás and Ed for proof-reading large parts of this thesis and putting up with so much for so long.

I have appreciated the support of my friends in Oxford and elsewhere during this project. Thank you!

Lastly, I would like to thank my parents and my sister Helen who have supported me throughout my education.

Contents

1 Carbon nanotubes **1**

1.1 Discovery of fullerenes and nanotubes 1

1.2 Observation and production of nanotubes 4

1.3 Potential nanotube applications 6

1.4 Studying nanotube caps 12

1.4.1 Motivation 12

1.4.2 Structure of the rest of this thesis 13

2 Tight-binding theory **16**

2.1 General tight-binding theory 18

2.1.1 Band structure energy 18

2.1.2 Repulsive energy 23

2.1.3 Forces 24

2.1.4 Tight-binding theory applications 26

2.1.5 Transferability 26

2.1.6 Tight-binding theory summary 28

2.2 Tight-binding schemes for carbon 29

2.2.1 Ho Hamiltonian 29

2.2.2 Porezag Hamiltonian 35
2.2.3 Ho versus Porezag 39
2.3 Summary 41

3 Limiting cases: Fullerenes and infinite nanotubes 42
3.1 Fullerenes 43
3.1.1 Cataloguing the fullerene isomers 44
3.1.2 Predicting the ground state isomer 45
3.1.3 Full valence tight-binding calculations 47
3.1.4 Fullerene summary 52
3.2 Infinite nanotubes 52
3.2.1 Hamada indices 53
3.2.2 Exploiting the helical and rotational symmetries 57
3.2.3 Electronic structure of infinite nanotubes 60
3.2.4 Energy per atom 66
3.2.5 Nanotube band gaps 71
3.3 Summary 73

4 Density matrix methods 74
4.1 Density matrix 75
4.1.1 Notation 75
4.1.2 Idempotency 77
4.1.3 Commutation with $\boldsymbol{H}$ 78
4.1.4 Localisation in real space 79
4.2 The LNV method 85
4.2.1 Grand canonical potential 85

4.2.2 Finding the density matrix 86
4.2.3 Minimisation 89
4.2.4 Cut-off distance 89
4.2.5 Forces 90
4.2.6 LNV conclusions 91
4.3 The PM method 91
4.3.1 Using the McWeeny transformation iteratively 91
4.3.2 Initial guess for the density matrix 92
4.3.3 Termination criterion 93
4.3.4 PM conclusions 95
4.4 The hybrid method 96
4.4.1 Connections between the LNV and PM methods 98
4.4.2 Hybrid conclusions 100
4.5 Demonstration of linear scaling 101
4.6 Applicability of density matrix methods 104
4.7 Applications of density matrix methods 106
4.8 Summary 107

5 Nanotube caps **109**
5.1 Motivation 110
5.1.1 Relevant phenomena 111
5.1.2 Previous studies 113
5.2 A census of nanotube caps 115
5.3 Extrapolating the energy of a cap 119
5.4 Results 128
5.4.1 (5,5) tube 128

5.4.2 (10,0) tube 130
5.4.3 (11,2) tube 135
5.5 Relevance of this survey to experiment 141
5.6 Density of states 145
5.6.1 Infinite nanotubes 148
5.6.2 Capped nanotubes 151
5.7 Summary 174

6 Conclusions **177**

A Hybrid method **180**
A.1 Displacements from the idempotency surface 180
A.2 PM path 183
A.3 LNV gradient 183

Bibliography **185**

Tight-binding studies of carbon nanotubes

Kirsty A. Skidmore, Linacre College

A thesis submitted in partial fulfilment of the requirements

for the degree of Doctor of Philosophy of the University of Oxford

Michaelmas Term, 2002

Abstract

Carbon nanotubes have been subject to a great many theoretical and experimental investigations and have many interesting properties. However, the caps found at the ends of nanotubes have been rather neglected. Previous work has established the possible caps for a given nanotube. This thesis seeks to build on this work to determine which caps are probable. Three representative nanotubes are considered: the (5,5) tube (an example of a metal); the (10,0) tube (an example of an insulator) and the (11,2) tube (for which experimental data is available).

A linear scaling density matrix method based on orthogonal tight-binding theory is used for a systematic study of doubly capped nanotubes. The energetically most stable caps are found for all three tubes and the isolated pentagon rule established for fullerenes is shown to be valid for nanotube caps. No simple rule governing the stability of isolated pentagon caps is found, although the stability of (11,2) caps correlates well with the number of hexagons adjacent to just one pentagon. The local densities of states (LDOS) are calculated for the most stable cap for each nanotube. A localised state is found for the capped (5,5) tube and a resonant state is observed for the capped (11,2) tube. The LDOS for the capped (11,2) tube is compared with experimental observations and questions are raised concerning the nature of the tube observed and the limitations of STM as a method for identifying nanotube caps.

Chapter 1

Carbon nanotubes

1.1 Discovery of fullerenes and nanotubes

In 1985, an analysis of the soot produced in a carbon arc discharge experiment identified a previously unknown allotrope of carbon: C_{60}.[1] Unlike diamond or graphite, C_{60} is a molecule, not an extended solid, although it does form a molecular solid with a fcc crystal structure.[2] C_{60} has a cage-like structure with the atoms arranged in hexagons and pentagons. In this structure, each carbon atom is connected to three others, as in graphite, but, owing to the presence of the pentagons, the molecule is spherical not planar. In fact, C_{60} has the same shape as a football and when Kroto, Heath, O'Brien, Curl and Smalley published the paper announcing the discovery of C_{60} it featured a photograph of a football.[1] This structure of C_{60} was later confirmed by the experimental determination of both the IR[3] and ^{13}C NMR[4] spectra of C_{60}.

At first, C_{60} was regarded as something of a curiosity, but it quickly became

apparent that C_{60} was only one in a whole family of molecules known as the fullerenes, all with similar cage-like structures. Further, it was realised that the fullerenes have some interesting properties, for example, the superconductivity of alkali metal doped solid C_{60} compounds such as K_3C_{60} and Rb_3C_{60}[2] and the perhaps surprising addition reactions of C_{60}.[5] It is also possible to form unusual metallofullerene compounds involving metal atoms within fullerene cages.[6,7] Indeed, in recognition of the importance of the discovery of C_{60}, the Nobel prize in chemistry 1996 was awarded to Kroto, Curl and Smalley.

So, the fullerenes themselves attracted much interest, but their discovery also led, in 1991, to the characterisation of carbon nanotubes by Iijima.[8] Nanotubes are similar to fullerenes; they too are composed of three-coordinate carbon atoms. The atoms are arranged in hexagons as for graphite except that the hexagons are on the surface of a cylinder. Multi-walled nanotubes consisting of a series of co-axial tubes with increasing diameter were discovered first,[8] however, single-walled nanotubes were soon detected.[9,10] The TEM (transmission electron microscope) image in Figure 1.1 (from Iijima's original paper[8]) shows three different multi-walled nanotubes of varying lengths. Figure 1.2 shows a TEM image of single-walled nanotubes (from a paper by Bonard *et al.*[11]). Comparing the length scales in both images, it is immediately apparent that single-walled nanotubes are narrower than multi-walled nanotubes.

Much of the theoretical work has focussed on single-walled infinite nanotubes (even though multi-walled nanotubes are easier to study experimentally).

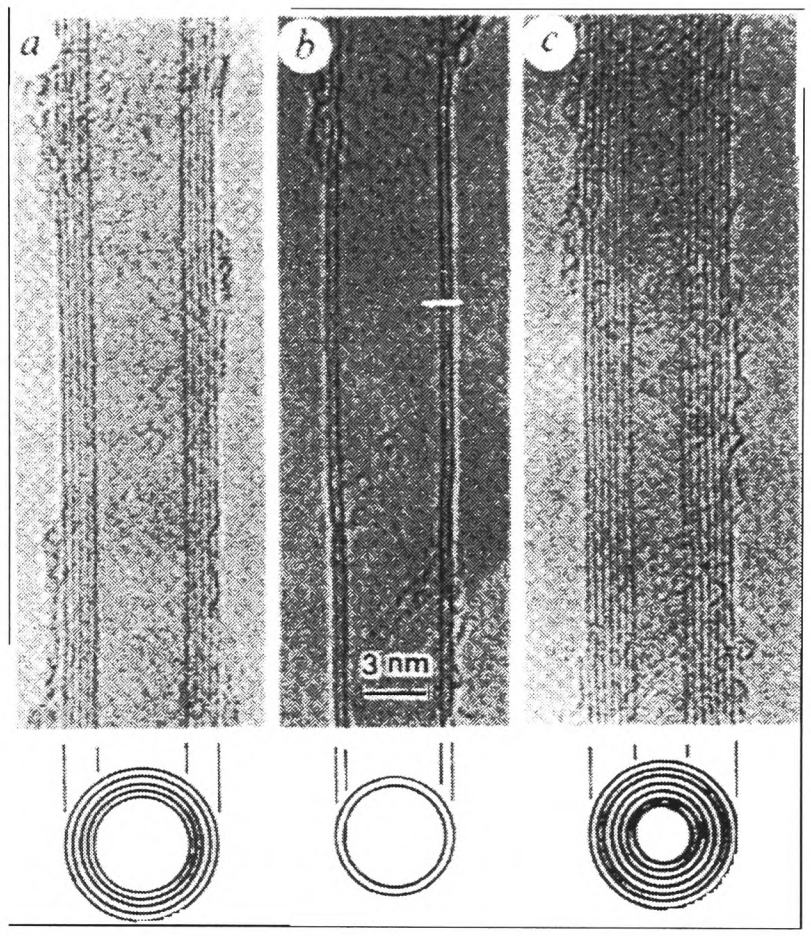


Figure 1.1: TEM images of multi-walled nanotubes taken from Iijima's original paper.[8]

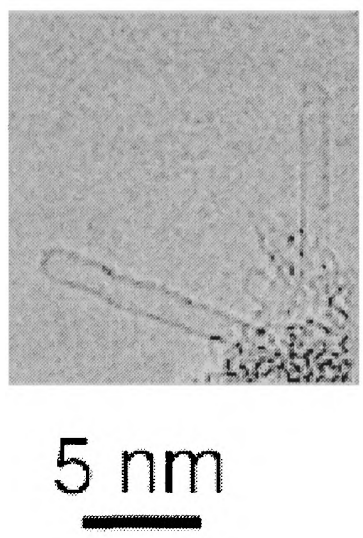


Figure 1.2: TEM image of single-walled nanotubes taken from a paper by Bonard *et al.*[11]

Single-walled nanotubes have been of particular interest to theoreticians for two (pragmatic) reasons. First, the dispersion interactions between the layers of multi-walled tubes[12] are not trivial to model successfully and second, the number of atoms in the unit cell of a single-walled nanotube is much less than in the case of multi-walled tubes. In this thesis, only single-walled nanotubes will be considered.

1.2 Observation and production of nanotubes

Nanotubes are observed experimentally using scanning tunnelling spectroscopies (STS) or electron microscopy, either high resolution transmission electron microscopy (HRTEM) or TEM. Electron microscopy was the first

technique used to explore nanotubes.[8–10] The atomic resolution of nanotubes using STS has only become possible recently.[13,14] However, STS can not only be used to image the tubes, but also to provide information about the electronic structure of nanotubes[13–17] and it is therefore an incredibly powerful technique for examining nanotubes.

There are three different methods[8–10,18,19] for making nanotubes experimentally: arc discharge, laser vaporisation and chemical vapour deposition (CVD). The method of choice depends on the reason the nanotubes are being manufactured, since the tubes produced vary in quality, cost and homogeneity according to the means of production.[20] Each of these methods can be fine-tuned to optimise production of particular nanotubes. The growth mechanism of nanotubes is an area of active research, but it is likely that it depends on the production method.[20] A better understanding of the growth mechanism is desirable as it would aid the design of synthesis methods for industrial scale production of nanotubes.[19]

The generic aspects of the three nanotube production methods are briefly explained below.

Arc discharge

An electric discharge between two graphitic electrodes results in nanotubes forming at the anode. This was the original production method for both multi- and single-walled tubes.[8–10] In order to produce single-walled nanotubes a metal catalyst is vaporised at the same time as the carbon. Originally iron[9] or cobalt[10] were used, but many other metals, or mixtures of

metals, may be used. The diameter of the tubes produced and the yield vary with the metal(s) chosen.[9,10,21]

Laser vaporisation

The laser vaporisation method is similar to the arc discharge method. As the name suggests, lasers are used to vaporise the graphite with cobalt or nickel or some other metal or mixtures of metals. The metal acts as a catalyst. This method is costly, owing to the use of lasers, but it is easier to control than the arc discharge method and produces more uniform nanotubes.[18]

CVD

This method involves passing small carbon containing gaseous molecules (such as carbon monoxide[22] or small hydrocarbons such as methane or acetylene[23]) with metal particles over a surface which has been patterned with a catalyst. The length and position of the tubes can be controlled by the location of the catalytic islands. This method is particularly cheap and might be especially useful in future nanotube applications where the precise position of the nanotubes is crucial.[23]

1.3 Potential nanotube applications

As a new field, nanotubes have naturally attracted a great deal of attention, both experimental and theoretical. As a result, it has been found that nanotubes have some extremely interesting properties. For example, most

nanotubes are semiconductors with varying bandgaps, but some are metallic. Nanotubes are very strong and are also good field emitters. In view of these interesting properties of nanotubes, it is no surprise that many applications have been suggested.

Nanotube research is indeed a very popular field: a search of the scientific abstracts in the ISI Web of Science database revealed that over 1700 papers on carbon nanotubes were published in 2002 alone. The following discussion of potential nanotube applications is by no means complete, but a comprehensive review is probably impossible and certainly outside the remit of this thesis.

Energy storage

In the search for cleaner fuels, hydrogen has often been suggested as a solution since its combustion produces water rather than harmful greenhouse gases. The main problem with using hydrogen is how to store it. Carbon nanotubes have been proposed as candidates for hydrogen storage, however, the results are still controversial and it is unclear as yet whether or not the uptake (and release) of hydrogen by nanotubes is great enough to justify attempting to build a hydrogen storage device using nanotubes.[24]

Alternatively, it may be that nanotubes can contribute to energy provision via batteries. Open multi-walled nanotubes can reversibly intercalate lithium ions which means they have potential as anodes for lithium-ion batteries.[25]

Composite materials

Adding nanotubes to polymers results in an electrically conducting composite. This is actually the major use of commercially manufactured nanotubes at the present time. Multi-walled nanotubes are commonly used to make conducting plastics for components in cars which are painted electrostatically.[24,26] The conducting plastic enables a uniform finish to be obtained. Multi-walled nanotubes are used owing to their lower cost. In fact, this application does not even require the nanotubes to be purified since it is sufficient that some metallic nanotubes are present. Continuing with this theme, it is hoped that adding nanotubes to plastics might provide stronger materials for construction.[24,26]

Scanning probe microscopies

Scanning probe microscopies are some of the most powerful methods for examining nanotubes. However, nanotubes themselves have been used as probes in scanning probe microscopies.[27] Scanning probe microscopies rely on a tunnelling current between the probe and the sample.[28] Both multi-walled[29] and single-walled[30] nanotubes can be used in this way. The nanotubes are either attached after they have been synthesised, or more recently, synthesised directly in place.[29,30]

Nanotubes make excellent probes because their cylindrical nature enables them to examine deeper features much more satisfactorily than more traditional probes. The small diameter of particularly single-walled nanotubes

means that the lateral resolution is very good. If a force which is greater than a certain critical value is applied to a nanotube, the nanotube will buckle. This limits the force which is applied to the sample and thus can prevent injury to fragile samples and also means that the probe itself is less likely to be damaged. In particular, the use of single-walled nanotubes has allowed DNA to be examined by AFM, something that was previously impossible because DNA is so easily damaged.[31]

This is one of the more important applications of nanotubes, at least in scientific terms. It is also one of the few applications that has actually been fully realised at the present time.

Field emission based devices

When a large electric field is applied to a nanotube, electrons from the tube are emitted into the vacuum by tunnelling.[21] This effect is known as field emission and has spawned intense interest and speculation about the possible uses of nanotubes. Nanotubes are particularly well suited for use in field emission devices because they are also chemically stable, strong and are very narrow compared to their width, maximising their field emission (the field, which determines field emission, is very sensitive to the local shape of the emitter).[21,22]

Field emitting nanotubes have a range of different applications in lamps,[21] cathode ray tubes,[21] flat screen displays[32] and x-ray generators.[24] These applications typically involve an assembly of nanotubes, rather than a single nanotube.[22]

Field emission displays are particularly desirable because they offer wide viewing angles, fast response times, operate under a wide range of temperatures and have low power consumption.[32] The electronics company Samsung have produced a nine inch flat screen display made of nanotubes which is capable of showing moving colour images.[32]

Lamps could use nanotubes to produce electrons to interact with a phosphorus coated surface.[11,21,22] Additionally, field emission lamps could be more environmentally sound than more traditional lamps.[21] This is both because their power consumption is low as there is no need to heat the nanotubes to induce the emission of electrons and because no toxic elements are required in the device.

Electronics

Since some nanotubes are metallic while others are semiconductors, the potential use of nanotubes in electronics is obvious. As nanotubes are extremely small, this would open the results for electronic circuits many times smaller than the current silicon chips. As the size of the silicon chips is rapidly becoming the limiting factor in speeding up computers, this application has excited much interest.[23,24,33]

Realistically, the electronic circuits would require the growth of nanotubes in situ (*i.e.* by CVD),[23] something that still needs to be developed. Assuming this can be achieved, formidable obstacles (such as how nanotubes could be used as the various components of an electrical circuit) remain. Defects in nanotubes can change the helicity of the nanotube, which affects whether

the tube is metallic or not. So, a semiconducting tube could be joined to a metallic tube through such a defect which would be useful for nanoscale electronic devices.[33,34] Tubes of this sort have been examined experimentally, but these tubes have only ever been made accidentally, rather than by design.[34]

Filling nanotubes

Both multi-walled[35] and single-walled[36] nanotubes may be filled with inorganic compounds. The filled single-walled nanotubes in particular have generated some interest regarding the mechanism of filling[37] and whether or not nanotubes could be used as a template for nanoscale fabrication.[35]

Gas sensors

The resistivity of a bundle of single-walled nanotube can change in the presence of certain gases. In fact, experiments have shown that not only does the resistivity change for N_2, H_2, O_2, He and NH_3, but that different types of temperature-dependent behaviour are expected when different gases are adsorbed on a nanotube bundle.[38] Indeed, it is believed that while N_2, H_2 and He are physisorbed, O_2 and NH_3 are chemisorbed. It is possible that these responses could be used as the basis for detectors of these gases.[38]

1.4 Studying nanotube caps

1.4.1 Motivation

In the early days, nanotube research focussed on the bulk properties of the tube and the ends of the tube caused little concern. However, nanotubes are usually terminated by a cap. A key issue, therefore, is whether a nanotube cap has any influence on the properties of the nanotube. This is of particular concern in studies of properties such as field emission, or on the use of nanotubes as STS probes, where the end of the tube is known to be important. Previous work has established the possible caps for a given nanotube.[39] Unfortunately, this is not the end of the discussion because there are many possible caps for a given nanotube. Different caps may affect the tube properties in different ways and it is thus imperative to determine which of the many possible caps for a tube are likely, before one draws any conclusions regarding the importance of the nanotube cap in influencing a property of a tube. However, most studies of carbon nanotubes, even those which consider the caps themselves, neglect to determine which cap is likely to be found experimentally.

This thesis seeks to address this issue by examining three representative nanotubes within the experimental range of diameters and determining the likely caps in each case, including one for which some experimental data is available. The tubes investigated are the (5,5) tube (an example of a metal); the (10,0) tube (an example of an insulator) and the (11,2) tube (for which experimental results are available). It is hoped that this study and

accompanying methodology will stimulate further research in this area.

1.4.2 Structure of the rest of this thesis

The experimentally observed isomer(s) of a given fullerene $C_{N_{atom}}$ are invariably found to be the most stable isomers.[40] In order to determine which isomer is the most stable one, the energy of each possible isomer must be calculated and compared to every other one for that fullerene. As nanotube caps have a similar structure to fullerenes, it is therefore plausible to suppose (in the absence of any other information) that the cap which is most likely to be observed experimentally for a given nanotube will be the most stable cap for that particular tube. This is the starting point for this thesis, although we shall return to this assumption in Chapter 5.

Having made this assumption, the first issue to be addressed is how to calculate this cap energy. A tight-binding scheme developed by Ho *et al.*[41] was chosen for this work. Chapter 2 reviews tight-binding theory, discusses the Ho Hamiltonian and the reasons why this particular parameterisation was chosen.

In order to justify the choice of the Ho Hamiltonian further, it is necessary to demonstrate good agreement between the results obtained with the Ho Hamiltonian and well established theoretical, or better experimental, results from the literature for molecules related to capped carbon nanotubes. As one reduces the length of a capped carbon nanotube the molecule becomes more and more similar to a fullerene. By the same argument, increasing the length of a capped carbon nanotube will result in something that increasingly

resembles an infinite nanotube. In the light of this, it is logical to expect that a Hamiltonian which correctly reproduces the properties of fullerenes and infinite nanotubes will correctly predict the energies of capped nanotubes. Therefore, Chapter 3 applies the Ho Hamiltonian to these two limiting cases and discusses their properties.

Any realistic calculation of the energy of a capped nanotube must consider a molecule with a large number of atoms since the lengths of experimentally observed nanotubes are generally of the order of at least a few hundred nanometres. Furthermore, the presence of the nanotube cap destroys the periodic nature of a nanotube which may be exploited to calculate the energy of an infinite tube very efficiently. Conventional tight-binding studies involve diagonalising an $(N \times N)$ matrix, a process which scales as N^3, where N is the number of orbitals. This places severe constraints on the number of atoms that may be considered. However, over the past ten years various linear scaling methods have been proposed which sidestep this problem. Chapter 4 discusses three such methods based on the density matrix.

The major work of this thesis is presented in Chapter 5. A method for extrapolating the energy of a nanotube cap (rather than the energy of a capped nanotube) is presented and used to calculate the cap energies from the results of the linear scaling tight-binding calculations on capped tubes. The most stable caps for the (5,5), (10,0) and (11,2) tubes are determined. Since density of states and local density of states for capped carbon nanotubes are readily obtainable from STM studies, these are calculated for the most stable caps for each tube studied. The general features of these density of states are

discussed, and in the case of the (11,2) tube, compared with experimental data available in the literature.

Finally, Chapter 6 summarises this work and the conclusions arising from it.

Chapter 2

Tight-binding theory

Ab initio methods are the most rigorous way to calculate the energy of a molecule or the energy per atom of a periodic system. However, *ab initio* studies can only be used for small molecules or simple periodic systems because these methods are extremely computationally intensive. Therefore, it is often necessary to use more approximate methods.

Some of the simplest methods are empirical force fields which do not even construct the Hamiltonian matrix of the molecule. Instead these models give a formula for the energy of the molecule (periodic system) as a function of the atomic coordinates. Models of this sort have been parameterised for a number of elements or combination of elements including silicon,[42,43] carbon,[44] hydrocarbons[45] and mixtures of silicon, carbon and germanium.[46] These methods are appealing because of their very modest computational requirements.

These simple empirical force fields have been used to examine the structure

of, for example, liquid silicon and various allotropes of solid silicon[42,43] including surfaces, and solid, liquid and amorphous carbon[44,46,47] and also in molecular dynamics simulations of, for example, C_{60} colliding with a diamond surface.[48] More importantly, as far as this thesis is concerned, carbon nanotubes have been studied using two of these empirical force fields[44,45] and the results are in qualitative agreement with more accurate methods.[49] Nevertheless, there are some properties of nanotubes which require knowledge of the electronic structure and which cannot be investigated by this type of method. Properties in this category include the band structures, band gaps and densities of states of infinite nanotubes. Since all of these have been probed experimentally to some degree, it is desirable to use a method capable of examining them. Moreover, in spite of the successes of empirical force fields, it has been found that they cannot successfully predict the ground state isomer of a fullerene.[50,51] It follows that they are unlikely to accurately predict the most stable cap(s) for a nanotube. Therefore this type of model is not a suitable candidate for the study undertaken in this work (although there have been some very recent developments which may improve the accuracy of these models[52]).

Tight-binding theory is a semi-empirical method for calculating the electronic structure of molecules or periodic systems. It is an orbital-based approach where the Hamiltonian matrix is constructed and then diagonalised to yield the energy of the molecule. Since tight-binding theory is orbital based, it is capable of examining the properties we are interested in, while not being as computationally demanding as *ab initio* methods.

This chapter starts by discussing tight-binding theory for molecules (periodic systems will not be discussed in this thesis, except for the particular case of infinite nanotubes which are dealt with in Chapter 3). We then compare two tight-binding parameterisations for carbon: the Ho Hamiltonian[41] and the Porezag Hamiltonian.[53] It is argued that the Ho Hamiltonian is more suitable for use in this thesis (Chapter 3 further justifies this choice by applying the Ho Hamiltonian to fullerenes and infinite nanotubes and comparing the results with experimental observations and previous theoretical studies). The chapter ends with a summary.

2.1 General tight-binding theory

Within tight-binding theory the total energy E_{tot} of a molecule is divided into two parts:

$$E_{\text{tot}} = E_{\text{bs}} + E_{\text{rep}}, \tag{2.1}$$

where E_{bs} is the band structure energy, which is the sum over all occupied electronic eigenvalues, and E_{rep} is a short-ranged repulsive energy. I shall discuss these separately.

2.1.1 Band structure energy

Tight-binding theory assumes a basis set of atomic-like orbitals. These are orbitals which share the symmetry properties of the atomic orbitals (and for this reason it is convenient to label them in the same way as atomic orbitals), but which may differ in other properties. For example, the basis orbitals are

often chosen to be orthogonal to one another,[54–56] but this is not generally true for the atomic orbitals in a molecule. Usually only valence orbitals are considered, so for carbon atoms only the $2s$ and $2p$ orbitals are included. These atomic orbitals are (in general) not explicitly constructed, but the wavefunction for a particular (molecular orbital) energy level is expressed as a linear combination of these N atomic-like orbitals,

$$|\Psi_n\rangle = \sum_{j=1}^{N} |j\rangle\, c_{jn}. \tag{2.2}$$

The band energy of the molecule can then be found by solving the Schrödinger equation:

$$\hat{H}\,|\Psi_n\rangle = \epsilon_n\,|\Psi_n\rangle\,, \tag{2.3}$$

with

$$E_{\mathrm{bs}} = 2\sum_{n=1}^{N_{\mathrm{occ}}} \epsilon_n, \tag{2.4}$$

where N_{occ} is the number of (doubly) occupied orbitals. Substituting Eqn. 2.2 in Eqn. 2.3 allows one to derive the secular equations as follows,

$$\begin{aligned}
\hat{H}\sum_{j=1}^{N} |j\rangle\, c_{jn} &= \epsilon_n \sum_{k=1}^{N} |k\rangle\, c_{kn}, \\
\sum_{j=1}^{N} \left(\left\langle i\left|\hat{H}\right|j\right\rangle - \epsilon_n \langle i|j\rangle\right) c_{jn} &= 0, \\
\sum_{j=1}^{N} \left(H_{ij} - \epsilon_n S_{ij}\right) c_{jn} &= 0,
\end{aligned} \tag{2.5}$$

where $H_{ij} = \left\langle i\left|\hat{H}\right|j\right\rangle$ and $S_{ij} = \langle i|j\rangle$ are the matrix elements of the Hamiltonian and overlap matrices respectively. The secular equations given above apply for non-orthogonal tight-binding. Often, orthogonal tight-binding theory is used where it is assumed that the elements of the overlap matrix are given by $S_{ij} = \delta_{ij}$. I shall discuss the validity of this assumption later.

The problem with setting up the secular equations exactly is that for even a moderately sized molecule there are many, many matrix elements to evaluate. However, all that enters the secular equations are the matrix elements H_{ij} and S_{ij} and not the orbitals themselves. Therefore, explicit construction of orbitals can be avoided by parameterising the matrix elements, and this is the essential idea of tight-binding theory.

In order to parameterise the Hamiltonian and overlap matrix elements, a number of approximations are usually made. The first of these is the so-called two-centred approximation,[55] which was suggested as long ago as 1954 in the original paper on tight-binding by Slater and Koster.[54] One of the problems of evaluating $\left\langle i \left| \hat{H} \right| j \right\rangle$ is that it depends not just on the orbitals i and j, and the atoms they are located on, but on other atoms since the Hamiltonian operator for the molecule contains a potential term made up of contributions from all the atoms in the molecule. Tight-binding theory replaces this part of the Hamiltonian with a potential that is dependent only on the atoms α and β that orbitals i and j are located on. The justification for this is that the three-centred interactions ought to be smaller in magnitude than the two-centred integrals.[54,55] Whilst this is true, it does not follow that the three-centred integrals are negligible.[55]

Having made this approximation, we now have matrix elements H_{ij} that only depend on the symmetry of the orbitals i and j and the interatomic distance $R_{\alpha\beta}$. This still leaves us with a large number of matrix elements to evaluate, but we can make two further simplifications. First, we assume that the matrix elements of the Hamiltonian matrix (and the overlap matrix

in non-orthogonal tight-binding theory) are zero beyond a certain cut-off distance. This is reasonable in that we are dealing with atomic-like orbitals and one would expect the Hamiltonian and overlap matrix elements between atomic orbitals on different atoms to decrease to zero fairly rapidly as the internuclear separation increases. It should be remembered that this is an approximation. The second simplification (which is not an approximation) involves further reducing the number of matrix elements considered to a small number of "fundamental" matrix elements. We discuss this next.

Slater-Koster integrals

In 1954, Slater and Koster published their celebrated paper on tight-binding theory.[54] They realised that the Hamiltonian and overlap matrix elements are a function of the internuclear distance and orientation of the orbitals. Thus the Hamiltonian matrix element $H_{\alpha\beta}^{sx}$ between an s orbital on atom α and a p_x orbital on atom β has the same value as the Hamiltonian matrix element $H_{\gamma\delta}^{sx}$ between an s orbital on a similar atom γ and a p_x orbital on atom δ, if $R_{\alpha\beta} = R_{\gamma\delta}$ and the relative orientations of the orbitals are the same. Secondly, they observed that such matrix elements can be expressed in terms of certain combinations of a small number of integrals. These combinations can be deduced by the use of simple geometry. In the case of a full valence carbon Hamiltonian, there are only four Slater-Koster integrals, $H_{\alpha\beta}^{ss\sigma}$, $H_{\alpha\beta}^{sp\sigma}$, $H_{\alpha\beta}^{pp\sigma}$ and $H_{\alpha\beta}^{pp\pi}$. The expressions for the matrix elements between $2s$, $2p_x$, $2p_y$ and $2p_z$ orbitals on different atoms are given in terms of these integrals in Table 2.1. Analogous expressions exist for the overlap matrix elements.

Hamiltonian element	Slater-Koster Integrals
$H_{\alpha\beta}^{ss}$	$H_{\alpha\beta}^{ss\sigma}$
$H_{\alpha\beta}^{sx}$	$lH_{\alpha\beta}^{sp\sigma}$
$H_{\alpha\beta}^{sy}$	$mH_{\alpha\beta}^{sp\sigma}$
$H_{\alpha\beta}^{sz}$	$nH_{\alpha\beta}^{sp\sigma}$
$H_{\alpha\beta}^{xx}$	$l^2H_{\alpha\beta}^{pp\sigma} + (1-l^2)H_{\alpha\beta}^{pp\pi}$
$H_{\alpha\beta}^{yy}$	$m^2H_{\alpha\beta}^{pp\sigma} + (1-m^2)H_{\alpha\beta}^{pp\pi}$
$H_{\alpha\beta}^{zz}$	$n^2H_{\alpha\beta}^{pp\sigma} + (1-n^2)H_{\alpha\beta}^{pp\pi}$
$H_{\alpha\beta}^{xy}$	$lmH_{\alpha\beta}^{pp\sigma} - lmH_{\alpha\beta}^{pp\pi}$
$H_{\alpha\beta}^{xz}$	$lnH_{\alpha\beta}^{pp\sigma} - lnH_{\alpha\beta}^{pp\pi}$
$H_{\alpha\beta}^{yz}$	$mnH_{\alpha\beta}^{pp\sigma} - mnH_{\alpha\beta}^{pp\pi}$

Table 2.1: Slater-Koster integrals in terms of the direction cosines, $l = \left(\frac{x_\beta - x_\alpha}{R_{\alpha\beta}}\right)$, $m = \left(\frac{y_\beta - y_\alpha}{R_{\alpha\beta}}\right)$ and $z = \left(\frac{z_\beta - z_\alpha}{R_{\alpha\beta}}\right)$.

Within tight-binding theory the four fundamental matrix elements $H_{\alpha\beta}^{ss\sigma}$, $H_{\alpha\beta}^{sp\sigma}$, $H_{\alpha\beta}^{pp\sigma}$ and $H_{\alpha\beta}^{pp\pi}$ are viewed as functions of the internuclear separation $R_{\alpha\beta}$ and parameterised accordingly. Therefore, no integrals between basis orbitals need be evaluated explicitly, which is a great saving in terms of computational cost. However, this does leave the problem of what functional form to use for the fundamental matrix elements. This has never been resolved in terms of one definitive answer; different parameterisation schemes use different functional forms.

Once functional forms for the matrix elements have been chosen and parameterised, the Hamiltonian and overlap matrices (if appropriate) may be generated. The Hamiltonian can then diagonalised and hence E_{bs} obtained via Eqn. 2.4.

2.1.2 Repulsive energy

Since we have obtained E_{bs}, all that remains to be done in order to calculate E_{tot} from Eqn. 2.1 is to calculate E_{rep}. It is normally assumed that this repulsive energy can be written as a sum of pair terms,

$$E_{\mathrm{rep}} = \sum_{\alpha=1}^{N_{\mathrm{atom}}} \sum_{\beta=1}^{N_{\mathrm{atom}}} U_{\alpha\beta}(R_{\alpha\beta}). \tag{2.6}$$

This assumption, and indeed the underlying assumption made by tight-binding theory that the energy can be written as Eqn. 2.1, have been examined by both Sutton *et al.*[57] and Foulkes and Haydock.[58] This was motivated partly by a desire to relate tight-binding theory to density functional theory (DFT), but also by the fact that the form of Eqn. 2.1 and the use of

a pairwise sum to approximate E_{rep} had never been properly justified. Both these groups showed that it was indeed possible to write the total energy of a molecule as a sum of occupied eigenvalues of a Hamiltonian plus a sum of pair terms,[55,57,58] *i.e.* in the form I have described. The sum of occupied eigenvalues of the Hamiltonian corresponds to a sum of bond energies (the energy of each covalent bond) and promotion energies (the energy arising from the change in orbital occupancy when the molecule is formed from the free atoms). The sum of pair terms, the repulsive energy, consists of the changes in the total electrostatic energy and in the exchange and correlation energy when the atoms condense to form a solid from free atoms. The electrostatic contribution is a sum of pair terms whereas the exchange and correlation terms may be approximated by a sum of pair terms.[55,57,58]

2.1.3 Forces

Sometimes, for example in molecular dynamics simulations or geometry optimisations, the forces are needed as well as the energy. The forces are given by differentiating the total energy with respect to the atomic coordinate of interest λ,

$$\frac{\partial E_{\text{tot}}}{\partial \lambda} = \frac{\partial E_{\text{bs}}}{\partial \lambda} + \frac{\partial E_{\text{rep}}}{\partial \lambda}. \tag{2.7}$$

To a certain extent, the ease of calculating the forces depends on the functions chosen for the parameterisation of the Hamiltonian and overlap matrix elements and the repulsive energy, and whether they are easily differentiated. The functional form of the repulsive energy ($U_{\alpha\beta}(R_{\alpha\beta})$ in Eqn. 2.6) is usually chosen with this in mind. In order to find the forces corresponding to the

band energy one must use the Hellmann-Feynman theorem.[59,60] The forces corresponding to the band energy are given by

$$\frac{\partial E_{\mathrm{bs}}}{\partial \lambda} = 2 \sum_{n=1}^{N_{\mathrm{occ}}} \frac{\partial \epsilon_n}{\partial \lambda}, \tag{2.8}$$

where $\partial \epsilon_n / \partial \lambda$ can be obtained by differentiating the secular equations given in Eqn. 2.5:

$$\sum_{j=1}^{N} \left\{ \left(\frac{\partial H_{ij}}{\partial \lambda} - \frac{\partial \epsilon_n}{\partial \lambda} S_{ij} - \epsilon_n \frac{\partial S_{ij}}{\partial \lambda} \right) c_{jn} + \left(H_{ij} - \epsilon_n S_{ij} \right) \frac{\partial c_{jn}}{\partial \lambda} \right\} = 0. \tag{2.9}$$

Multiplying by c_{in}, summing over all possible values of i and rearranging yields:

$$\begin{aligned} \sum_{i=1}^{N} \sum_{j=1}^{N} c_{in} \left(\frac{\partial H_{ij}}{\partial \lambda} - \epsilon_n \frac{\partial S_{ij}}{\partial \lambda} \right) c_{jn} \quad - \quad \frac{\partial \epsilon_n}{\partial \lambda} \sum_{i=1}^{N} \sum_{j=1}^{N} c_{in} S_{ij} c_{jn} \\ + \sum_{j=1}^{N} \sum_{i=1}^{N} c_{in} \left(H_{ij} - \epsilon_n S_{ij} \right) \frac{\partial c_{jn}}{\partial \lambda} \quad = \quad 0. \end{aligned} \tag{2.10}$$

However, from Eqn. 2.2 it follows that $\sum_{i=1}^{N} \sum_{j=1}^{N} c_{in} S_{ij} c_{jn} = 1$, and from the secular equations $\sum_{i=1}^{N} c_{in} \left(H_{ij} - \epsilon_n S_{ij} \right) = 0$. Therefore, we are left with the Hellmann-Feynman expression for the forces:[59,60]

$$\frac{\partial \epsilon_n}{\partial \lambda} = \sum_{i=1}^{N} \sum_{j=1}^{N} c_{in} \left(\frac{\partial H_{ij}}{\partial \lambda} - \epsilon_n \frac{\partial S_{ij}}{\partial \lambda} \right) c_{jn}. \tag{2.11}$$

Thus it is possible to evaluate the forces by differentiating the matrix elements of the Hamiltonian and overlap matrices with respect to each atomic coordinate λ in turn. Since within tight-binding theory the Hamiltonian and overlap matrix elements are simply functions of the vector between the atoms with atomic orbitals i and j, this is comparatively straightforward.

2.1.4 Tight-binding theory applications

Tight-binding theory has been used extensively to study a wide range of different systems. The review by Goringe *et al.*[55] gives an excellent overview of this topic and, indeed, tight-binding theory in general. Tight-binding theory has been applied to both metals and semi-conductors. Topics that have been studied include extended solids, molecules, clusters, surfaces, defects, liquids and amorphous systems.

2.1.5 Transferability

A parameterised tight-binding Hamiltonian should yield results in good agreement with more exact methods (or experiments) for molecules or periodic systems which were used in the parameterisation process. Unfortunately, this does not necessarily mean the scheme will be applicable to other molecules or periodic systems. This lack of transferability has a number of sources.[55]

It is often assumed that the basis functions are orthogonal. This assumption is based on using Löwdin functions[55,56] which have the same symmetry properties as atomic orbitals and have been constructed from a basis set of non-orthogonal orbitals as follows:

$$|j\rangle = \sum_{i=1}^{N} \left(\boldsymbol{S}^{-\frac{1}{2}}\right)_{ji} |i) \,, \tag{2.12}$$

where $|i)$ are the original non-orthogonal orbitals, and $|j\rangle$ are the orthogonal orbitals. It is, of course, true that such a transformation exists (provided that $\boldsymbol{S}$ is not singular), yet there are two problems that arise from using this assumption in tight-binding theory. First, the orthogonalisation transformation

is only valid in the configuration in which it was carried out. Therefore, when one changes the position of the atoms, the transformation needed will be different, so in widely differing arrangements the basis functions are unlikely to be orthogonal. Secondly, the range of the orthogonal orbitals produced in this way is longer than that of the non-orthogonal atomic-like orbitals. The usual assumption that the basis function can be both orthogonal and short-ranged is incorrect. In general, it is only possible to have one or the other and not both.[54,55]

The use of the two-centred approximation, where three-centred integrals are neglected, can also result in a lack of transferability. Even though these three-centred integrals are certainly much less important than two-centred integrals, three-centred integrals are not negligible. In particular, three-centred integrals are rather sensitive to the structure which leads to problems with transferability.[55]

The problem of how to make a Hamiltonian transferable has dogged research into tight-binding methods.[55,61,62] One way to improve transferability is to use a more flexible functional form, *i.e.* to include more parameters and to fit to a wider range of data. This has been done; examples of this approach include the parameterisation for silicon by Goodwin *et al.*[62] and the Ho Hamiltonian for carbon.[41] The drawback of this type of approach is the fact that it is unsystematic and it is by no means obvious how it should be extended to other elements. Alternative approaches to improving transferability involve removing the assumptions (such as those listed above) which lead to this lack of transferability. The most important improvement is using

non-orthogonal orbitals.[63]

2.1.6 Tight-binding theory summary

It is clear, then, that tight-binding theory is a popular choice for electronic structure calculations involving large numbers of atoms. It has been widely applied to all sorts of different problems. Tight-binding theory is much less computationally demanding than *ab initio* methods and yet much more reliable than empirical force field methods as it retains an orbital description of the bonding. In addition, tight-binding theory has recently been implemented in various linear scaling methods which avoid the (direct) diagonalisation of the Hamiltonian matrix. Chapter 4 deals with linear scaling methods based on the density matrix.

However, there are still many challenges facing tight-binding theory. The issue of transferability is an on-going challenge, but some progress has been made in this area (fifteen years ago most potentials were not transferable[55,61,62]) Also, tight-binding theory parameterisation schemes are often not readily applicable to different elements and are often invented on a case by case basis. However, Porezag *et al.* have formulated a scheme (based on DFT) for systematically generating matrix elements for the Hamiltonian and overlap matrices, which should improve matters.[53]

Finally, it must be remembered that any tight-binding theory scheme is only as good as its parameterisation, including the functional forms used to fit the data.

2.2 Tight-binding schemes for carbon

The number of tight-binding studies of carbon systems has rapidly increased over the last fifteen years, something which can be attributed, at least in part, to the discoveries of C_{60} and other fullerenes and carbon nanotubes. There are at least two tried and tested tight-binding Hamiltonians for carbon in the literature: the orthogonal Ho Hamiltonian[41] and the non-orthogonal Porezag Hamiltonian.[53] For this work, I have chosen to use the Ho Hamiltonian for reasons which will be discussed shortly (see Section 2.2.3), and I will therefore describe this Hamiltonian in more detail; however, a review of the key features of the Porezag Hamiltonian will assist the later discussion.

2.2.1 Ho Hamiltonian

The Ho Hamiltonian[41] is an orthogonal tight-binding Hamiltonian for carbon. As usual in tight-binding theory, the total energy is expressed as in Eqn. 2.1,

$$E_{tot} = E_{bs} + E_{rep},$$

where E_{bs} is the band structure energy which is the sum of electronic eigenvalues over all occupied electronic states and E_{rep} is a short-ranged repulsive energy. The form of these functions shall be discussed presently.

The functional forms used in parameterising the Ho Hamiltonian were first suggested by Goodwin, Skinner and Pettifor[62] and were used by them to construct a tight-binding Hamiltonian for silicon. Goodwin *et al.* demonstrated that these functional forms resulted in a transferable tight-binding parameterisation for silicon, something that had been somewhat elusive previously.

Band structure parameterisation

The Ho Hamiltonian is a full valence tight-binding Hamiltonian which means that the basis functions are the $2s$ and $2p$ orbitals of carbon. In the Ho parameterisation scheme, the diagonal elements of the Hamiltonian matrix are the energy of the $2s$ orbital $\epsilon_s (= -2.99\,\text{eV})$ and the $2p$ orbital $\epsilon_p (= 3.71\,\text{eV})$ of a free carbon atom. Of course, these do not correspond to the actual energies of the $2s$ and $2p$ orbitals in carbon, they are merely parameters selected such that the cohesive energies of certain reference structures (including graphite and diamond) are fitted correctly. The off-diagonal Hamiltonian matrix elements are determined by a set of parameters $V_{ss\sigma}$ ($= -5$ eV), $V_{sp\sigma}$ ($= 4.7$ eV), $V_{pp\sigma}$ ($= 5.5$ eV) and $V_{pp\pi}$ ($= -1.55$ eV) which are multiplied by a scaling function $s\left(R_{\alpha\beta}\right)$ that depends on the internuclear separation $R_{\alpha\beta}$. The scaling function is given by

$$s\left(R_{\alpha\beta}\right) = \left(\frac{r_0}{R_{\alpha\beta}}\right)^n \exp\left\{n\left[-\left(\frac{R_{\alpha\beta}}{r_c}\right)^{n_c} + \left(\frac{r_0}{r_c}\right)^{n_c}\right]\right\}, \qquad (2.13)$$

where r_0 ($= 1.536329$ Å) is the nearest neighbour distance in diamond. The parameters n, n_c and r_c are given in Table 2.2. It is desirable not only that $s\left(R_{\alpha\beta}\right)$ should be zero at the cut-off distance r_m but also that it should be a smooth function for all $R_{\alpha\beta}$. To this end, Ho *et al.* replaced the tail of $s\left(R_{\alpha\beta}\right)$ with a third order polynomial $t_s(R_{\alpha\beta} - r_1)$ when $r_1 < R_{\alpha\beta} < r_m$. Since we know that $s\left(R_{\alpha\beta}\right) = t_s(R_{\alpha\beta} - r_1)$ when $R_{\alpha\beta} = r_1$ (for the resulting function to be smooth, we also require the first derivatives to be equal at $R_{\alpha\beta} = r_1$) and that $t_s(r_m - r_1)$ must be zero (as must its first derivative in order for the function to be smooth), it is possible to determine the coefficients for $t_s(R_{\alpha\beta} - r_1)$. The values for these coefficients are given in Table 2.3 and the

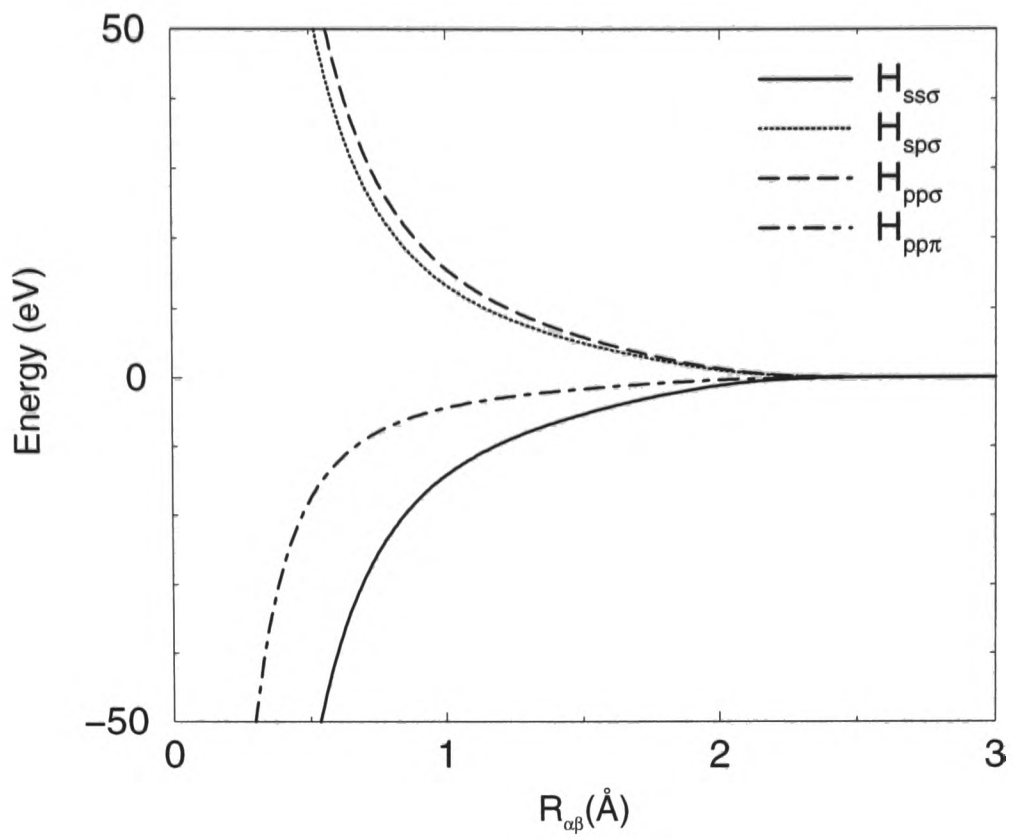


Figure 2.1: Ho Hamiltonian matrix elements.

values for r_1 and r_m are in Table 2.2. Fig. 2.1 shows the matrix elements $H_{\alpha\beta}^{ss\sigma}$, $H_{\alpha\beta}^{sp\sigma}$, $H_{\alpha\beta}^{pp\sigma}$ and $H_{\alpha\beta}^{pp\pi}$ as functions of the internuclear separation.

Form of the repulsive energy

Unusually for a tight-binding Hamiltonian, the repulsive energy for the Ho Hamiltonian is not given by a sum of pair terms, but rather a fourth order polynomial of a sum of pair terms:

$$E_{rep} = \sum_{\alpha=1}^{N_{\text{atom}}} f\left(\sum_{\beta=1}^{N_{\text{atom}}} \phi\left(R_{\alpha\beta}\right)\right), \tag{2.14}$$

where $\phi\left(R_{\alpha\beta}\right)$ is a pairwise potential between atoms α and β and f is a fourth order polynomial with argument $\sum_{\beta=1}^{N_{\text{atom}}} \phi\left(R_{\alpha\beta}\right)$. The pairwise potential

$s\,(R_{\alpha\beta})$ parameters		$\phi\,(R_{\alpha\beta})$ parameters	
r_1	2.45 Å	d_1	2.57 Å
n	2	ϕ_0	8.18555 eV
n_c	6.5	d_0	1.64 Å
r_c	2.18 Å	d_c	2.1052 Å
		m	3.30304
		m_c	8.6655

Table 2.2: Parameters for $s\,(R_{\alpha\beta})$ and $\phi\,(R_{\alpha\beta})$.[41]

	$t_s(R_{\alpha\beta} - r_1)$	$t_\phi(R_{\alpha\beta} - d_1)$	$f(x)$
c_0	$6.7392620074314 \times 10^{-3}$	$2.2504290109 \times 10^{-8}$	-2.5909765118191
c_1	$-8.1885359517898 \times 10^{-2}$	$-1.4408640561 \times 10^{-6}$	0.5721151498619
c_2	0.1932365259144	$2.1043303374 \times 10^{-5}$	$-1.7896349903996 \times 10^{-3}$
c_3	0.3542874332380	$6.6024390226 \times 10^{-5}$	$2.3539221516757 \times 10^{-5}$
c_4			$-1.24251169551587 \times 10^{-7}$

Table 2.3: Coefficients for the polynomials $t_s(R_{\alpha\beta} - r_1)$, $t_\phi(R_{\alpha\beta} - d_1)$ and $f(x)$.[41]

$\phi(R_{\alpha\beta})$ has the same form as $\phi(R_{\alpha\beta})$, albeit with different constants:

$$\phi(R_{\alpha\beta}) = \phi_0 \left(\frac{d_0}{R_{\alpha\beta}}\right)^m \exp\left\{m\left[-\left(\frac{R_{\alpha\beta}}{d_c}\right)^{m_c} + \left(\frac{d_0}{d_c}\right)^{m_c}\right]\right\}. \qquad (2.15)$$

where ϕ_0, m, d_c, d_0 and m_c are parameters determined by the fitting process; these are listed in Table 2.2.[41] The tail of the scaling function is replaced by the polynomial $t_\phi(R_{\alpha\beta} - d_1)$ when $d_1 < R_{\alpha\beta} < r_m$ as previously. The coefficients for $t_\phi(R_{\alpha\beta} - d_1)$ are given in Table 2.3 and the appropriate value for d_1 can be found in Table 2.2.[41]

Applications of the Ho Hamiltonian

The Ho Hamiltonian was parameterised by fitting to first principles calculations of the energy as a function of nearest-neighbour interatomic separation. These calculations used density-functional theory within the local-density approximation (LDA). A variety of different allotropes of carbon were used: diamond, graphite, linear chain, simple cubic and face-centred cubic structures. The diamond, graphite and linear chain structures were given particular attention. Figure 2.2 (taken from a paper by Ho *et al.*[41]) shows the energy per atom as a function of the nearest neighbour distance obtained from the Ho Hamiltonian and the LDA results. Unsurprisingly, the Ho Hamiltonian reproduces the graphite and diamond energy curves as a function of nearest neighbour internuclear separation particularly well and the curve for the linear chain structure is also satisfactory. The energy curves for higher coordinated structures are not so well predicted, but, the equilibrium cohesive energies of these structures are nevertheless acceptable.

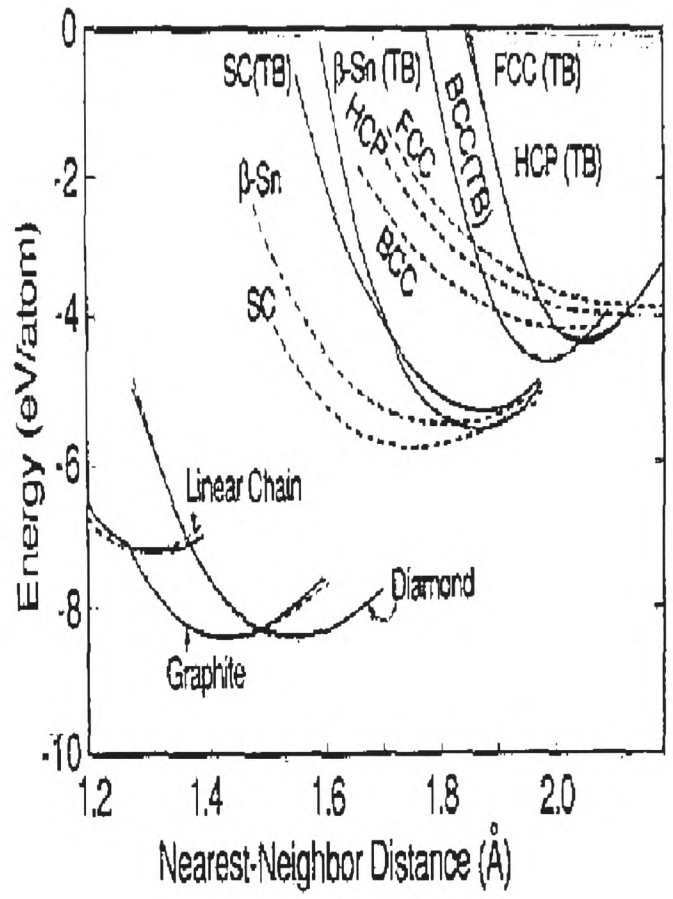


Figure 2.2: Energy curves against nearest neighbour internuclear separation for a variety of carbon allotropes, including the face-centred cubic (FCC), body-centred cubic (BCC), hexagonal close-packed (HCP), β-tin (β-Sn) and simple cubic (SC) structures. The LDA results are shown as unbroken lines, the Ho Hamiltonian results as dashed lines. This figure is taken from a paper by Ho *et al.*[41]

The Ho Hamiltonian has been widely applied to a large number of systems. It was originally tested on liquid carbon and small carbon clusters (for $N_{\text{atom}} < 11$, a Hubbard-like term to handle charge transfer is necessary),[41] and has since been applied to experimentally relevant fullerenes,[64–66] giant fullerenes,[67] crystalline and amorphous carbon.[68] The Ho Hamiltonian has been used for geometry optimisations,[67] simulated annealing[66] and tight-binding molecular dynamics simulations.[61,68] The next chapter will demonstrate the successful application of the Ho Hamiltonian to fullerenes and infinite carbon nanotubes.

2.2.2 Porezag Hamiltonian

A non-orthogonal tight-binding Hamiltonian for carbon was proposed by Porezag *et al.* in 1995.[53] This Hamiltonian is an example of what is sometimes called *ab initio* tight-binding. The *ab initio* calculations are used to calculate the matrix elements themselves, rather than to calculate cohesive energies which are then used to parameterise the matrix elements. The primary advantage of this is that there is a clear method for determining the matrix elements which should facilitate parameterisations for more than one element, as Porezag *et al.* demonstrated in the original paper by providing parameters for both carbon and hydrogen.

Band structure

The first stage in determining the Hamiltonian and overlap matrix elements is finding the orbitals using self-consistent DFT.[53] It is desirable to have or-

bitals which decay rapidly, so Porezag *et al.* artificially compress the atomic electronic density, by forcing the wavefunction to avoid regions at a great distance from the nucleus. Using these compressed orbitals is actually more accurate for solid state applications than using the atomic orbitals because the repulsion between the atoms in a solid results in contracted electron densities relative to free atoms.[53] The wavefunction for the molecule is written as a linear combination of these orbitals, as usual only the valence orbitals are included.

The overlap matrix elements can now be calculated easily. The diagonal elements of the Hamiltonian matrix are simply the eigenvalues of the free atoms. It is necessary to alter the potential part of the Hamiltonian in order to generate the off-diagonal elements of the Hamiltonian. The two-centred approximation is used to simplify the off-diagonal Hamiltonian matrix elements which are then given by

$$H_{ij} = \begin{cases} \epsilon_i & \text{if } i = j \\ \left\langle \phi_i^\alpha \left| \hat{H} \right| \phi_j^\beta \right\rangle & \text{if } \alpha \neq \beta \\ 0 & \text{otherwise} \end{cases} \tag{2.16}$$

where i labels the orbital and α the atom.

These expressions are used to tabulate the Slater-Koster matrix elements of the Hamiltonian and overlap matrices as a function of interatomic separation. The value of a matrix element for any interatomic separation can be interpolated from these values (which need be calculated only once and then stored). This is achieved by fitting the matrix elements to Chebyshev

polynomials,[69] for example,

$$H^{sp\sigma}(R_{\alpha\beta}) = \sum_{m=1}^{10} h_m^{sp\sigma} T_{m-1}(y) - \frac{h_1^{sp\sigma}}{2}, \tag{2.17}$$

and

$$S^{sp\sigma}(R_{\alpha\beta}) = \sum_{m=1}^{10} s_m^{sp\sigma} T_{m-1}(y) - \frac{s_1^{sp\sigma}}{2}, \tag{2.18}$$

where

$$y = \frac{R_{\alpha\beta} - \frac{b+a}{2}}{\frac{b-a}{2}}, \tag{2.19}$$

and

$$T_m(y) = \cos(m \arccos y), \tag{2.20}$$

and a (= 1.0 a_0) and b (= 7 a_0) define the range of the Hamiltonian and overlap matrices. The coefficients $h_m^{sp\sigma}$ and $s_m^{sp\sigma}$ are given in the paper by Porezag *et al.*[53] The matrix elements for $\boldsymbol{H}$ and $\boldsymbol{S}$ are shown in Figure 2.3.

Repulsive energy

The repulsive energy is determined by subtracting $E_{\rm bs}$ from the total energy obtained from the *ab initio* calculations:

$$E_{\rm rep}(R_{\alpha\beta}) = E_{\rm LDA}(R_{\alpha\beta}) - E_{\rm bs}(R_{\alpha\beta}). \tag{2.21}$$

The repulsive energy is calculated for different values of $R_{\alpha\beta}$ and fitted to a polynomial,

$$E_{\rm rep}(R_{\alpha\beta}) = \begin{cases} \sum_{n=2}^{\rm NP} d_n \left(R_c - R_{\alpha\beta}\right)^n & (R_{\alpha\beta} < R_c) \\ 0 & \text{otherwise} \end{cases} \tag{2.22}$$

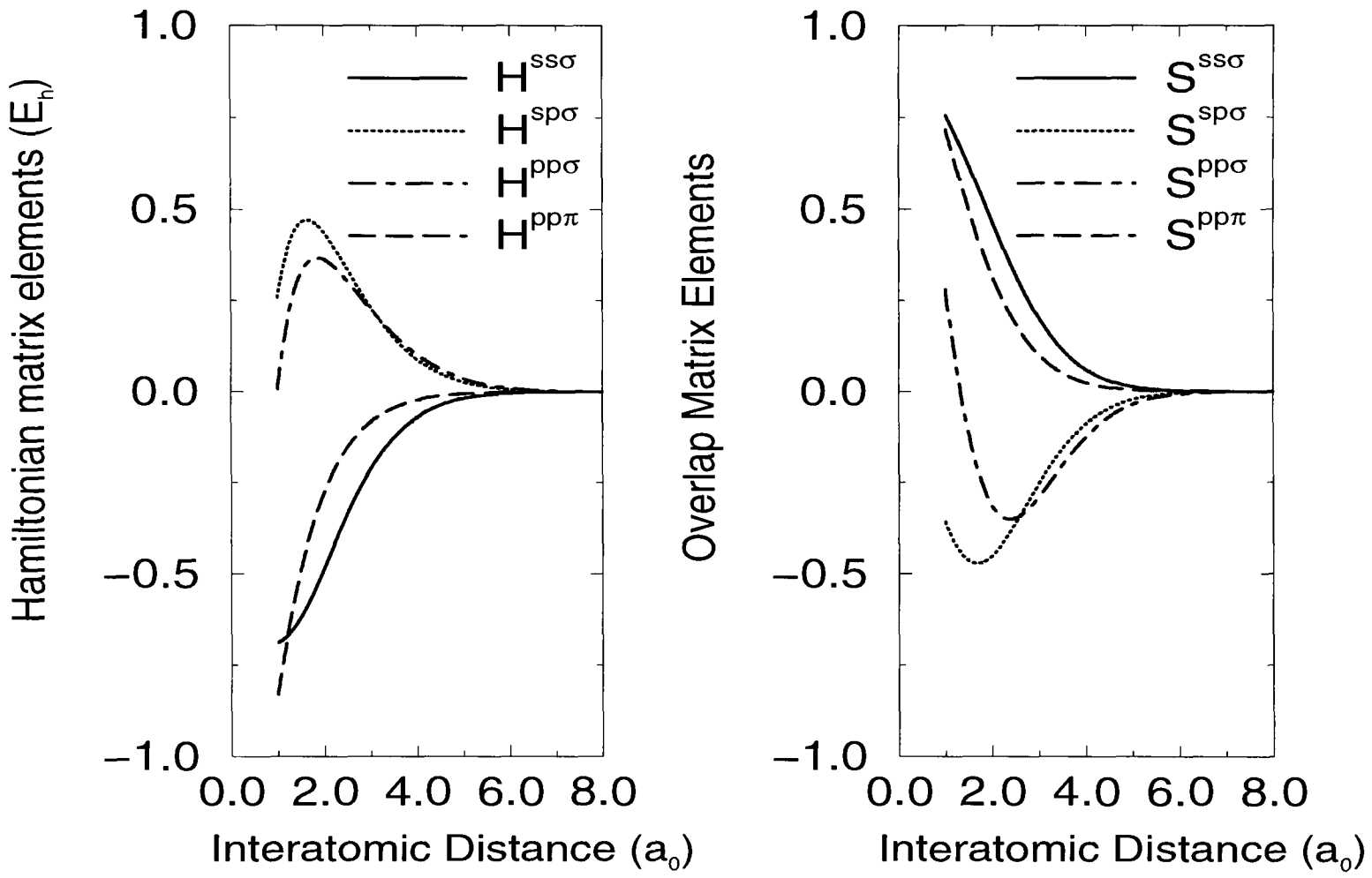


Figure 2.3: Porezag Hamiltonian matrix elements.[53]

where NP is the degree of the polynomial; a polynomial of degree five is usually sufficient. The repulsive potential is then fitted to a Chebyshev polynomial, where the coefficients are given in the literature[53] and the bounds of the Chebyshev expansion defined by Eqn. 2.19 are $a = 1\ a_0$ and $b = 4.1\ a_0$.[53] Data from LDA calculations relating to C_2, graphite and diamond is used to fit the repulsive potential.

Applications

The Porezag Hamiltonian has been used to model C_{60} , diamond and graphite.[53] As it also contains parameters for carbon-hydrogen and hydrogen-hydrogen interactions, it has been used to model hydrocarbons.[53] Figure 2.4 (from a paper by Porezag *et al.*[53]) shows the Porezag predictions for the cohesive energies for the diamond, graphite, linear chain, simple cubic, body-centred cubic and face-centred cubic structures of carbon. Again, the more highly coordinated allotropes of carbon are less well described than the lower coordinated allotropes such as graphite and diamond. Finally, the Porezag Hamiltonian also successfully predicts which of the twenty-four possible isomers of C_{84} is the ground state isomer (see Section 3.1.3 where C_{84} is discussed in the context of the Ho Hamiltonian).[70]

2.2.3 Ho versus Porezag

The Ho Hamiltonian was chosen for this work for a number of reasons. Although the Porezag Hamiltonian is non-orthogonal and therefore expected to be more readily transferable, this is not an important consideration for this

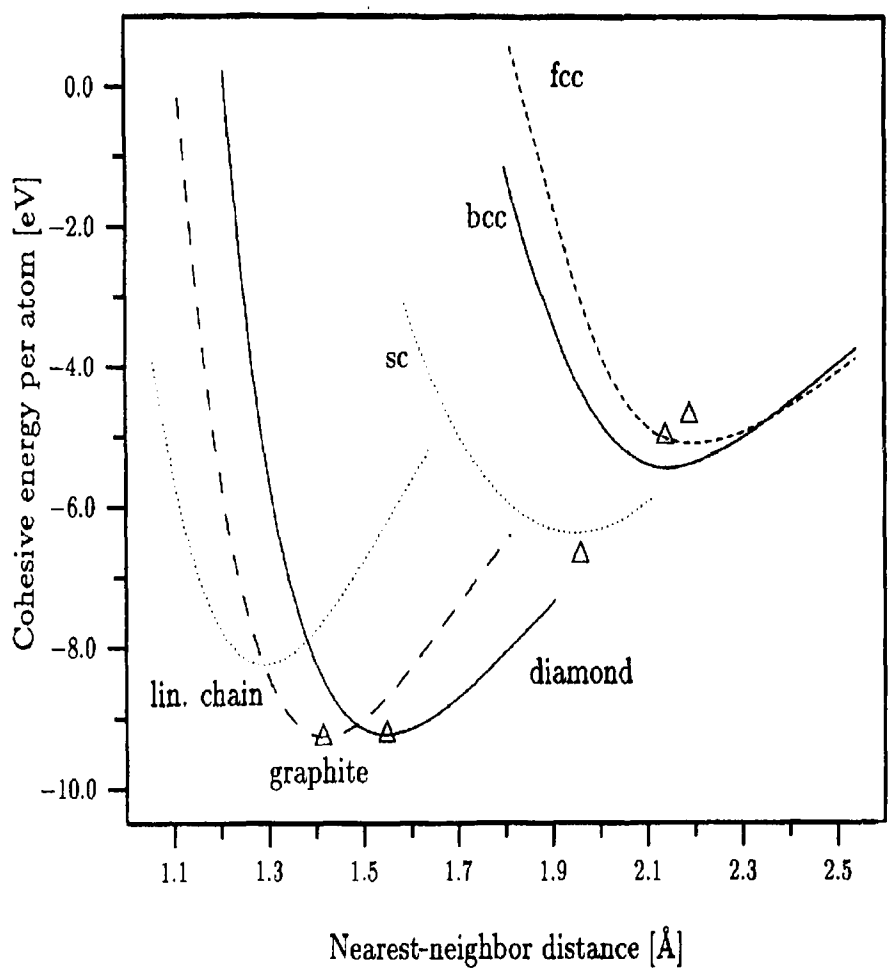


Figure 2.4: Energy curves against nearest neighbour internuclear separation for a variety of carbon allotropes, including the face-centred cubic (FCC), body-centred cubic (BCC), hexagonal close-packed (HCP), and simple cubic (SC) structures. The SCF-LDA results are shown as triangles. This figure is taken from a paper by Porezag *et al.*[53]

thesis. This is because this thesis examines capped carbon nanotubes which involve three-valent carbon and the Ho Hamiltonian is fitted to data from graphite which also involves three-valent carbon. Therefore, since the local environment of the carbon atoms is likely to be similar to the parameterisation data, the Hamiltonian should yield accurate results. Chapter 3 will demonstrate the excellent agreement between the Ho Hamiltonian and both experimental and theoretical results for fullerenes and infinite nanotubes.

Furthermore, although the linear scaling density matrix method presented in Chapter 4 can be generalised to non-orthogonal systems, the algorithm is much more complicated and requires more memory and CPU time than in the orthogonal case. This is an important consideration when molecules with large number of atoms are to be studied. It is for these reasons that the Ho Hamiltonian was chosen for this work.

2.3 Summary

Tight-binding theory is a very popular method for studying the electronic structure of a molecule containing large numbers of atoms. While it is not as accurate as *ab initio* methods, the results obtained are reasonable in many circumstances. The Ho Hamiltonian has been chosen for the work in this thesis. This is in part owing to its simplicity versus the other candidate, but mostly owing to its success in reproducing the electronic properties of low coordinated carbon systems. The next chapter discusses its application to fullerenes and infinite nanotubes, which can be viewed as the limiting cases for capped nanotubes.

Chapter 3

Limiting cases: Fullerenes and infinite nanotubes

In the introduction to this thesis it was suggested that fullerenes and infinite nanotubes could be regarded as limiting cases of capped carbon nanotubes. In this chapter both fullerenes and infinite nanotubes are reviewed. The results obtained using the Ho Hamiltonian introduced in the previous chapter are compared with other theoretical calculations and experimental evidence in order to test the suitability of this Hamiltonian for calculations on capped nanotubes. I shall start by discussing the fullerenes and then move on to discuss infinite nanotubes.

3.1 Fullerenes

Fullerenes are cage-like molecules consisting of three-valent carbon atoms arranged in hexagons and pentagons. For any spherical polyhedron (*e.g.* a fullerene), the number of vertices v, the number of faces f and the number of edges e are related by Euler's theorem:

$$v + f = e + 2. \tag{3.1}$$

Since, for a fullerene, the number of vertices is clearly just the number of atoms, the number of faces is the sum of the number of hexagons h and pentagons p, and the number of edges is $3N_{\text{atom}}/2$, Eqn. 3.1 becomes

$$h + p = \frac{N_{\text{atom}}}{2} + 2. \tag{3.2}$$

Furthermore, since each pentagon (hexagon) has five (six) vertices, each of which is shared by three faces, it follows that

$$6h + 5p = 3N_{\text{atom}}. \tag{3.3}$$

On solving Eqns. 3.2 and 3.3, one obtains the result that any fullerene must contain $p = 12$ pentagons and $h = (N_{\text{atom}}/2 - 10)$ hexagons. The arrangement of these pentagons and hexagons controls the shape and symmetry of the fullerene isomer.

However, for a given $C_{N_{\text{atom}}}$ fullerene there are (in general) a number of possible isomers satisfying the criteria given above. There is also no simple way of predicting beyond doubt the ground state isomer without considering all the possible isomers. Therefore, in order to make a systematic study of the fullerenes it is crucial to be able to identify all these isomers.

3.1.1 Cataloguing the fullerene isomers

There are two approaches to this problem. One can either find all the possible fullerene isomers (with any number of atoms) with a given symmetry, or one can find all possible isomers for a given number of atoms. Indeed, both these approaches have been pursued.[40] Finding all the possible isomers with a given symmetry suffers from two disadvantages. First, this method is increasingly complicated to apply as the symmetry decreases. Secondly, one is usually interested in all the isomers of a particular fullerene $C_{N_{\text{atom}}}$ and not all the fullerenes with a given symmetry; finding these may result in excessive work.

The archetypal method for finding all the possible isomers for a given fullerene is the spiral algorithm.[71] This method rests on the spiral conjecture: any fullerene can be represented by a code consisting of the numbers '5' (for a pentagon face) and '6' (for a hexagon face) which is generated by choosing a starting face (pentagon or hexagon) and then winding through neighbouring faces in a spiral fashion such that each face is listed once and only once. From this spiral code it is possible to construct an adjacency list detailing the atoms next to each atom and hence the bonding connectivity of the isomer.[72] Since the pentagons are represented by a '5' and the hexagons by a '6' each isomer is represented by a many-digit number which may be used to label the isomer, thus the spiral code serves a dual purpose. For some isomers, many spirals exist so the spiral that is lowest numerically is taken as the definitive one. Full details of this method (including a computer program which implements it) can be found in the literature.[40]

The spiral conjecture does break down for larger fullerenes[73] but, neverthe-

less, the method remains very useful as counterexamples are likely to be very large fullerenes where there are so many almost isoenergetic isomers that identifying the unique "most stable" isomer is less relevant to experimental results. For example, the first fullerene isomer which was shown to be without a spiral was an isomer of C_{380} with T symmetry.[73] Moreover, a more recent and completely rigorous fullerene enumeration method developed by Brinkmann and Dress[74] has shown that the spiral algorithm predictions are correct up to $N_{\text{atom}} = 100$.

3.1.2 Predicting the ground state isomer

The number of isomers of a fullerene $C_{N_{\text{atom}}}$ can be large and it increases rapidly with N_{atom}. For instance, there are 1812 possible isomers of C_{60}, but 8149 possible isomers for C_{70}.[40] It is clear that *ab initio* calculations on this number of isomers are impossible (and likely to remain so for some time to come), and it is not even realistic to expect to do full valence tight-binding calculations on every isomer. However, though there are a large number of C_{60} isomers, only one is ever reported experimentally. This suggests that there may be some way of reducing the number of isomers that need to be considered.

Indeed, it turns out that there is a simple, yet powerful, rule which governs the stability of fullerenes: the isolated pentagon rule (IPR).[75,76] This predicts that isomers with adjacent pentagons will be less stable than those where the pentagons are surrounded by hexagons. This can be rationalised on both steric and electronic grounds.[40] The carbon atoms in a sheet of graphite

are sp^2 hybridised. However, in fullerenes, some of bond angles are reduced from 120° which induces rehybridisation away from the sp^2 ideal. The π orbitals in the fullerene will no longer be ideally placed for maximal overlap, resulting in weaker bonding. Since the pentagons induce curvature, the bond angles will be more distorted where two pentagons are adjacent and therefore the rehybridisation will be greater. The electronic argument in favour of the isolated pentagon rule is as follows: two pentagons form a 8π electron system which is anti-aromatic according to the Hückel $4n$ rule and hence can be expected to be destabilising.

However, the isolated pentagon rule is of no use when $N_{\text{atom}} < 60$ (since C_{60} is the first fullerene to have any isolated pentagon isomers) and it is of limited use as $N_{\text{atom}} \to \infty$ (because the number of isolated pentagon isomers increases rapidly). There have been various attempts to develop simple rules which go beyond the isolated pentagon rule. For example, the isolated pentagon criterion can be generalised for isomers with pentagon adjacencies[40] and it has been suggested that the local environment of the hexagons in isolated pentagon isomers determines stability.[40,77]

In fact, the generalised isolated pentagon criterion has proved to be very effective. As the number of adjacent pentagons increases, the degree of rehybridisation is expected to increase and hence also the orbital misalignment between neighbouring atoms in this region, which will destabilise the isomer. This is the idea behind the generalised isolated pentagon criterion. The local environment of the pentagons can be characterised using pentagon neighbour indices.[40] The number of pentagons with k adjacent pentagons is p_k where

$k = 0, 1 \ldots 5$. The total number of fused pentagon pairs can written in terms of these indices as

$$N_p = \sum_{k=1}^{5} \frac{kp_k}{2}. \tag{3.4}$$

Now, the smaller N_p is, the fewer adjacent pentagons are present and the more stable the isomer should be. In fact, when this criterion was used to study the isomers of all fullerenes in the range $N_{\text{atom}} = 20$ to 70,[40] the ground state isomer predicted by the Ho Hamiltonian[66] for each $C_{N_{\text{atom}}}$ was always one of the optimum minimum N_p isomers. Furthermore, where several isomers possessed the optimum N_p value, the ground state isomer was always found to be one of those with the minimum p_k value for the maximum value of k. Thus, the generalised isolated pentagon rule is an accurate, dependable predictor of fullerene stability.[40]

Nevertheless, the generalised isolated pentagon rule becomes less and less selective as N_{atom} increases because the number of isolated pentagon isomers increases rapidly with N_{atom} when $N_{\text{atom}} > 60$. Unfortunately, there is no such reliable criterion for distinguishing between isolated pentagon isomers, the relative energies of which are generally much more similar than those of isomers with adjacent pentagons. Thus there is a real need for explicit electronic structure calculations (such as tight-binding calculations).

3.1.3 Full valence tight-binding calculations

The most stable isomer(s) of a fullerene can generally be predicted quite satisfactorily by full valence tight-binding calculations[40] (Hückel calculations cannot be relied on owing to the $\sigma - \pi$ mixing induced by the curvature of

the molecule). As an illustration of this, and in order to demonstrate the accuracy of the Ho Hamiltonian,[41] I shall discuss two fullerenes: C_{76} and C_{84}. The optimum structures of the fullerene isomers examined in this work were obtained using the conjugate gradient method.[69] This method uses the forces to find the minimum energy and hence the optimum structure. The last chapter discussed how the energy of a molecule is found within the tight-binding model and also how the forces may be constructed using the Hellmann-Feynman theorem.[59,60] This still requires an initial guess for the coordinates; these were generated by the CaGe program[78] which uses the fullerene generation method of Brinkmann and Dress.[74]

Actually, Ho *et al.* have already used their Hamiltonian (along with simulated annealing) to predict the ground state structures of both smaller ($C_{N_{\rm atom}<70}$)[66] and larger ($C_{N_{\rm atom}>60}$)[65] fullerenes, but they did not systematically consider all the (IPR) isomers of a given $C_{N_{\rm atom}}$. However, they did carry out such an analysis for C_{84}.[64] Therefore, the following demonstrations serve mainly to check the coding of the energy and the forces in the Ho Hamiltonian.

C_{76}

C_{76} has two isolated pentagon isomers, an isomer with D_2 symmetry and an isomer with T_d symmetry, both of which are shown in Figure 3.1. The tight-binding calculations performed in this work predict that the D_2 isomer is the ground state isomer. *Ab initio* calculations also predict that the D_2 isomer will be more stable and experimental evidence supports this conclusion.[79]

The *ab initio* calculations also predict that the T_d isomer will be unstable with respect to a Jahn-Teller distortion to D_{2d} symmetry.[79]

This distortion is due to the fact that the undistorted T_d isomer contains two electrons in a T_1 symmetry orbital manifold. A Jahn-Teller distortion lowers the symmetry of the molecule to a subgroup of the original symmetry group. According to the epikernel principle,[80] the distorted molecule will have the highest possible symmetry group that resolves the degeneracy. For the case in hand, a distortion to T symmetry will not resolve degeneracy, however, a distortion to D_{2d} symmetry results in a doubly occupied A_2 orbital and an unoccupied pair of E orbitals. An investigation of the eigenvalues obtained in the tight-binding calculation carried out for this work showed that the optimised geometry of the "T_d" isomer is distorted in this way. However, the effect is small and cannot be seen graphically in Figure 3.1.

The energy difference between the distorted T_d isomer and the D_2 isomer is 1.85 eV according to *ab initio* calculations,[79] but the Ho Hamiltonian only predicts an energy difference of 0.23 eV. This is a large discrepancy but, nevertheless, the predictions of the Ho Hamiltonian are qualitatively correct.

C_{84}

The ground state of C_{84} was a hotly debated topic for some time.[64,65,77,81] In part, this was because there are 24 IPR isomers,[72] which is considerably more than the previous example, and all of these need to be considered. The two ground state isomers, labelled C_{84}: 22 (D_2) and C_{84}: 23 (D_{2d}) in the spiral notation of "An Atlas of Fullerenes",[40] were finally identified by

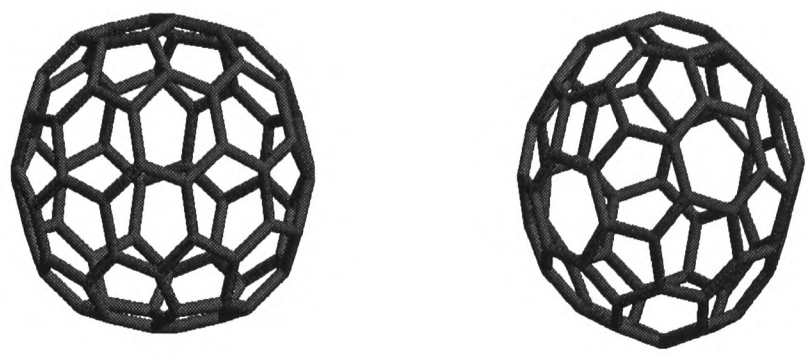


Figure 3.1: The two lowest isomers of C_{76}: D_2 (left) and T_d (right). These structures have been geometry optimised and the T_d isomer has undergone a (slight) Jahn-Teller distortion to D_{2d} symmetry.

a combination of energetic calculations,[64,77] experimental ^{13}C NMR spectra and thermodynamic reasoning.[81]

Since there are more isomers, C_{84} provides a more rigorous test for the Ho Hamiltonian. The energies per atom of these isomers (calculated using the Ho Hamiltonian) are plotted in Figure 3.2. The Ho Hamiltonian correctly predicts the two lowest energy structures, both of which are found experimentally,[81] and there is good agreement with previous work as to the ordering of the other 22 isomers.[64,70] The two lowest energy isomers are illustrated in Figure 3.3.

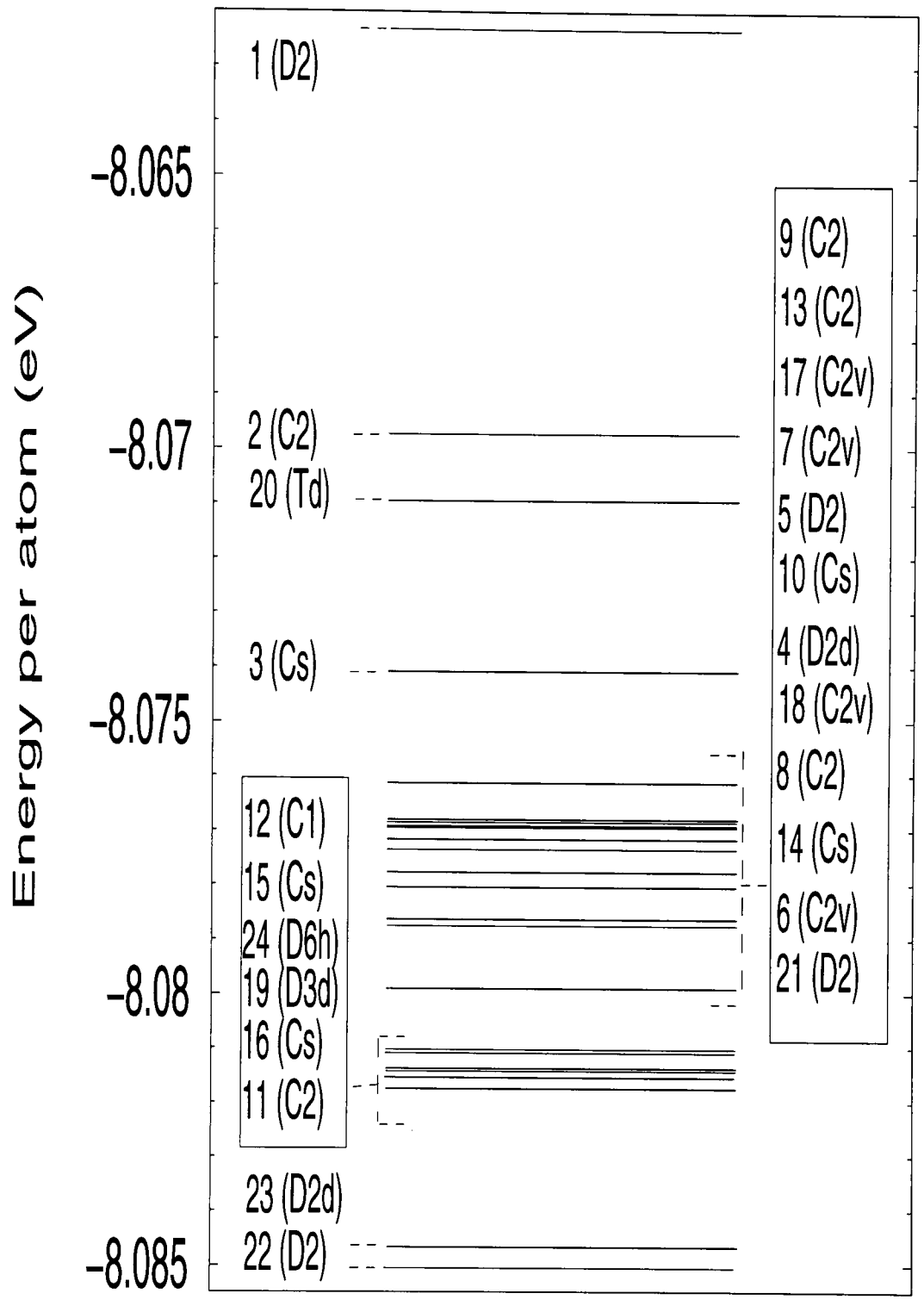


Figure 3.2: Calculated energies per atom of the 24 isolated pentagon isomers of C_{84}, labelled according to the IUPAC convention as set out in "An Atlas of Fullerenes".[40]

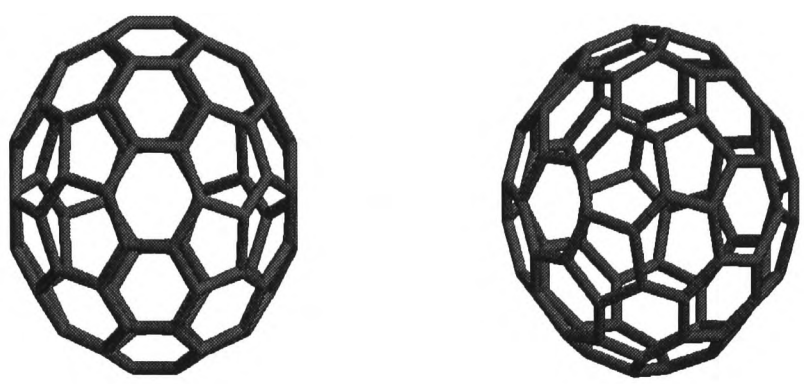


Figure 3.3: The two lowest energy isomers of C_{84}: 22 (D_2, left) and 23 (D_{2d}, right).

3.1.4 Fullerene summary

The results I have just presented for selected fullerenes demonstrate that the Ho Hamiltonian gives a reasonable description of the energetics of different IPR fullerene isomers. The relative stabilities of IPR fullerene isomers is rather subtle and it is something for which there are no simple rules. This is also likely to be the case for nanotube caps with isolated pentagons and it is therefore essential that any Hamiltonian used to study nanotube caps is able to distinguish between IPR isomers in a realistic way.

3.2 Infinite nanotubes

Nanotubes have been subject to intensive experimental and theoretical investigation ever since they were discovered in 1991.[8] Nanotubes have often

been approximated as infinite because the diameters of the tubes are typically two or more orders of magnitude smaller than their lengths. This section examines some of the properties of infinite nanotubes.

3.2.1 Hamada indices

Let us assume that the nanotube body is perfect, that is all the carbon atoms are arranged in hexagons (some experimental observations of nanotubes show defects in the structure resulting from the presence of heptagons and pentagons,[82] but these will not be dealt with in this thesis). The structure implied by this assumption is reminiscent of the structure of a sheet of graphite, a planar structure where all the carbon atoms are also arranged in (regular) hexagons. The only difference is that the nanotubes are cylindrical and not planar. This suggests that it should be possible to somehow relate nanotubes to graphite. In fact, a nanotube can be thought of as a rolled-up strip of graphene. This is a rather profitable comparison because it leads directly to a very useful classification for nanotubes *i.e.* in terms of their Hamada indices.[83]

Consider Fig. 3.4, which shows a single sheet of graphite (otherwise known as graphene) with the lattice vectors $\boldsymbol{a}_1 = (a, 0)$ and $\boldsymbol{a}_2 = (a/2, \sqrt{3}a/2)$, which form the edges of the unit cell where $a = \sqrt{3}\, r_{CC}$ and $r_{CC} = 1.42$ Å, the bondlength in a sheet of graphite. When a nanotube is mapped onto the planar graphene, the circumference of the nanotube is represented by a vector on the lattice (the vector $\boldsymbol{R}$ in Fig. 3.4), the starting and finishing points of which are equivalent by translational symmetry. Since any such

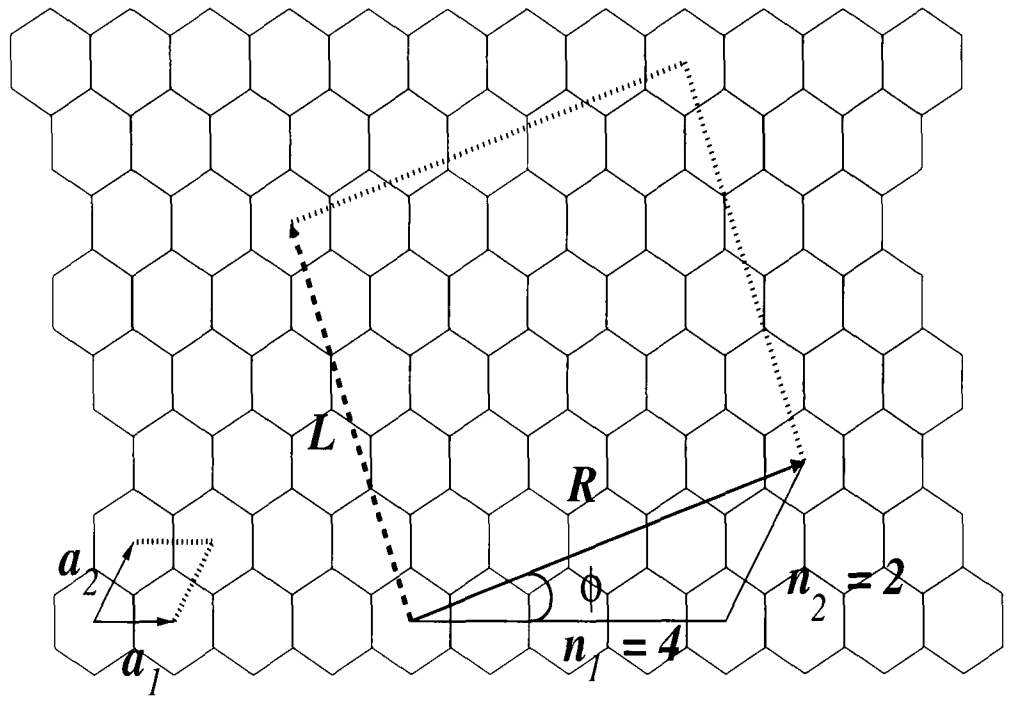


Figure 3.4: A sheet of graphene where the lattice vectors $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ define the unit cell of graphene. The translational unit cell of the illustrated (4,2) nanotube is defined in the text.

vector may be expressed in terms of the lattice vectors $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$, we can write $\boldsymbol{R} = n_1\boldsymbol{a}_1 + n_2\boldsymbol{a}_2$, where n_1 and n_2 are integers. The nanotube (n_1,n_2) is uniquely specified by these two numbers, known as the Hamada indices; for example, Fig. 3.4 shows the translational unit cell of the (4,2) nanotube. In this thesis only nanotubes with $n_1 \geq n_2 \geq 0$ will be considered. Nanotubes which do not fulfil this criterion are simply enantiomers of tubes within our definition and since this does not affect the electronic structures it is not necessary to consider these tubes as separate cases.

As well as specifying which nanotube is under consideration, many of the structural properties of a tube may be written in terms of these Hamada indices using simple geometry. Since a nanotube can be alternatively specified in terms of the radius r and the chiral angle ϕ *i.e.* the angle between $\boldsymbol{a}_1$ and $\boldsymbol{R}$, it is particularly useful to be able to express these in terms of the Hamada indices. The radius is given by

$$r = \frac{|\boldsymbol{R}|}{2\pi} = \frac{a}{2\pi}\left(n_1^2 + n_1 n_2 + n_2^2\right)^{\frac{1}{2}}, \tag{3.5}$$

and the chiral angle by

$$\phi = \arctan\left(\frac{\sqrt{3}n_2}{2n_1 + n_2}\right), \tag{3.6}$$

where $0 \leq \phi \leq \pi/6$.

The number of atoms in the translational unit cell of a nanotube may also be obtained from n_1 and n_2. Fig. 3.4 shows the unit cell for the (4,2) nanotube. The length of the unit cell $|\boldsymbol{L}|$ can be calculated by writing $\boldsymbol{L} = l_1\boldsymbol{a}_1 + l_2\boldsymbol{a}_2$ (where l_1 and l_2 are integers) and using the fact that $\boldsymbol{R}$ and $\boldsymbol{L}$ are orthogonal

to deduce that

$$\frac{l_1}{l_2} = -\frac{2n_2 + n_1}{2n_1 + n_2}. \tag{3.7}$$

It then follows that the smallest possible values of l_1 and l_2 are given by

$$l_1 = -\frac{2n_2 + n_1}{N_\mathrm{r}} \ ,$$

and

$$l_2 = \frac{2n_1 + n_2}{N_\mathrm{r}} \ ,$$

where N_r is the highest common factor of n_1 and n_2, and the order of the rotational symmetry group of the nanotube. The number of atoms in the translational unit cell of the nanotube is just twice $|\boldsymbol{R}| \times |\boldsymbol{L}|$ divided by the area of the unit cell of graphene, that is,

$$N_\mathrm{uc} = \frac{4\left(n_1^2 + n_1 n_2 + n_2^2\right)}{N_\mathrm{r}}. \tag{3.8}$$

However, this leads us to a problem. Techniques exist for elucidating the electronic structure of an infinite periodic structure, but these still depend on the number of atoms in the unit cell. Eqn. 3.8 shows that the number of atoms in the translational unit cell increases rapidly as the diameter increases, which severely limits such calculations. For example, the (11,2) nanotube has 688 atoms in its unit cell and yet has a radius of just 0.95 nm which is by no means the largest in the experimental range.[23] Moreover, in the previous section we saw how nanotubes may be related to graphene, which contains only two atoms per unit cell suggesting that such large unit cells based on translation along the tube axis are not the best possible choice. In 1993, White, Robertson and Mintmire showed that fully exploiting the helical and

rotational symmetries of a nanotube enables one to fully specify the nanotube structure with a unit cell of only two atoms.[84] We shall discuss this next.

3.2.2 Exploiting the helical and rotational symmetries

Once again the key idea is the relationship between graphite and a nanotube. However White, Robertson and Mintmire explored this relationship more fully. When a nanotube is projected onto a graphene sheet, the strip of graphene contains all the necessary information to construct the nanotube and since the graphene strip can be obtained using a unit cell containing just two atoms, it follows that this is also true for the nanotube.

The carbon atoms in the unit cell of graphene must first be mapped onto the surface of the nanotube. The first atom may be placed at any arbitrary point, the second must be positioned relative to the first. This is accomplished by a rotation about the tube axis followed by a translation along the same axis. In the graphene sheet, this is equivalent to a translation along $\boldsymbol{R}$ and then a perpendicular translation. The length of the translation along $\boldsymbol{R}$ may be found by simple geometry, but when we consider the tube it corresponds to the length of an arc, so the angle of rotation can be found by multiplying by 2π and dividing by the circumference $|\boldsymbol{R}|$. Thus, the angle of rotation about the tube axis is

$$\theta = \frac{2\pi\,(\boldsymbol{d}\cdot\boldsymbol{R})}{|\boldsymbol{R}|^2}, \tag{3.9}$$

(in radians) and the distance of the translation T along the tube axis can be determined to be

$$T = \frac{|\boldsymbol{d}\times\boldsymbol{R}|}{|\boldsymbol{R}|}, \tag{3.10}$$

where $\boldsymbol{d} = (\boldsymbol{a}_1 + \boldsymbol{a}_2)/3$ is the vector connecting the bottom corner of the unit cell and the first atom, or, equivalently, the first and second atom in the unit cell.

The helical motif generated in this fashion can be used to position all the atoms in the rest of the tube by repeatedly applying a rotation operator $\hat{\mathcal{C}}_{N_\mathrm{r}}$ and a screw operation $\hat{\mathcal{S}}(\alpha, h)$, which is composed of a translation h along the tube axis and a rotation α about this axis. It is possible to determine h and α because there must be a vector $\boldsymbol{H} = p_1\boldsymbol{a}_1 + p_2\boldsymbol{a}_2$ in the graphene lattice such that

$$h = \frac{|\boldsymbol{H} \times \boldsymbol{R}|}{|\boldsymbol{R}|}, \tag{3.11}$$

and

$$\alpha = \frac{2\pi\,(\boldsymbol{H} \cdot \boldsymbol{R})}{|\boldsymbol{R}|^2}. \tag{3.12}$$

The problem of finding p_1 and p_2 can be solved by noting that the area covered by the helical motif is given by $|\boldsymbol{H} \times \boldsymbol{R}|$ and also by N_r times the area of the unit cell of graphene which, by elementary geometry, is $|\boldsymbol{a}_1 \times \boldsymbol{a}_2| N_\mathrm{r}$. Algebraic manipulation results in the following relation:

$$p_2 n_1 - p_1 n_2 = \pm N_\mathrm{r}. \tag{3.13}$$

In order to define p_1 and p_2 uniquely, White *et al.* suggest taking $n_1 \geq n_2 \geq 0$, $p_1 \geq 0$, assuming the positive sign in Eqn. 3.13 and then finding the solution yielding the minimum value of $|\boldsymbol{H}|$. It should be noted that only α is affected by these conditions; h is independent of $\boldsymbol{H}$.

Now, we can define the (m, l) cell on the surface of the nanotube as that obtained by mapping the $[0, 0]$ cell of graphene onto the surface of the nanotube and then applying the rotation operator $\hat{\mathcal{C}}_{N_\mathrm{r}}$ l times, followed by the

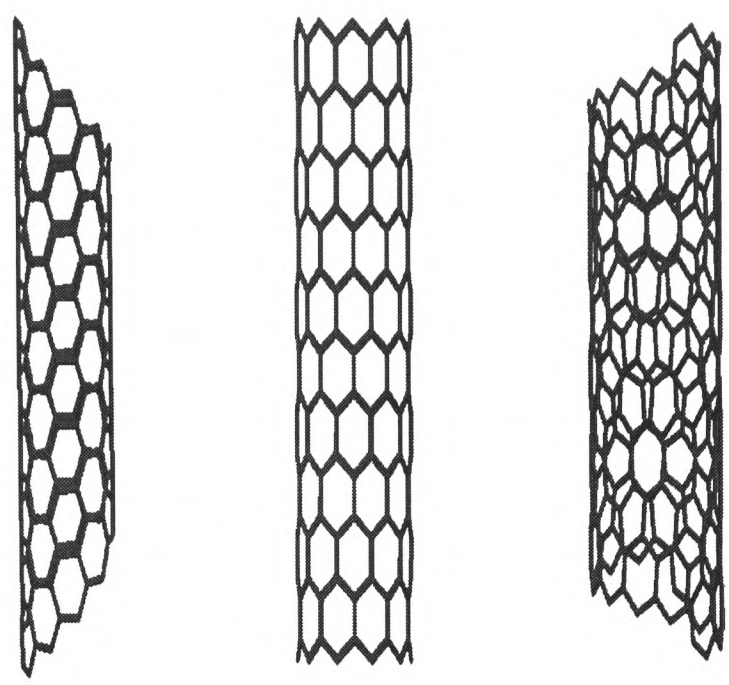


Figure 3.5: Fragments of the nanotubes corresponding to the Hamada indices (5,5) (left);(10,0) (centre) and (11,2) (right).

screw operator $\hat{\mathcal{S}}(\alpha, h)$ m times. These operations completely specify the tube and yet there are only two atoms in the unit cell. It is useful to relate the position of the unit cell on the surface of the nanotube to a unit cell on the graphene sheet. If the cell on the surface of the tube located at (m, l) corresponds to the cell in the plane labelled as $\boldsymbol{r}_{t_1 t_2} = t_1 \boldsymbol{a}_1 + t_2 \boldsymbol{a}_2$, it follows from the definition of $\boldsymbol{R}$ and $\boldsymbol{H}$ and Eqn. 3.13 that

$$m = \frac{t_2 n_1 - t_1 n_2}{N_\mathrm{r}}, \tag{3.14}$$

and

$$l = t_1 p_2 - t_2 p_1, \tag{3.15}$$

where l is defined modulo N_r.

It is clear that the work of White *et al.*[84] allows us to construct a nanotube knowing only some simple properties of a graphene sheet and the Hamada indices of the tube in question. In this thesis, all properties discussed for infinite nanotubes are calculated from nanotubes with coordinates obtained in this manner with no geometry optimisation. Previous investigations have concluded that omitting the geometry optimisation of a nanotube has minimal effect on the results,[85] explained by the experimental observation that the C-C bondlength in a nanotube is very close to the graphite ideal (1.42 Å).[13] Fig. 3.5 shows the (5,5); (10,0) and (11,2) nanotubes constructed using this method.

Furthermore, the work of White *et al.* is actually much more powerful, since it simplifies the calculation of the electronic structure as well as the geometric structure. The unit cell of the nanotube has now been reduced to one containing just two atoms which can be implemented in the electronic structure calculation. This enables us to find analytic expressions for the energy bands within the Hückel model and to perform full valence tight-binding calculations which involve the diagonalisation of no more than a (8×8) matrix.

3.2.3 Electronic structure of infinite nanotubes

In order to simpliy the notation, the following explanation considers the case of one orbital per unit cell. The generalisation to more than one orbital per unit cell (which is the case we are, in fact, interested in) will be given at the end.

Let $|m, l\rangle$ be a basis function for the unit cell on the tubule labelled by (m, l)

where m and l are defined by Eqns. 3.14 and 3.15. Symmetry adapted Bloch sums[86] may be constructed to give the wavefunctions $|\kappa, n\rangle$

$$|\kappa, n\rangle = \lim_{M\to\infty} \frac{1}{\sqrt{N_{\rm r}\,(2M+1)}} \sum_{m=-M}^{M} \sum_{l=0}^{N_{\rm r}-1} {\rm e}^{{\rm i}\kappa m} {\rm e}^{2\pi {\rm i} n l/N_{\rm r}} \,|m, l\rangle\,, \tag{3.16}$$

such that

$$\hat{\mathcal{C}}_{N_{\rm r}} \,|\kappa, n\rangle = {\rm e}^{-2\pi {\rm i} n/N_{\rm r}} \,|\kappa, n\rangle\,, \tag{3.17}$$

where $n = 0, \cdots, N_{\rm r} - 1$, and

$$\hat{\mathcal{S}}(\alpha, h) \,|\kappa, n\rangle = {\rm e}^{-{\rm i}\kappa} \,|\kappa, n\rangle\,, \tag{3.18}$$

where $-\pi < \kappa \leq \pi$. It is possible to write down the wavefunction in such a way because $\hat{\mathcal{C}}_{N_{\rm r}}$ and $\hat{\mathcal{S}}(\alpha, h)$ commute with the Hamiltonian. Now, given an expression for the wavefunction, an expression for the Hamiltonian matrix elements can be derived.

Hamiltonian matrix elements

The matrix elements of the Hamiltonian are given by

$$\begin{aligned} \langle \kappa', n'| \hat{H} \,|\kappa, n\rangle \;=\; & \lim_{M\to\infty} \frac{1}{N_{\rm r}\,(2M+1)} \sum_{m'=-M}^{M} \sum_{l'=0}^{N_{\rm r}-1} \sum_{m=-M}^{M} \sum_{l=0}^{N_{\rm r}-1} {\rm e}^{-{\rm i}\kappa' m'} \\ & \times {\rm e}^{-2\pi {\rm i} n' l'/N_{\rm r}} {\rm e}^{{\rm i}\kappa m} {\rm e}^{2\pi {\rm i} n l/N_{\rm r}} \,\langle m', l'| \hat{H} \,|m, l\rangle\,. \end{aligned} \tag{3.19}$$

However, group theory dictates that matrix elements involving wavefunctions of different symmetries are zero, thus

$$\begin{aligned} \langle \kappa', n'| \hat{H} \,|\kappa, n\rangle \;=\; & \delta_{\kappa'\kappa}\delta_{n'n} \lim_{M\to\infty} \frac{1}{N_{\rm r}\,(2M+1)} \sum_{m'=-M}^{M} \sum_{l'=0}^{N_{\rm r}-1} \sum_{m=-M}^{M} \sum_{l=0}^{N_{\rm r}-1} {\rm e}^{{\rm i}\kappa(m-m')} \\ & \times {\rm e}^{2\pi {\rm i} n(l-l')/N_{\rm r}} \,\langle m', l'| \hat{H} \,|m, l\rangle\,, \end{aligned} \tag{3.20}$$

since κ labels the screw symmetry and n the rotational symmetry. However, the matrix element $\left\langle m', l' \left| \hat{H} \right| m, l \right\rangle$ is identical to the matrix element $\left\langle 0, 0 \left| \hat{H} \right| m - m', l - l' \right\rangle$ because of the underlying periodic symmetry. Using this relationship and substituting $m'' = m - m'$ and $l'' = l - l'$ in Eqn. 3.20 results in the following expression for the matrix elements (dropping the primes in the final result):

$$\left\langle \kappa', n' \right| \hat{H} \left| \kappa, n \right\rangle = \delta_{\kappa'\kappa}\delta_{n'n} \lim_{M\to\infty} \sum_{m=-M}^{M} \sum_{l=0}^{N_r-1} \mathrm{e}^{\mathrm{i}\kappa m} \mathrm{e}^{2\pi \mathrm{i} n l/N_r} \left\langle 0, 0 \right| \hat{H} \left| m, l \right\rangle . \quad (3.21)$$

The generalised form of Eqn. 3.21 for N_{orb} orbitals per unit cell is simply

$$\left\langle \kappa', n'; p' \left| \hat{H} \right| \kappa, n; p \right\rangle = \delta_{\kappa'\kappa}\delta_{n'n} \lim_{M\to\infty} \sum_{m=-M}^{M} \sum_{l=0}^{N_r-1} \mathrm{e}^{\mathrm{i}\kappa m} \mathrm{e}^{2\pi \mathrm{i} n l/N_r} \left\langle 0, 0; p' \left| \hat{H} \right| m, l; p \right\rangle , \quad (3.22)$$

where $p = 1, \ldots N_{\mathrm{orb}}$.

It seems as though we have an infinite sum to evaluate in order to obtain the matrix element $\left\langle \kappa, n; p' \right| \hat{H} \left| \kappa, n; p \right\rangle$. However, since the range of the Hamiltonian is finite, there are actually only a few terms. Within Hückel theory (which considers nearest neighbour interactions only) the matrix elements between π orbitals on atoms i and j are $\left\langle i \left| \hat{H} \right| i \right\rangle = \left\langle j \left| \hat{H} \right| j \right\rangle = \alpha$ and $\left\langle i \left| \hat{H} \right| j \right\rangle = \left\langle j \left| \hat{H} \right| i \right\rangle = \beta$ if i and j are adjacent and zero otherwise. Therefore, there are only five unit cells which need concern us: [0,0] of the graphene lattice which becomes (0,0) on the tube surface when Eqns. 3.14 and 3.15 are used to determine m and l; [0,1] $\equiv (n_1/N_r, -p_1)$; [0,-1] $\equiv (-n_1/N_r, p_1)$;

[1,0] $\equiv (-n_2/N_\mathrm{r}, p_2)$ and [-1,0] $\equiv (n_2/N_\mathrm{r}, -p_2)$. Hence

$$\begin{aligned}\left\langle \kappa, n; p' \left| \hat{H} \right| \kappa, n; p \right\rangle &= \left\langle 0,0; p' \left| \hat{H} \right| 0,0; p \right\rangle + \left\langle 0,0; p' \left| \hat{H} \right| \frac{n_1}{N_\mathrm{r}}, -p_1; p \right\rangle \\ &+ \left\langle 0,0; p' \left| \hat{H} \right| -\frac{n_1}{N_\mathrm{r}}, p_1; p \right\rangle + \left\langle 0,0; p' \left| \hat{H} \right| -\frac{n_2}{N_\mathrm{r}}, p_2; p \right\rangle \\ &+ \left\langle 0,0; p' \left| \hat{H} \right| \frac{n_2}{N_\mathrm{r}}, -p_2; p \right\rangle . \end{aligned} \tag{3.23}$$

The Hamiltonian matrix for a given κ and n is then given within Hückel molecular orbital theory by

$$\boldsymbol{H}^{(\kappa,n)} = \begin{pmatrix} \alpha & \beta' \\ \beta'^{*} & \alpha \end{pmatrix},$$

where

$$\beta' = \beta \left(1 + \mathrm{e}^{\mathrm{i}(\kappa n_1 - 2\pi n p_1)/N_\mathrm{r}} + \mathrm{e}^{\mathrm{i}(\kappa n_2 - 2\pi n p_2)/N_\mathrm{r}} \right),$$

follows from Eqn. 3.22. Diagonalising this (2×2) matrix yields the eigenvalues[84]

$$\begin{aligned} E_{\pm}(\kappa, n) &= \alpha \pm \beta \left\{ 3 + 2\cos\left(\frac{(n_1 + n_2)\,\kappa - 2\pi n\,(p_1 + p_2)}{N_\mathrm{r}} \right) \right. \\ &\left. + 2\cos\left(\frac{n_1\kappa - 2\pi n p_1}{N_\mathrm{r}} \right) + 2\cos\left(\frac{n_2\kappa - 2\pi n p_2}{N_\mathrm{r}} \right) \right\}^{\frac{1}{2}}, \end{aligned} \tag{3.24}$$

hence providing us with an analytic solution to the band structure of any nanotube within the Hückel molecular orbital approximation.

Band structure within full valence tight-binding theory

We are now in possession of an expression for the π energy bands of a nanotube. Whilst an analysis within Hückel theory yields a qualitatively correct

account of the electronic structure of graphite, nanotubes are curved, not planar, and this curvature is expected to induce $\sigma - \pi$ mixing which Hückel theory cannot account for. This motivates a full valence tight-binding calculation and we therefore return to Eqn. 3.22. In this case, it is necessary to include unit cells from the graphene sheet to beyond the cut-off distance of the Hamiltonian and an 8×8 matrix must be diagonalised to obtain the eigenvalues $E_k(\kappa, n)$ for $k = 1, 2 \ldots, N_{\text{orb}} = 8$.

As nanotubes are cylindrical, one must use radial and tangential $2p$ orbitals rather than $2p_x$ and $2p_y$ (where the z-axis is the tube axis) because under a rotation or screw operation a p_x orbital in the nanotube will in general be converted to some combination of p_x and p_y whereas a radial p_r orbital on the first (second) atom in a unit cell is equivalent to the radial orbital on the first (second) atom in any other unit cell. The same is true of the tangential p_t orbitals (and also the s and p_z orbitals). From Fig. 3.6 it is clear that

$$\begin{pmatrix} p_r \\ p_t \\ p_z \end{pmatrix} = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix},$$

and therefore the matrix elements between radial and tangential orbitals may be expressed in terms of matrix elements between the more familiar p_x and p_y orbitals. For example, the matrix element between two radial orbitals on

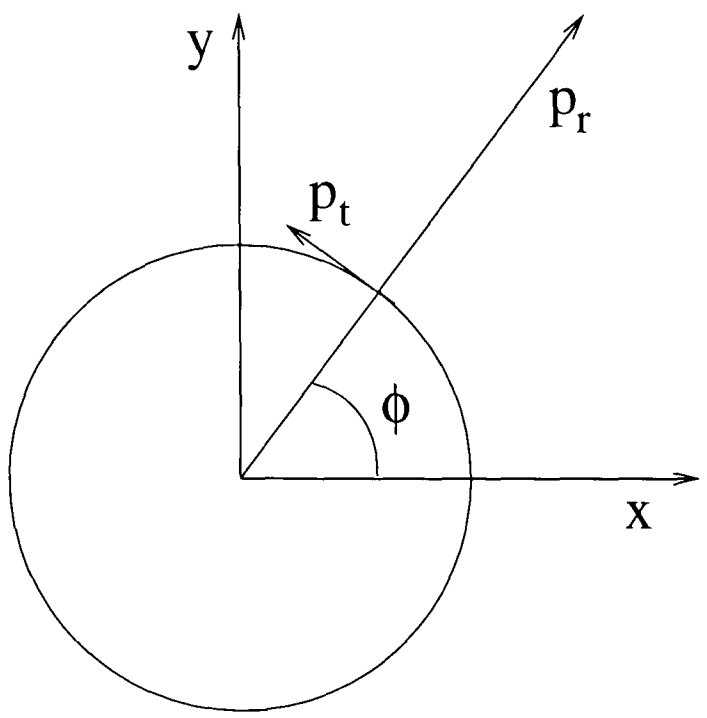


Figure 3.6: Relationship between radial and tangential orbitals and p_x and p_y.

atoms i and j is given by

$$\begin{aligned}\left\langle p_{ri}\left|\hat{H}\right|\hat{p}_{rj}\right\rangle &= \left\langle \cos\phi_i p_{xi} + \sin\phi_i p_{yi}\left|\hat{H}\right|\cos\phi_j p_{xj} + \sin\phi_j p_{yj}\right\rangle \\ &= \left\langle p_{xi}\left|\hat{H}\right|p_{xj}\right\rangle \cos\phi_i\cos\phi_j + \left\langle p_{xi}\left|\hat{H}\right|p_{yj}\right\rangle \cos\phi_i\sin\phi_j \\ &\quad + \left\langle p_{yi}\left|\hat{H}\right|p_{xj}\right\rangle \sin\phi_i\cos\phi_j + \left\langle p_{yi}\left|\hat{H}\right|p_{yj}\right\rangle \sin\phi_i\sin\phi_j.\end{aligned} \tag{3.25}$$

It is no longer instructive (or practical) to seek an analytical form for the energy, but the problem is trivially solved numerically. Figure 3.7 shows the band structure for two nanotubes: the (5,5) nanotube which is metallic as demonstrated by the fact that two bands cross at the Fermi level ($E = 0$) and the (9,2) nanotube which is a semi-conductor. The number of bands is given by $8N_{\mathrm{r}}$ in each case, where N_{r} is the order of the rotational symmetry

group of the nanotube. Thus the (5,5) tube has 40 bands whereas the (9,2) tube only has 8 bands as can be seen from Fig. 3.7. The bands are symmetric about $\kappa = 0$ because the sign of κ corresponds the direction of the screw operation which clearly has no effect on the energy bands of an infinite nanotube. These results are in keeping with other theoretical studies published in the literature.[85] It has been much more difficult to study dispersion curves experimentally. However, recent progress in this area has shown that the two dispersion curves for the metallic (13,13) tube are linear when they cross at the Fermi level,[17] and it is likely that this is the case for all (n_1,n_1) tubes, an idea which Figure 3.7 supports.

These results enable us to calculate other, more interesting, properties of nanotubes some of which I shall briefly discuss.

3.2.4 Energy per atom

The first property of nanotubes we shall examine here is the energy per atom, obtained by filling the bands with the available electrons. Fig. 3.8 shows how the energy per atom varies with radius. As the radius increases, the energy falls to that of graphene because as the curvature of the tube decreases, the strain lessens. The energy per atom of a given nanotube E_{nt} is approximately given by

$$E_{\mathrm{nt}} = E_{\mathrm{gr}} + \frac{b}{|\boldsymbol{R}|^2},$$

where E_{gr} is the energy per atom for a sheet of graphene. This relationship has been justified theoretically by reference to a continuum elastic

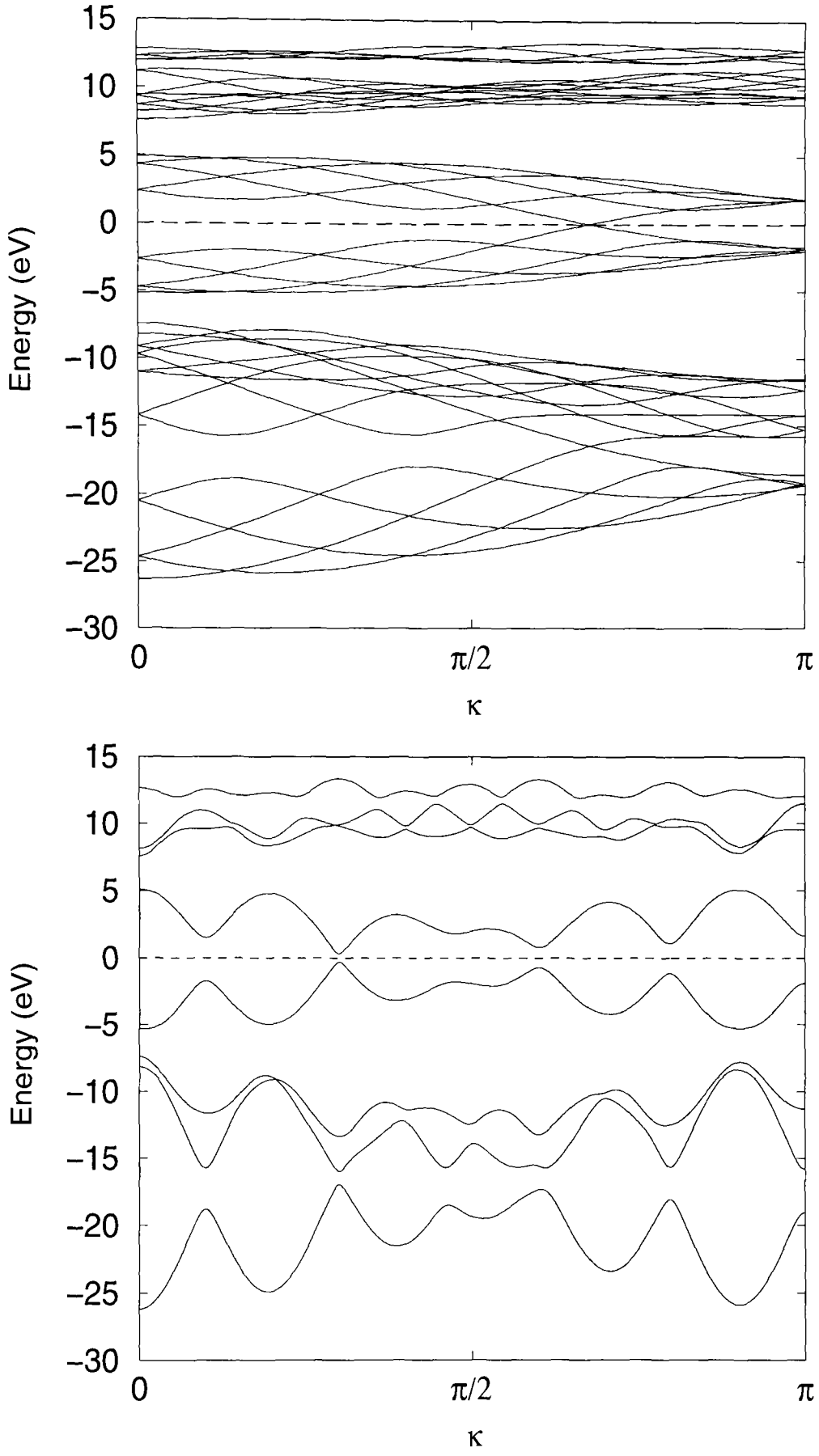


Figure 3.7: Dispersion curves for the (5,5) tube (top) and the (9,2) tube (below), obtained using the Ho Hamiltonian.

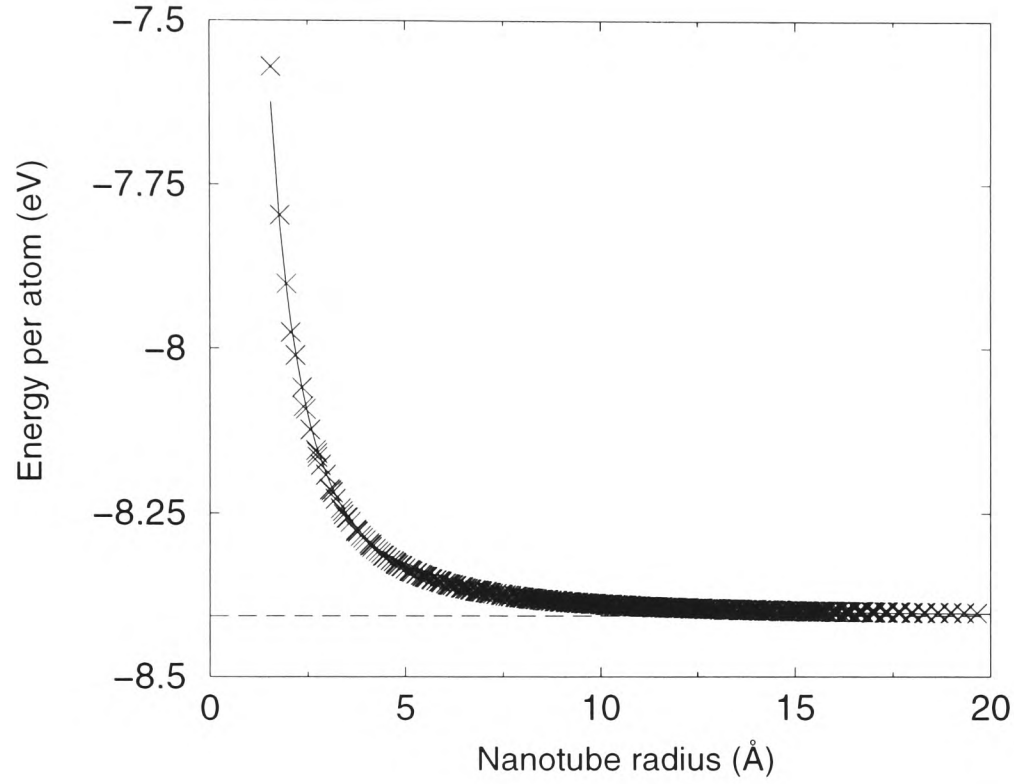


Figure 3.8: Variation of the energy per atom of a nanotube with radius. The solid line is a fit to a function of the form $E_{\mathrm{nt}} = E_{\mathrm{gr}} + b/|\boldsymbol{R}|^2$. The energy of graphite is shown as a dashed line for comparison.

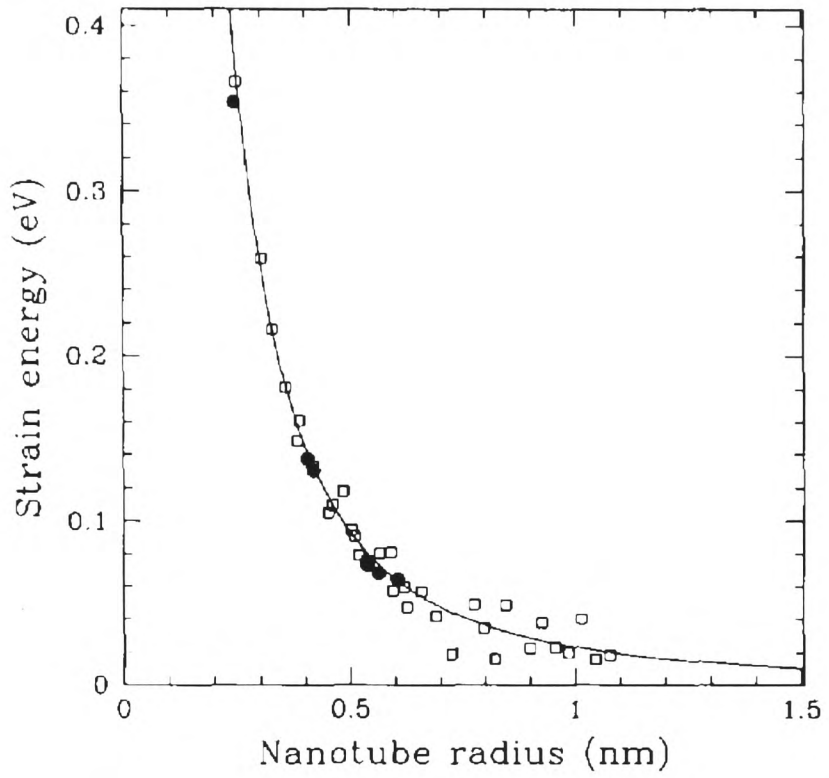


Figure 3.9: The strain energy per atom as a function of tube radius calculated for unoptimised tubes (open squares) and optimised tubes (closed squares) as obtained by Mintmire and White[85] using first principles methods. The solid line is a fit to a function of the form $E_{\rm nt} = b/|\boldsymbol{R}|^2$. This figure is from a paper by Mintmire and White.[85]

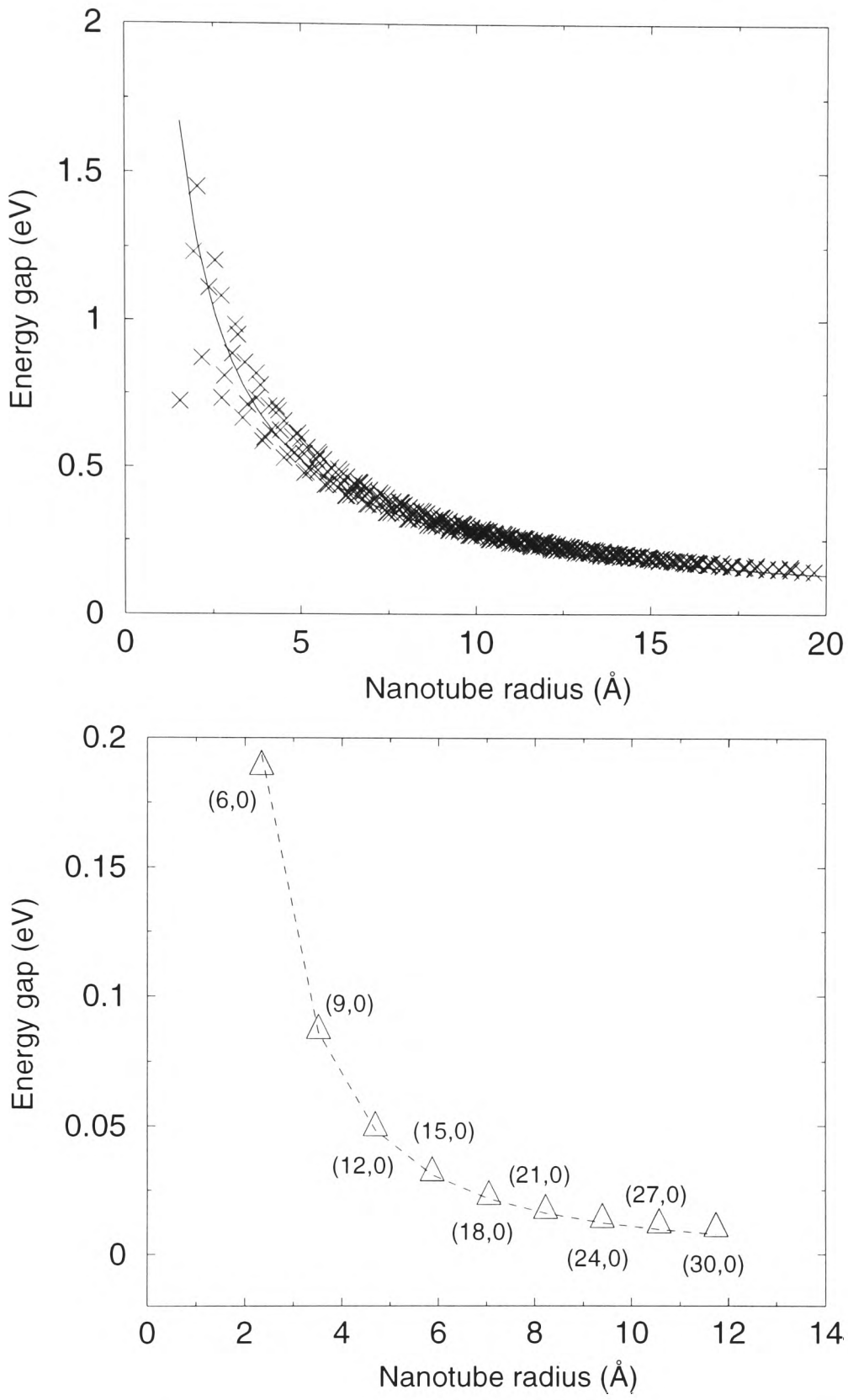


Figure 3.10: Variation of the energy gap of a nanotube with radius for nanotubes with $n_1 - n_2 \neq 3q$ (top) and (n_1,0) tubes with $n_1 = 3q$ (bottom). For $n_1 - n_2 \neq 3q$, $E_{\text{gap}} \propto |\boldsymbol{R}|^{-1}$, however, for ($n_1$,0) tubes with $n_1 = 3q$, $E_{\text{gap}} \propto |\boldsymbol{R}|^{-2}$. The data has been fitted to functions of the appropriate form.

model.[49,85,87] Fig. 3.9 shows the first principles results for the strain energy (the energy per atom minus the energy per atom of a graphene sheet).[85] The general trend is exactly the same as the Ho Hamiltonian predicts. The results presented here are in excellent qualitative agreement with these first-principles methods,[85] and also with the results in the literature for a different tight-binding model.[88]

3.2.5 Nanotube band gaps

According to Hückel theory, nanotubes with

$$n_1 - n_2 = 3q, \tag{3.26}$$

where q is an integer are metallic and those with $n_1 - n_2 \neq 3q$ are semi-conductors.[89] Figure 3.10 shows the calculated full valence tight-binding band gaps for a number of tubes which do not satisfy Eqn. 3.26. The calculated points have been fitted to a curve of the form $E_{\text{gap}} = a/|\boldsymbol{R}|$, which gives an approximate fit to the data.[84,88] This $1/|\boldsymbol{R}|$ dependence can be rationalised by considering the Hückel electronic structure of graphite.[85,90] Band gaps for semi-conducting nanotubes have been measured experimentally.[13,14,91] The predictions of this tight-binding model agree qualitatively with both the theoretical[88] and experimental results.[13,14,91] Furthermore, the predicted and experimental band gaps are roughly comparable. For example, for a tube diameter of 8 Å the measured gap is ~0.85 eV as shown in Fig. 3.11 and the predicted gap is in the range $0.5 - 0.8$ eV (dependent on the tube helicity) as can be seen in Fig. 3.10. The form of the curve $E_{\text{gap}} = a/|\boldsymbol{R}|$ is also observed to fit the experimental data.

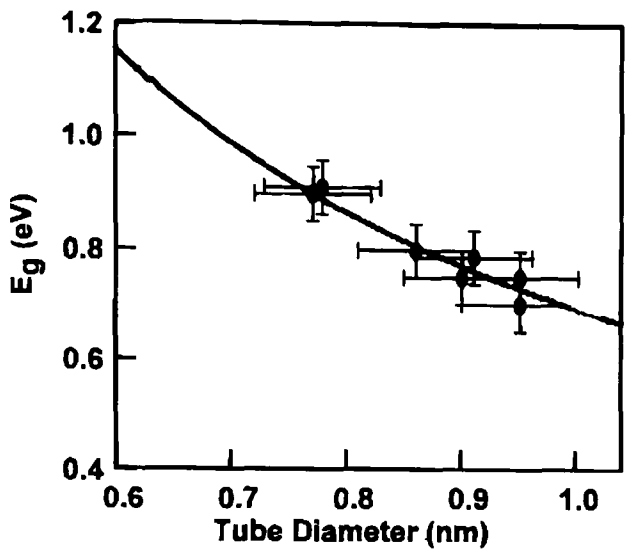


Figure 3.11: STM measurements of the band gap for semi-conducting nanotubes as reported by Odom *et al.*[91] The data has been fitted to a function of the form $E_{\text{gap}} = 2\gamma_0 r_{CC}/|\boldsymbol{R}|$ where γ_0 is 2.5 eV according to this data. This figure is from a paper by Odom *et al.*[91]

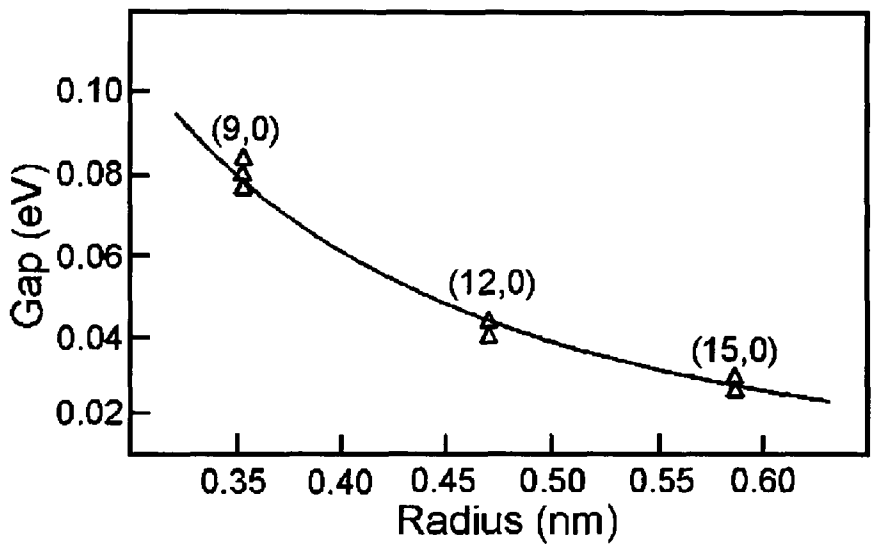


Figure 3.12: Low temperature STM measurements of curvature induced band gap for (n_1,0) tubes as measured by Ouyang *et al.*[17] The data has been fitted to $3\gamma_0 r_{CC}^2/(16|\boldsymbol{R}|^2)$ where γ_0 is 2.60 eV. This figure is from a paper by Ouyang *et al.*[17]

In reality, a small band gap opens up for nanotubes satisfying Eqn. 3.26, unless $n_1 = n_2$, because of $\sigma - \pi$ mixing. Figure 3.10 shows the energy gaps predicted using the Ho Hamiltonian for $(n_1, 0)$ tubes satisfying Eqn. 3.26. The gaps for these tubes are an order of magnitude smaller than those for $n_1 - n_2 \neq 3q$, a hangover of the Hückel prediction that $n_1 - n_2 = 3q$ tubes will be metallic. The experimental results obtained by Ouyang *et at.*[17] are shown in Fig. 3.12 for comparison. Consideration of the effect of misalignment of the π orbitals leads to the prediction that the band gap for the tubes should depend on $|\boldsymbol{R}|^{-2}$.[17] Both the experimental measurements and the Ho Hamiltonian results support this conclusion.

3.3 Summary

This chapter has discussed fullerene isomers and some of the properties of infinite carbon nanotubes. In particular, the Ho Hamiltonian has been shown to give the correct energy ordering of IPR fullerene isomers for both C_{76} and C_{84}. Furthermore, the electronic energy per atom and band gaps for a wide variety of infinite nanotubes are in good agreement with previous calculations and experiments respectively. This excellent qualitative agreement between experiments and the results presented clearly demonstrate that the Ho Hamiltonian is appropriate for studies of capped carbon nanotubes.

Chapter 4

Density matrix methods

The previous chapters have discussed a tight-binding Hamiltonian for carbon, the Ho Hamiltonian, and illustrated its applicability to both infinite nanotubes and fullerenes. The reliability of the Ho Hamiltonian for these classes of materials strongly suggest that it is an appropriate choice for studying capped carbon nanotubes. Since the Hamiltonian has been chosen, in theory it is relatively easy to calculate the energy for the molecule as detailed in Chapter 2. However, when this strategy is implemented for larger and larger systems, it quickly becomes too slow for convenience and the whole method eventually becomes impractical. One reason why the calculation becomes so time consuming as N_{atom} increases is that it involves diagonalising a matrix and the effort required to diagonalise a matrix scales as $O(N^3)$, where N is the dimension of the (square) matrix. This is the underlying problem which limits the number of atoms that can be considered and therefore even the types of molecules studied. This is a very real issue for our study of capped carbon nanotubes as we wish to include molecules with $N_{\text{atom}} \approx 1000$ (and

therefore $N \approx 4000$).

In light of this limitation on molecule size, there has been considerable recent interest in linear scaling methods.[55,92–100] This chapter focuses on three schemes utilising the density matrix. These methods avoid diagonalising the Hamiltonian matrix by expressing the energy of the molecule in terms of the density matrix and achieve $O(N)$ scaling by truncating the density matrix. Since the density matrix is a vital component of these methods I shall begin by reviewing its properties in Section 4.1, before discussing the method proposed by Li, Nunes and Vanderbilt (LNV)[94] in Section 4.2 and the Palser-Manolopoulos (PM) method[96] in Section 4.3. The hybrid method suggested by Bowler and Gillan,[97] which is a combination of the two earlier methods, is then considered. The linear scaling of all three methods is demonstrated in tight-binding calculations on doubly capped carbon nanotubes. Finally, applications and limitations of density matrix methods are reviewed and the chapter ends with a short summary.

4.1 Density matrix

4.1.1 Notation

Throughout this chapter we shall consider an orthogonal tight-binding Hamiltonian in a basis of N orthonormal atomic orbitals. Each molecular orbital $|\psi_n\rangle$ is expanded as a linear combination of these atomic orbitals,

$$|\psi_n\rangle = \sum_{i=1}^{N} |\phi_i\rangle \, c_{in}. \tag{4.1}$$

The expansion coefficients c_{in} and the molecular orbital energy levels ϵ_n are obtained by solving the secular equations:

$$\sum_{j=1}^{N} H_{ij} c_{jn} = c_{in} \epsilon_n, \tag{4.2}$$

where $H_{ij} = \langle \phi_i | \hat{H} | \phi_j \rangle$ is an element of the $N \times N$ (real and symmetric) Hamiltonian matrix $\boldsymbol{H}$.

Since $\boldsymbol{H}$ is real and symmetric, its eigenvectors are orthonormal:

$$\sum_{i=1}^{N} c_{in} c_{im} = \delta_{nm}, \tag{4.3}$$

$$\sum_{n=1}^{N} c_{in} c_{jn} = \delta_{ij}, \tag{4.4}$$

and its eigenvalues ϵ_n are real. The band structure energy E is simply the sum of the occupied molecular orbital energy levels as given in Eqn. 2.4,

$$E = 2 \sum_{n=1}^{N_{\text{occ}}} \epsilon_n,$$

where N_{occ} is the number of occupied orbitals.

The elements ρ_{ij} of the density matrix $\boldsymbol{\rho}$ are defined by

$$\rho_{ij} = \sum_{n=1}^{N_{\text{occ}}} c_{in} c_{jn}. \tag{4.5}$$

The density matrix is useful since all the electronic properties of the molecule may be expressed in terms of it. Thus, the number of electrons N_e is given by

$$N_e = 2 \text{ tr} [\boldsymbol{\rho}], \tag{4.6}$$

and the energy is

$$E = 2 \text{ tr} [\boldsymbol{\rho H}]. \tag{4.7}$$

The proofs of these statements are trivial, and are based on the secular equations in Eqn. 4.2 and the orthonormality of eigenvectors in Eqns. 4.3 and 4.4:

$$2\ \mathrm{tr}\left[\boldsymbol{\rho}\right] = 2\sum_{i=1}^{N}\ \rho_{ii} = 2\sum_{i=1}^{N}\ \sum_{n=1}^{N_{\mathrm{occ}}}\ c_{in}c_{in} = 2\sum_{n=1}^{N_{\mathrm{occ}}}\ 1 = N_e, \tag{4.8}$$

and

$$\begin{aligned} 2\ \mathrm{tr}\left[\boldsymbol{\rho H}\right] &= 2\sum_{i=1}^{N}\ \sum_{j=1}^{N}\ H_{ij}\rho_{ji} = 2\sum_{i=1}^{N}\ \sum_{j=1}^{N}\ \sum_{n=1}^{N_{\mathrm{occ}}}\ \left(H_{ij}c_{jn}\right)c_{in} \\ &= 2\sum_{i=1}^{N}\ \sum_{n=1}^{N_{\mathrm{occ}}}\ c_{in}\epsilon_n c_{in} = 2\sum_{n=1}^{N_{\mathrm{occ}}}\ \epsilon_n = E. \end{aligned}$$

In addition, there are three properties of the density matrix which are important to linear scaling methods: the fact that it is idempotent; that it commutes with the Hamiltonian matrix $\boldsymbol{H}$ and that it is localised in real space.

4.1.2 Idempotency

The density matrix $\boldsymbol{\rho}$ has the following property,

$$\boldsymbol{\rho}^2 = \boldsymbol{\rho}, \tag{4.9}$$

that is, it is idempotent. This follows immediately from the definition of $\boldsymbol{\rho}$ in Eqn. 4.5, which gives

$$\left(\boldsymbol{\rho}^2\right)_{ij} = \sum_{k=1}^{N}\ \rho_{ik}\rho_{kj}$$

$$
\begin{aligned}
&= \sum_{k=1}^{N} \left(\sum_{n=1}^{N_{\mathrm{occ}}} c_{in} c_{kn} \right) \left(\sum_{m=1}^{N_{\mathrm{occ}}} c_{km} c_{jm} \right) \\
&= \sum_{n=1}^{N_{\mathrm{occ}}} \sum_{m=1}^{N_{\mathrm{occ}}} c_{in} \left(\sum_{k=1}^{N} c_{kn} c_{km} \right) c_{jm} \\
&= \sum_{n=1}^{N_{\mathrm{occ}}} \sum_{m=1}^{N_{\mathrm{occ}}} c_{in} \delta_{nm} c_{jm} \\
&= \left(\boldsymbol{\rho} \right)_{ij} .
\end{aligned}
$$

From the definition of $\boldsymbol{\rho}$ in Eqn. 4.5, it is clear that $\boldsymbol{\rho}$ is real and symmetric, and it follows that its eigenvalues are real. The fact that $\boldsymbol{\rho}$ is idempotent imposes a stringent constraint on these eigenvalues r_n, which must satisfy $r_n^2 = r_n$. This clearly implies that $r_n = 0$ or 1.

4.1.3 Commutation with $\boldsymbol{H}$

The density matrix also commutes with the Hamiltonian matrix $\boldsymbol{H}$, *i.e.*

$$
\boldsymbol{\rho H} = \boldsymbol{H \rho}. \tag{4.10}
$$

This can be shown by making use of Eqn. 4.2 as follows:

$$
\begin{aligned}
\left(\boldsymbol{\rho H} \right)_{ij} &= \sum_{k=1}^{N} \rho_{ik} H_{kj} \\
&= \sum_{k=1}^{N} \sum_{n=1}^{N_{\mathrm{occ}}} c_{in} c_{kn} H_{kj} \\
&= \sum_{n=1}^{N_{\mathrm{occ}}} c_{in} \left(\sum_{k=1}^{N} H_{jk} c_{kn} \right) \\
&= \sum_{n=1}^{N_{\mathrm{occ}}} c_{in} \epsilon_n c_{jn},
\end{aligned}
$$

but

$$
\begin{aligned}
(\boldsymbol{H}\boldsymbol{\rho})_{ij} &= \sum_{k=1}^{N} H_{ik}\rho_{kj} \\
&= \sum_{k=1}^{N}\sum_{n=1}^{N_{\rm occ}} H_{ik}c_{kn}c_{jn} \\
&= \sum_{n=1}^{N_{\rm occ}} c_{jn}\sum_{k=1}^{N} (H_{ik}c_{kn}) \\
&= \sum_{n=1}^{N_{\rm occ}} c_{jn}\epsilon_n c_{in} \\
&= (\boldsymbol{\rho}\boldsymbol{H})_{ij}\,,
\end{aligned}
$$

which is the desired result. One consequence of this is that the eigenvectors which diagonalise $\boldsymbol{H}$ will also diagonalise $\boldsymbol{\rho}$. Actually, this follows more directly from Eqns. 4.2, 4.3 and 4.4; which imply the following spectral representation of $\boldsymbol{H}$:

$$H_{ij} = \sum_{n=1}^{N} c_{in}\epsilon_n c_{jn}. \tag{4.11}$$

The corresponding spectral representation of $\boldsymbol{\rho}$ in Eqn. 4.5 is

$$\rho_{ij} = \sum_{n=1}^{N_{\rm occ}} c_{in}c_{jn} = \sum_{n=1}^{N} c_{in}r_n c_{jn}, \tag{4.12}$$

where $r_n = 1$ if $n \leq N_{\rm occ}$ (*i.e.* if molecular orbital n is occupied) and 0 otherwise (*i.e.* if molecular orbital n is unoccupied).

4.1.4 Localisation in real space

The density matrix is localised in real space. This means that as the distance R_{ij} between atomic orbitals i and j increases, the corresponding element of

the density matrix ρ_{ij} tends to zero. It is often said that in metals the elements of the density matrix decay according to the power law $R_{\alpha\beta}^{-d}$ (where d is the number of dimensions), whereas in insulators the decay is exponential.[101] However, this has only been proved for an isolated band in a 1-D crystal with inversion symmetry.[99] What happens in 3-D crystals with overlapping bands has not been fully examined, and also, the effect of defects in the structure, such as caps, is unclear.[93,99]

In view of this, we have examined the decay of the density matrix for some representative doubly capped carbon nanotubes. The elements of the density matrix were calculated in a single geometry direct diagonalisation calculation using the Ho Hamiltonian (the tube geometries having previously been optimised using the hybrid method described later in this chapter). $\boldsymbol{\rho}_{ij}$ are calculated as a function of the internuclear distance from an atom in the centre of the tube. There were 1040 atoms in the (5,5) tube, 1020 atoms in the (10,0) tube and 1340 atoms in the (11,2) tube.

Figure 4.1 shows the decay of the density matrix elements between two s, two p_x, two p_y and two p_z orbitals for a doubly capped (5,5) tube with C_{5v} caps, a doubly capped (10,0) tube with C_{5v} caps and a doubly capped (11,2) nanotube with $C_1(78)$ caps. In each case, the density matrix element quickly tends towards zero as the internuclear distance increases. Furthermore, the density matrix elements are actually remarkably similar for all the tubes. The differences between the matrix elements involving p orbitals are due to the different orientations of these orbitals. Figure 4.2 shows the decay of the density matrix elements in more detail.

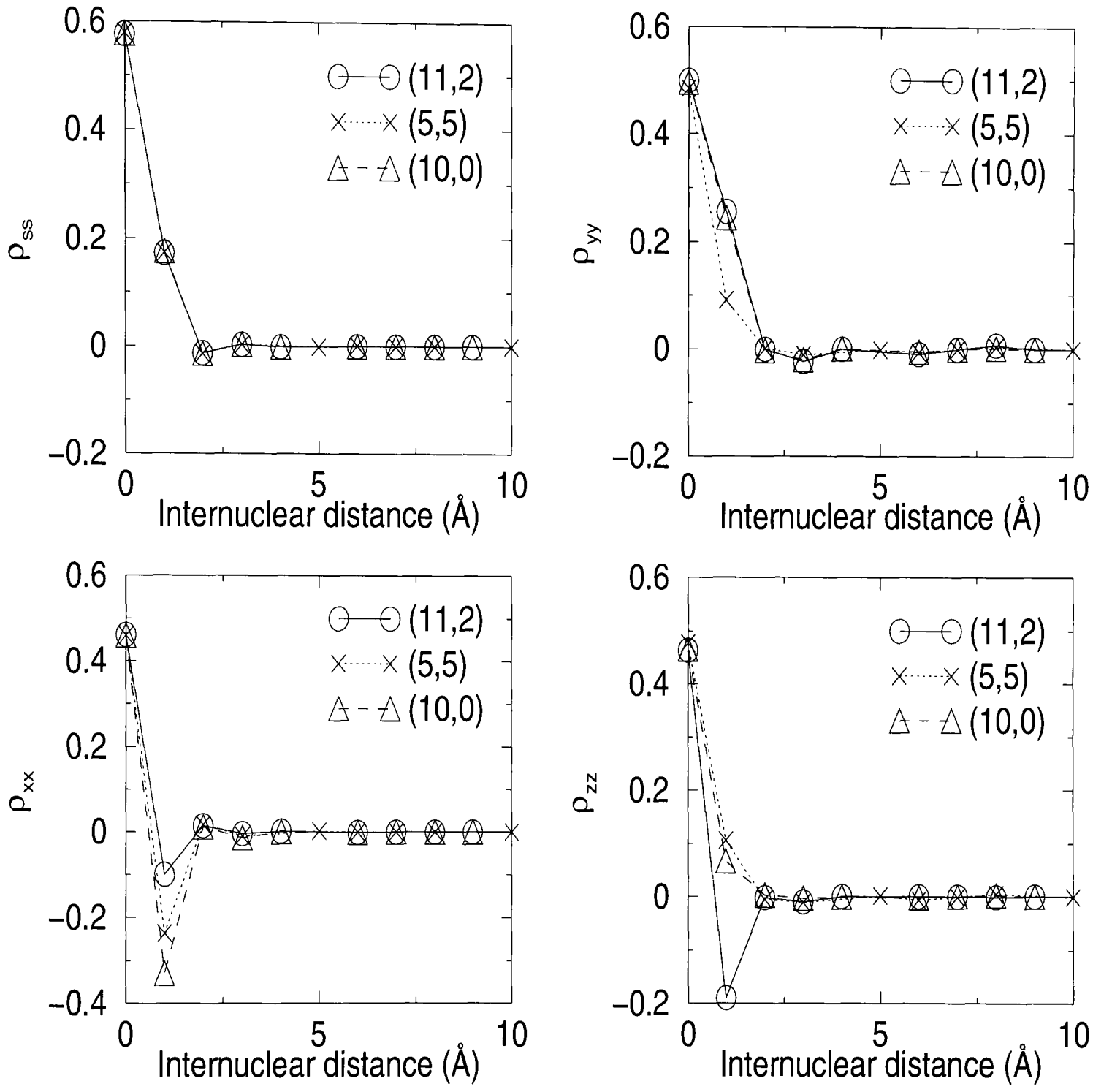


Figure 4.1: Elements of the density matrix for the (5,5) tube capped with the C_{5v} cap, the (10,0) tube capped with the C_{5v} cap and the (11,2) tube capped with cap 78.

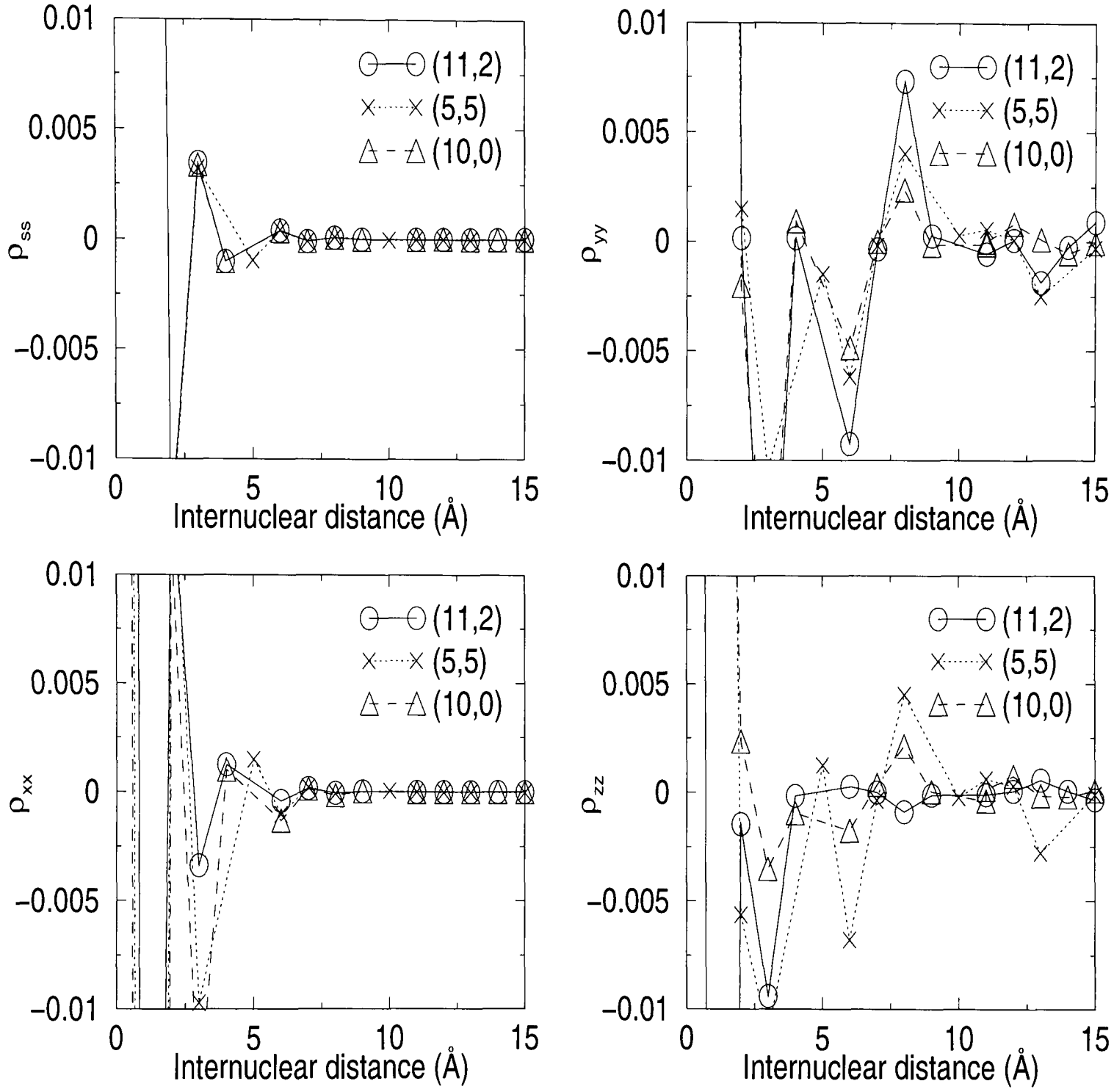


Figure 4.2: Enlarged view of the elements of the density matrix for the (5,5) tube capped with the C_{5v} cap, the (10,0) tube capped with the C_{5v} cap and the (11,2) tube capped with cap 78.

Another way of investigating the decay of the density matrix is to consider the magnitude of the density matrix element[93] $|\rho_{\alpha\beta}|$ between atoms α and β as defined by Nunes and Vanderbilt:[95]

$$|\rho_{\alpha\beta}| = \sqrt{\frac{\sum_{i\in\alpha, j\in\beta} |\rho_{ij}|^2}{n_{\text{orb}}}}, \tag{4.13}$$

where $n_{\text{orb}} = 4$ is the number of orbitals per atom. This quantity is shown as a function of internuclear separation in Figure 4.3 for the three tubes considered above. The decay of the density matrix is very similar for all the tubes examined, and is found to be very rapid. Small oscillations do persist to quite large distances, however.

Bowler *et al.*[93] have also examined the magnitude of the density matrix elements for carbon and silicon in the diamond structure and hexagonally close packed titanium as a function of the internuclear separation. They found that $|\rho_{\alpha\beta}|$ died away extremely quickly for carbon, but more slowly for silicon. However, even for silicon, the fluctuations in the density matrix were insignificant for internuclear separations greater than about three bondlengths. The density matrix elements for titanium did not vanish so soon, and Bowler *et al.* suggest that a finite electronic temperature may be necessary to localise the density matrix for metals.[93]

Lest the titanium result seem to disagree with our calculated density matrix elements for the (5,5) tube (Figures 4.1 and 4.2), which is metallic, or cast doubt on the wisdom of using this method for (metallic) capped nanotubes, it should be noted that a small band gap opens for finite "metallic" nanotubes[102–105] (This will be discussed in more detail in Chapter 5). This is

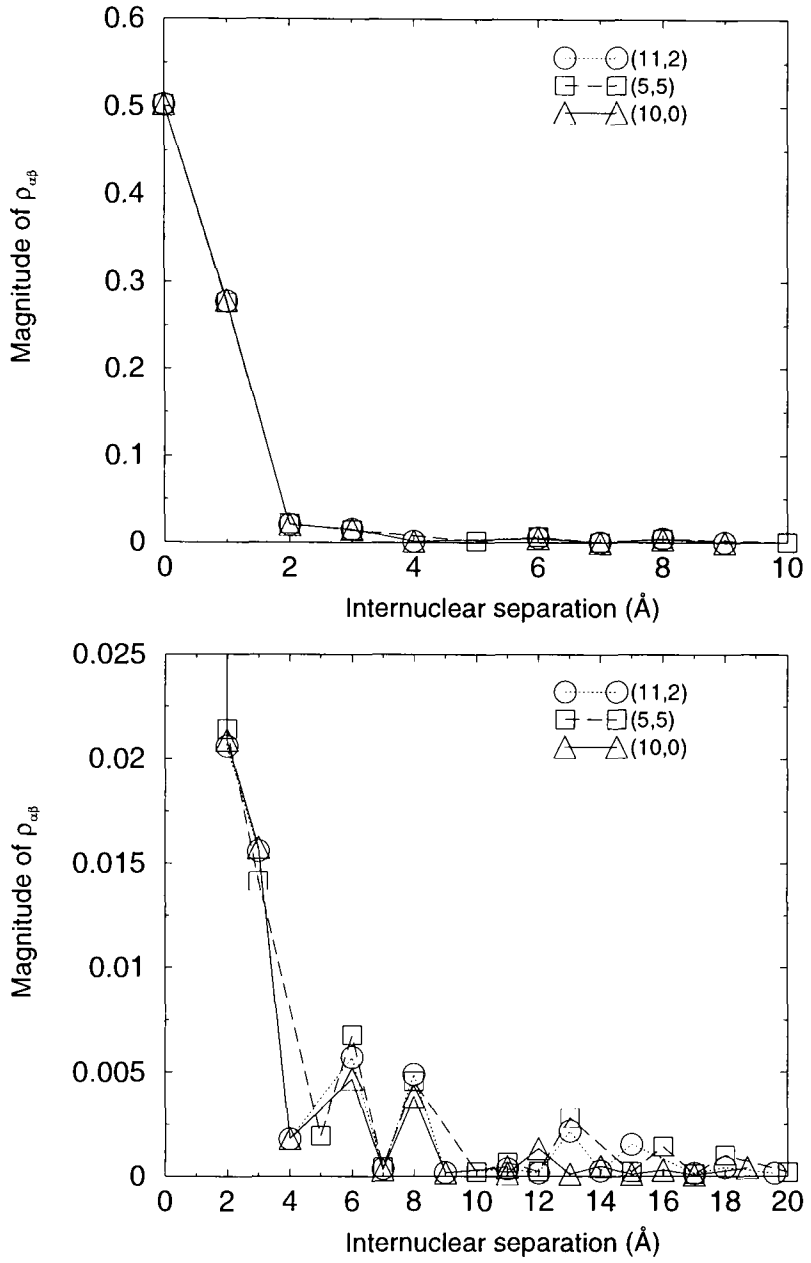


Figure 4.3: Magnitude of the density matrix elements $|\rho_{\alpha\beta}|$ between atoms α and β as a function of the internuclear separation $R_{\alpha\beta}$.

actually rather convenient for our purposes as it means that density matrix methods are more efficient and more accurate for finite nanotubes than would be expected to be in the limit of an infinite metallic tube.

4.2 The LNV method

This method was proposed by Li, Nunes and Vanderbilt in 1993[94] and as such was the first linear scaling density matrix method to be put forward (although at around the same time an equivalent method was independently suggested by Daw[106]). The method was originally proposed for orthogonal tight-binding Hamiltonians and, although it was subsequently extended to the non-orthogonal case,[95] this is what we shall focus on here.

4.2.1 Grand canonical potential

The electronic energy of a molecule can be expressed, in terms of its density matrix, as Eqn. 4.7,

$$E = 2 \text{ tr} \left[\boldsymbol{\rho} \boldsymbol{H} \right].$$

In our carbon nanotube applications we desire a fixed Fermi level μ rather than a constant number of electrons N_e, and it is therefore more convenient to work with the grand potential

$$\Omega = E - \mu N_e = 2 \text{ tr} \left[\boldsymbol{\rho} \left(\boldsymbol{H} - \mu \boldsymbol{I} \right) \right] = 2 \text{ tr} \left[\boldsymbol{\rho} \boldsymbol{H}' \right]. \tag{4.14}$$

This has the effect of shifting the zero of energy to the Fermi level and hence occupied levels have negative energy and unoccupied levels have positive energy.

4.2.2 Finding the density matrix

Given the above expression for calculating the grand potential Ω in terms of the density matrix $\boldsymbol{\rho}$, all that is needed is a method for calculating the density matrix itself. However, this is not completely straightforward because the definition of $\boldsymbol{\rho}$ given in Eqn. 4.5 involves the eigenvectors of the Hamiltonian, which can only be obtained by diagonalisation. The heart of the LNV method is to take an initial guess for the density matrix and then variationally minimise the grand potential with respect to the elements of the density matrix rather than obtaining $\boldsymbol{\rho}$ by diagonalisation. On its own, however, this does not place any constraints on the density matrix, and yet $\boldsymbol{\rho}$ must be idempotent. Finding the eigenvalues of $\boldsymbol{\rho}$, and hence determining whether it is idempotent or not, would necessitate diagonalising the matrix which is the very process we wish to avoid.

The solution to this problem is the McWeeny purification transformation,[107]

$$\boldsymbol{\rho} = 3\boldsymbol{\sigma}^2 - 2\boldsymbol{\sigma}^3. \tag{4.15}$$

Eqn. 4.15 is used to produce a more nearly idempotent matrix $\boldsymbol{\rho}$ from a nearly idempotent matrix $\boldsymbol{\sigma}$. In order to show that the McWeeny transformation produces a more nearly idempotent matrix, consider a nearly idempotent matrix $\boldsymbol{\sigma}$ with eigenvalues $s_n = 1 + \delta$ or δ. Since $\boldsymbol{\sigma}$ and $\boldsymbol{\rho}$ commute (because

$\boldsymbol{\rho}$ is a polynomial in $\boldsymbol{\sigma}$), it is possible to work in a basis where they are both diagonal. When $s_n = 1 + \delta$, the corresponding eigenvalue r_n of $\boldsymbol{\rho}$ is given by Eqn. 4.15,

$$\begin{aligned} r_n &= 3(1+\delta)^2 - 2(1+\delta)^3 \\ &= 1 - 3\delta^2 - 2\delta^3 \\ &= 1 + O(\delta^2), \end{aligned} \tag{4.16}$$

Thus the idempotency error in r_n is $O(\delta^2)$. It is also trivial to see that $r_n = O(\delta^2)$ when $s_n = \delta$.

Li, Nunes and Vanderbilt combined this idea with the minimisation of Ω in Eqn. 4.14 by writing

$$\Omega = 2 \text{ tr}\left[\boldsymbol{\rho}\boldsymbol{H}'\right] = 2 \text{ tr}\left[\left(3\boldsymbol{\sigma}^2 - 2\boldsymbol{\sigma}^3\right)\boldsymbol{H}'\right], \tag{4.17}$$

and then minimising Ω with respect to the elements of $\boldsymbol{\sigma}$ as variational parameters rather than the elements of $\boldsymbol{\rho}$.[94]

The reason why this works is as follows. Consider once more the eigenvalues of $\boldsymbol{\rho}$,

$$r_n = 3s_n^2 - 2s_n^3,$$

and the function $f(x) = 3x^2 - 2x^3$ shown in Figure 4.4. Now, if s_n is between $-\frac{1}{2}$ and $\frac{3}{2}$, then r_n will lie between 0 and 1. Furthermore, if $-\frac{1}{2} < s_n < \frac{1}{2}$, then r_n will be closer to zero than s_n, whereas if $\frac{1}{2} < s_n < \frac{3}{2}$, then r_n will be closer to one than s_n. Hence, successive applications of the McWeeny transformation will drive the eigenvalues towards zero or one.

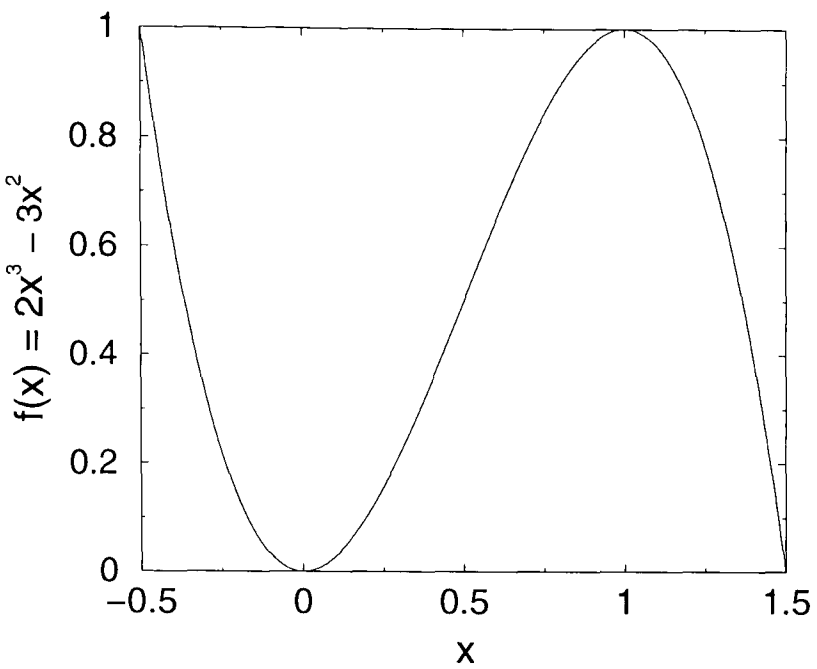


Figure 4.4: The function $f(x) = 3x^2 - 2x^3$.

In other words, the McWeeny purification ensures that the eigenvalues of $\boldsymbol{\rho}$ will lie close to 0 and 1 without explicitly imposing this condition. The initial guess does not even have to be (nearly) idempotent, it merely needs to have eigenvalues in the range $-\frac{1}{2}$ to $\frac{3}{2}$. However, the LNV method will only give the correct solution if the eigenvalues of $\boldsymbol{\sigma}$ remain between $-\frac{1}{2}$ and $\frac{3}{2}$ throughout the minimisation of Ω in Eqn. 4.17; if an eigenvalue leaves this range the McWeeny transformation will diverge to give unphysical solutions with $r_n \to \pm\infty$.

In practice, it is usual to start the LNV iteration with $\boldsymbol{\sigma} = \frac{1}{2}\boldsymbol{I}$. However, in some cases this will not be very like the "true" density matrix which will result in an increased number of iterations.

4.2.3 Minimisation

The minimisation of Ω in Eqn. 4.17 with respect to the elements of $\boldsymbol{\sigma}$ can be performed using standard techniques such as the steepest descent or conjugate gradient methods.[69] The conjugate gradient method is often chosen because it is more efficient.[69] Both methods can be modified to exploit the fact that Ω is a cubic function of $\boldsymbol{\sigma}$.[68] Each of these minimisation techniques also requires the derivatives of Ω with respect to the elements of $\boldsymbol{\sigma}$, which are found by differentiating Eqn. 4.17:

$$\frac{\mathrm{d}\Omega}{\mathrm{d}\sigma_{ij}} = 2\left[3(\boldsymbol{\sigma}\boldsymbol{H}' + \boldsymbol{H}'\boldsymbol{\sigma})_{ji} - 2(\boldsymbol{\sigma}^2\boldsymbol{H}' + \boldsymbol{\sigma}\boldsymbol{H}'\boldsymbol{\sigma} + \boldsymbol{H}'\boldsymbol{\sigma}^2)_{ji}\right]. \qquad (4.18)$$

4.2.4 Cut-off distance

Whilst the LNV method as described above will work, it is still does not scale linearly with N_{atom}. Recall that in Section 4.1.4 it was stated that the density matrix is localised in real space. This means that most of the matrix elements σ_{ij} are almost zero. This motivates the introduction of a cut-off radius parameter R_c such that σ_{ij} is assumed to be exactly zero if the distance $R_{\alpha\beta}$ between the two atoms α and β (with orbital i on atom α and orbital j on atom β) is greater than R_c. This is the central approximation in this method and in the other strategies based on the density matrix. The grand potential is minimised with respect to the remaining non-zero elements of the density matrix, the number of which increases linearly with the size of the system. Sparse matrix algebra may be used to obtain linear scaling and reduce the memory requirements. The other matrices which are used in this

algorithm are also subject to an appropriate cut-off distance, for example R_H for $\boldsymbol{H}$ and $R_c + R_H$ for $\boldsymbol{\rho H}$. In this work, specially adapted sparse matrix routines were used. These took advantage of the fact that, for example, $\boldsymbol{H}$ is symmetric to reduce the storage requirements still further and also fully utilised the fact that all the orbitals on atom α are at the same distance from all the orbitals on atom β to speed up the program.

4.2.5 Forces

For molecular dynamics calculations or geometry optimisations of a molecule, calculating the electronic energy alone is not enough; the atomic forces are also required.

Since the LNV method is variational it is possible to use the Hellmann-Feynman theorem[59,60] to calculate the derivatives of the energy with respect to some parameter ζ, for example, an atomic coordinate. Differentiating the expression for Ω in Eqn. 4.17 with respect to ζ yields

$$\frac{\mathrm{d}\Omega}{\mathrm{d}\zeta} = \frac{\partial\Omega}{\partial\boldsymbol{\sigma}}\frac{\partial\boldsymbol{\sigma}}{\partial\zeta} + \frac{\partial\Omega}{\partial\boldsymbol{H'}}\frac{\partial\boldsymbol{H'}}{\partial\zeta}, \tag{4.19}$$

but since the LNV minimisation is variational, $\partial\Omega/\partial\boldsymbol{\sigma} = 0$ at the minimum, so Eqn. 4.19 becomes

$$\frac{\mathrm{d}\Omega}{\mathrm{d}\zeta} = \frac{\partial\Omega}{\partial\boldsymbol{H'}}\frac{\partial\boldsymbol{H'}}{\partial\zeta} \tag{4.20}$$

$$= 2\ \mathrm{tr}\left[\boldsymbol{\rho}\frac{\partial\boldsymbol{H'}}{\partial\zeta}\right], \tag{4.21}$$

which is the Hellmann-Feynman expression.[59,60] This is an important result because it means that the calculated forces in this method are exactly consistent with changes in Ω.

4.2.6 LNV conclusions

Before leaving the LNV method there are several points to note. If it works correctly, the grand canonical potential found by the LNV method is a variational upper bound on the exact potential, but this solution is not the global minimum of Ω in Eqn. 4.17. Unphysical "runaway" solutions are also possible in which the eigenvalues of the density matrix tend towards $\pm\infty$.

Putting these unphysical solutions aside, the only approximation (within the TB formalism) in this method is truncating the density matrix. The method therefore becomes exact as $R_c \to \infty$.

4.3 The PM method

Palser and Manolopoulos[96] have also proposed a linear scaling method based on the density matrix. The PM method, which is much simpler than the LNV method, uses the McWeeny transformation iteratively. The crucial new twist added by Palser and Manolopoulos is a prescription for an appropriate guess for the initial density matrix.

4.3.1 Using the McWeeny transformation iteratively

The McWeeny transformation given in Eqn. 4.15 was originally conceived as an iterative scheme for generating an idempotent matrix from an approximately idempotent matrix.[107] Eqn. 4.15 is rewritten as

$$\boldsymbol{\rho}_{k+1} = 3\boldsymbol{\rho}_k^2 - 2\boldsymbol{\rho}_k^3, \tag{4.22}$$

and it follows from Section 4.2.2 that if the initial guess $\boldsymbol{\rho}_0$ for the density matrix is nearly idempotent, then the density matrix $\boldsymbol{\rho}_1$ at the next iteration will be more nearly idempotent (as shown in Section 4.2.2) and so on. It also follows from the arguments in Section 4.2.2 that the McWeeny transformation used iteratively will eventually generate an exactly idempotent matrix provided that the eigenvalues of $\boldsymbol{\rho}_0$ are between $-\frac{1}{2}$ and $\frac{3}{2}$. Therefore, Eqn. 4.22 can be used iteratively to obtain the ground state density matrix, if a suitable starting guess is used. This leaves us with two remaining issues: an initial guess for the density matrix and the criterion for terminating the iterative process.

4.3.2 Initial guess for the density matrix

The true density matrix of the system is defined by the fact that it commutes with $\boldsymbol{H}$ and that it is idempotent which means that its eigenvalues are either 1 (occupied orbitals) or 0 (unoccupied orbitals). Since $\boldsymbol{\rho}_k$ is a polynomial in $\boldsymbol{\rho}_0$, by virtue of Eqn. 4.22, it follows that if $\boldsymbol{\rho}_0$ commutes with $\boldsymbol{H}$, then so will the final density matrix. This can be achieved by choosing $\boldsymbol{\rho}_0$ to be a linear function of $\boldsymbol{H}$, which leads to $\boldsymbol{\rho}_k$ being a polynomial in $\boldsymbol{H}$ of order $3k$. Also, we know that $\boldsymbol{\rho}_k$ will be idempotent if the eigenvalues of $\boldsymbol{\rho}_0$ are between $-\frac{1}{2}$ and $\frac{3}{2}$. The following initial guess, suggested by Palser and Manolopoulos, satisfies both these conditions:

$$\boldsymbol{\rho}_0 = \frac{\xi}{2}(\mu \boldsymbol{I} - \boldsymbol{H}) + \frac{1}{2}\boldsymbol{I}, \tag{4.23}$$

where

$$\xi = \min\left\{\frac{1}{H_{\text{max}} - \mu}, \frac{1}{\mu - H_{\text{min}}}\right\}, \tag{4.24}$$

and H_{max} (H_{min}) is the upper (lower) bound on the eigenvalue spectrum of $\boldsymbol{H}$.

However, this leads to another problem: how can we find the upper and lower bounds of the eigenvalues of the Hamiltonian matrix when diagonalising this matrix is the very thing we wish to avoid? Fortunately, reasonable approximations for H_{max} and H_{min} may be obtained using Gershgorin's formulas,[96]

$$H_{\text{min}} = \min_i \left\{ H_{ii} - \sum_{j \neq i} |H_{ij}| \right\}, \tag{4.25}$$

$$H_{\text{max}} = \max_i \left\{ H_{ii} + \sum_{j \neq i} |H_{ij}| \right\}, \tag{4.26}$$

which can be evaluated in $O(N)$ operations for a typical (sparse) tight-binding Hamiltonian matrix.

This guess for $\boldsymbol{\rho}_0$ will give the correct ground state density matrix when the calculation is done exactly. However, the exact calculation does not scale linearly. As with the LNV method, linear scaling is achieved by truncating the density matrix and using sparse matrix algebra as discussed above.

4.3.3 Termination criterion

At the k^{th} iteration of the McWeeny transformation, the corresponding value of the grand canonical potential Ω_k is given by

$$\Omega_k = 2 \text{ tr} \left[\boldsymbol{\rho}_k \boldsymbol{H}' \right]. \tag{4.27}$$

If the iteration is performed without truncating the density matrix, it can be shown that Ω converges monotonically *i.e.* $\Omega_{k+1} \leq \Omega_k$. The argument that

establishes this is as follows. Since $\boldsymbol{\rho}_k$ will be diagonal in any basis in which $\boldsymbol{H}$ is diagonal (because they commute), it is convenient to use such a basis. If the eigenvalues of $\boldsymbol{H}$ are arranged in ascending order

$$H_{\min} \leq \epsilon_1 \leq \cdots \leq \epsilon_j \leq \epsilon_{j+1} \leq \cdots \leq \epsilon_N \leq H_{\max},$$

then the eigenvalues $r_{j,0}$ of $\boldsymbol{\rho}_0$ in Eqn. 4.23 will be arranged in descending order, and (since $H_{\min} \leq \mu \leq H_{\max}$) they will be bounded by 0 and 1:

$$1 \geq r_{1,0} \geq \cdots \geq r_{j,0} \geq r_{j+1,0} \geq \cdots \geq r_{N,0} \geq 0.$$

Moreover, subsequent iterations do not change the order of the eigenvalues because the iteration monotonically maps $\boldsymbol{\rho}_k$ onto $\boldsymbol{\rho}_{k+1}$ when $0 \leq r_n \leq 1$ as can be seen in Figure 4.4. So, after the k^{th} iteration,

$$1 \geq r_{1,k} \geq \cdots \geq r_{j,k} \geq r_{j+1,k} \geq \cdots \geq r_{N,k} \geq 0.$$

If the j^{th} eigenvalue of $\boldsymbol{\rho}_k$, $r_{j,k}$, is greater than, or equal to, $\frac{1}{2}$, then the j^{th} eigenvalue of $\boldsymbol{\rho}_{k+1}$ will be greater than or equal to $r_{j,k}$. Similarly if $r_{j,k} \leq \frac{1}{2}$, then $r_{j,k+1} \leq r_{j,k}$. Furthermore, if $\epsilon_j > \mu$, Eqn. 4.23 implies that $r_{j,0,\leq\frac{1}{2}}$ and, in the same way, if $\epsilon_j < \mu$, then $r_{j,0,\geq\frac{1}{2}}$. Therefore,

$$\sum_k r_{j,k+1}\left(\epsilon_j - \mu\right) \;\leq\; \sum_k r_{j,k}\left(\epsilon_j - \mu\right), \tag{4.28}$$

where $k = 0, 1, 2, \ldots$, which implies that $\Omega_{k+1} \leq \Omega_k$ *i.e.* convergence is monotonic. This proof of monotonic convergence only holds when the calculation is exact. This fact can be used to provide a termination criterion for the McWeeny iteration. For a calculation with a truncated density matrix, Palser and Manolopoulos suggest that it is logical to assume that truncation error is dominating when $\Omega_{k+1} \geq \Omega_k$, and therefore to stop the iteration at this stage.

4.3.4 PM conclusions

The PM method has a number of advantages over the LNV method: the starting guess for the density matrix is guaranteed to work, and the PM method is faster than the LNV method, for two reasons. First, the algorithm itself is simpler than the LNV recipe, and fewer (sparse) matrix multiplications need be performed per iteration. Secondly, the PM method converges faster. The convergence of the LNV method is determined by the convergence of the minimisation method which, for the conjugate gradient method, is linear (*i.e.* the error decreases linearly with each iteration). However, the PM method converges quadratically. In order to see this, consider the $(k+1)^{\text{th}}$ iteration where $\boldsymbol{\rho}_k = \boldsymbol{\rho} + \boldsymbol{\delta}_k$ in a basis where $\boldsymbol{\rho}_k$ and $\boldsymbol{\rho}$ are diagonal (which implies that $\boldsymbol{\delta}_k$ is also diagonal). We have:

$$\begin{aligned} \boldsymbol{\rho}_{k+1} &= 3\boldsymbol{\rho}_k^2 - 2\boldsymbol{\rho}_k^3 \\ &= 3\left(\boldsymbol{\rho} + \boldsymbol{\delta}_k\right)^2 - 2\left(\boldsymbol{\rho} + \boldsymbol{\delta}_k\right)^3 \\ &= \boldsymbol{\rho} + \left(3\boldsymbol{I} - 6\boldsymbol{\rho}\right)\boldsymbol{\delta}_k^2 - 2\boldsymbol{\delta}_k^3, \end{aligned} \tag{4.29}$$

and therefore,

$$\boldsymbol{\rho}_{k+1} = \boldsymbol{\rho} + O(\boldsymbol{\delta}_k^2). \tag{4.30}$$

Therefore, not only is each PM iteration faster than each LNV iteration, but the PM method generally requires fewer iterations than the LNV method.

However, the PM method is not without its disadvantages. The PM method is not variational with respect to the elements of the purified density matrix $\boldsymbol{\rho}_k$, which means that the Hellmann-Feynmann expression is not valid and the forces derived in this way will not be exactly consistent with changes in

Ω. This has serious implications for the use of method in number of applications, such as molecular dynamics simulations and optimising the geometry (and, therefore, energy) of molecules, where forces are an integral part of the calculation.

Finally, it should be noted that although this discussion has focussed exclusively on the fixed Fermi level (μ) calculation, it is possible to reformulate the algorithm and the starting guess for a fixed electron count N_e.[96] The resulting algorithm is more complicated than that presented here, but is more useful in applications where the Fermi level is either unknown or difficult to calculate. Both the fixed Fermi level and the fixed electron count algorithms may be adapted for non-orthogonal Hamiltonians.[96]

4.4 The hybrid method

In the previous two sections I have described two density matrix methods. Neither of these methods is perfect and both have strengths and weaknesses, as summarised in Table 4.1. The only entry in this table which needs further comment is that concerning the initial density matrix. The fact that the LNV method does not have a prescription for the starting guess can be a disadvantage because an unfortunate initial choice of $\boldsymbol{\sigma}$ will lead to unphysical solutions. However, the freedom to choose any initial density matrix can be an advantage in, for example, molecular dynamics simulations. This is the case if the density matrix at each timestep is very similar to the density matrix from the previous timestep. In these circumstances it is more efficient to start with the density matrix from the previous timestep.[68] The other en-

LNV method	PM method
Variational	Not variational
Unphysical solutions possible	No unphysical solutions
More matrix operations per iteration	Fewer matrix operations per iteration
Linear convergence	Quadratic convergence
Arbitrary initial density matrix $\boldsymbol{\sigma}$	Prescribed initial density matrix $\boldsymbol{\rho}_0$

Table 4.1: Comparison of the advantages and disadvantages of the LNV and PM methods.

tries in Table 4.1 show quite clearly that the drawbacks of the LNV method are compensated for by the PM method and vice versa. Thus a strategy that could combine the two methods, retaining the best features of each, would be highly desirable.

Bowler and Gillan have suggested just such a strategy for combining the PM and LNV methods.[97] They proposed using the PM method to generate a good approximation to the density matrix and then refining $\boldsymbol{\rho}$, thereby improving the estimate of Ω, with the LNV method. Since the PM method should have done most of the optimising of $\boldsymbol{\rho}$, this method should be fast, but as the LNV method finishes the process, the overall scheme is variational, allowing the forces to be calculated using Hellmann-Feynman theory (Eqn. 4.20).

Furthermore, it makes sense to do this because the PM method is essentially very good at finding idempotent matrices whereas the LNV method is good at finding *the* idempotent matrix *i.e.* the density matrix of the system.

Therefore, the PM method should generate an idempotent matrix (quickly) and the LNV method should cycle through (all) the possible idempotent matrices to find the one corresponding to the "true" ground state. Bowler and Gillan have cast these arguments in more mathematical language and we shall examine this next.

4.4.1 Connections between the LNV and PM methods

In order to take this discussion further, it is useful (following Bowler and Gillan[97]) to view matrices as points in vector space. Within this vector space it is possible to define geometrical concepts such as straight lines. For example, the straight line between two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ can be defined as the set of matrices $(1-\lambda)\,\boldsymbol{A} + \lambda \boldsymbol{B}$, where $0 < \lambda < 1$. Thus, the sequence of matrices $\boldsymbol{\rho}_0, \boldsymbol{\rho}_1, \cdots, \boldsymbol{\rho}_k, \cdots, \boldsymbol{\rho}_\infty$ which is generated by repeated application of the McWeeny iteration in the PM method, can be thought of as a series of points in vector space which can be joined by lines to give the PM path. The limit of this series $\boldsymbol{\rho}_\infty$ is idempotent and it lies on what is known as the idempotency surface, a region consisting of all the idempotent matrices.

Further, it is possible to define tangent planes to the idempotency surface. Now, for any point on the idempotency surface, $\boldsymbol{P} = \boldsymbol{P}^2$. Consider a straight line $\boldsymbol{P}' = \boldsymbol{P} + \alpha \boldsymbol{B}$ where $\boldsymbol{B}$ is a Hermitian matrix and α is a real (scalar) variable. In general, $\boldsymbol{P}'$ is not idempotent, but there will be some matrices $\boldsymbol{B}$ for which $\boldsymbol{P}'$ is idempotent to linear order *i.e.* $\boldsymbol{P}'^2 - \boldsymbol{P}' = \alpha^2 \boldsymbol{B}^2$. In this case, the shortest distance between $\boldsymbol{P}'$ and the idempotency surface is of order α^2, so the straight line can be considered to be a tangent line to the idempotency

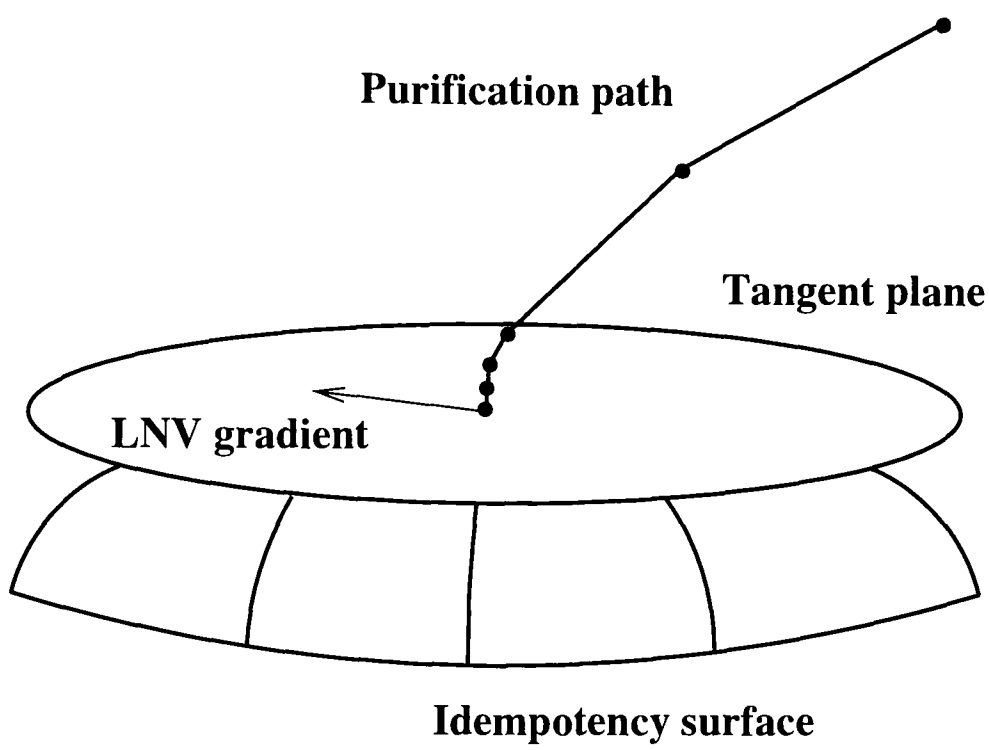


Figure 4.5: A geometrical representation of the PM (purification path) and the LNV gradient taken from Bowler and Gillan's paper.[97]

surface. The tangent plane to the idempotency surface at point $\boldsymbol{P}$ consists of all the matrices $\boldsymbol{P}'$ on all the tangent lines passing through $\boldsymbol{P}$.

These concepts were used by Bowler and Gillan to prove the following:

1. All PM paths meet the idempotency surface orthogonally;

2. For any point on the idempotency surface, the gradient of the LNV function is tangential to the surface.

These statements are justified in Appendix A, but illustrated pictorially in Figure 4.5. Taken together, these statements mean that the LNV gradient is orthogonal to the PM path. Additionally, the first statement implies that the PM convergence in the final stages will be extremely fast and direct. The

second indicates that the LNV iteration will not interfere with the PM iteration because (to first order) it remains on the idempotency surface. Since the LNV iteration is constrained to the idempotency surface, there is no possibility of reaching the unphysical "runaway" solutions. Hence this method is robust.

The discussion thus far has focussed on the exact calculation. Truncating the density matrix will affect both the PM and the LNV methods. The PM iteration will no longer result in an exactly idempotent matrix, but the matrix obtained should be very nearly idempotent. Assuming it is very close to the idempotency surface, there will still be no danger of the LNV iteration giving unphysical solutions. The convergence of the LNV iteration is expected to be reasonably fast since the matrix obtained by the PM iteration should be close to the ground state density matrix.

4.4.2 Hybrid conclusions

The central idea of the hybrid method is to use the PM method followed by the LNV method, hence overcoming the weaknesses of the individual methods without losing their desirable properties. In particular, the method is robust, linear scaling and variational. The method is expected to get more efficient as R_c increases because more of the work will be done by the (faster) PM method.

Fixed electron count calculations are also possible using the hybrid method. The PM fixed electron method is used, followed by the LNV method. Even though the LNV method is a calculation at fixed Fermi level, this method

should conserve the electron count because the PM method will generate an approximately idempotent matrix. Since the electron count is constant on the idempotency surface, it will be roughly constant near the surface.

Finally, the arguments used to underpin the hybrid method may be generalised to non-orthogonal tight-binding theory.[97]

4.5 Demonstration of linear scaling

This section discusses the results for some real tight-binding calculations on doubly capped carbon nanotubes. In particular, these studies examine the relative speed of the three methods discussed in this chapter. These calculations were performed using the orthogonal Ho tight-binding potential[41] discussed in Chapter 2. These calculations were performed on (5,5) nanotubes doubly capped with the C_{5v} cap. The length of these nanotubes, and hence the size of the system N_{atom}, was varied between calculations. The Fermi level for the nanotubes was set at that for the infinite tube.

The two factors determining the speed of a single geometry calculation using any of the three methods discussed in this chapter are the number of atoms in the molecule and the value of R_c. Increasing either N_{atom} or R_c will increase the time taken as shown in Figures 4.6 and 4.7. These figures clearly demonstrate the linear scaling of the LNV, PM and hybrid methods with the number of atoms. All the methods show a large increase in computational effort between $R_c = 6$ Å and $R_c = 7$ Å. This can be ascribed to the fact that the diameter of this particular tube is between 6 and 7 Å.

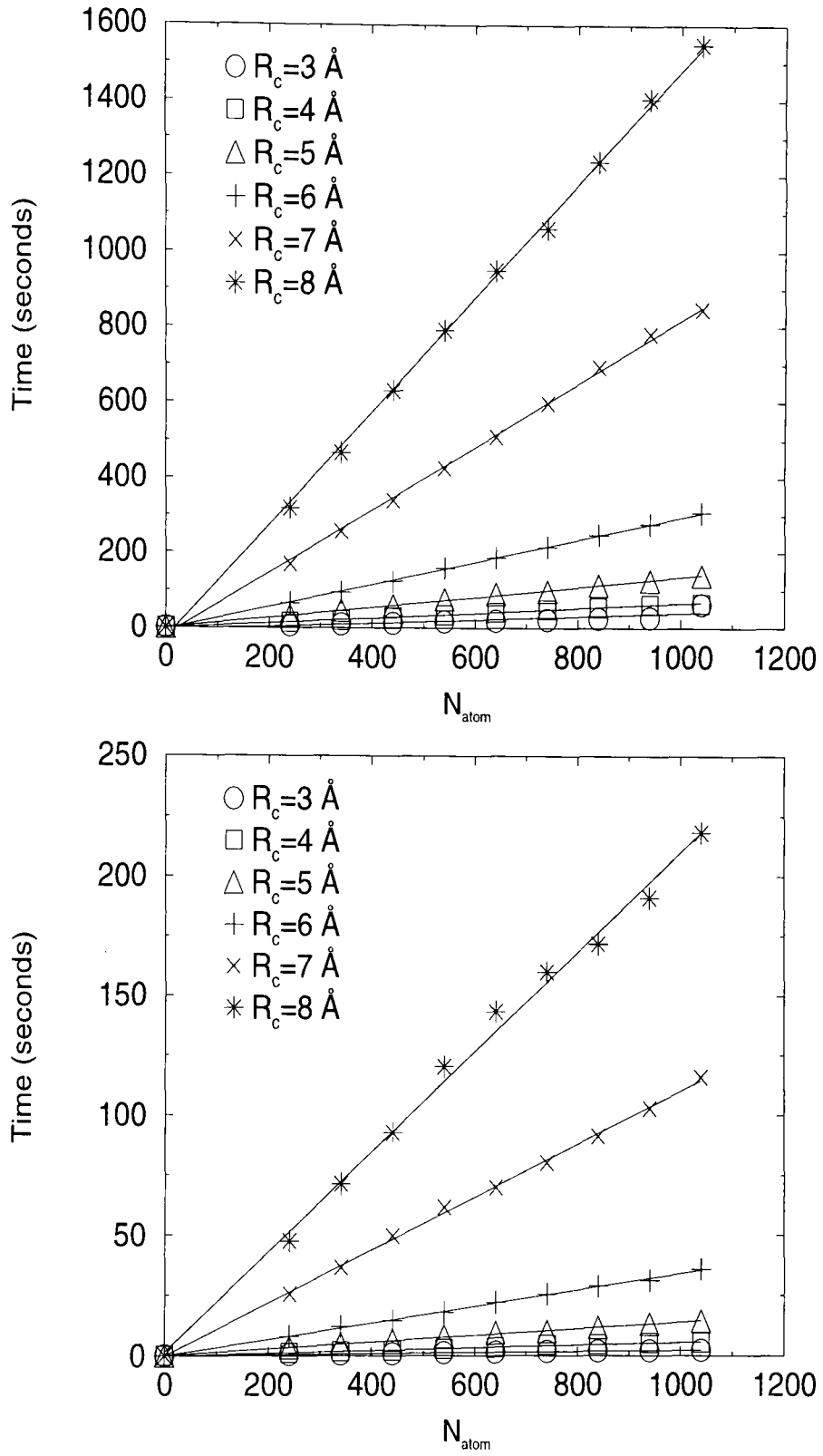


Figure 4.6: Time taken for single geometry LNV (top) and PM (bottom) calculations as a function of system size with a variety of cut-off distances. The results for each cut-off parameter are fitted by linear regression.

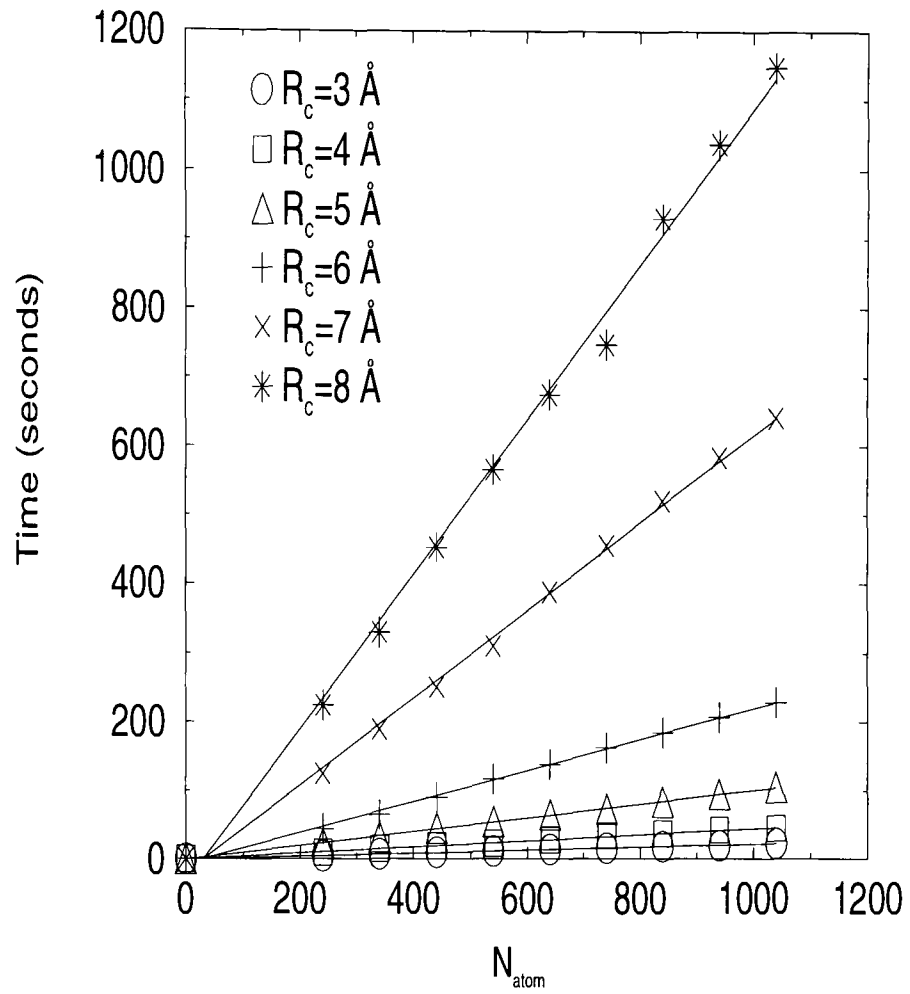


Figure 4.7: Time taken for a single geometry hybrid calculation as a function of system size with a variety of cut-off distances. The results for each cut-off parameter are fitted by linear regression.

The results for the LNV, PM and hybrid calculations with $R_c = 4$ Å and $R_c = 6$ Å are shown in Figure 4.8 to facilitate comparisons between the methods. These two values of R_c were chosen because most of the calculations detailed in Chapter 5 used these parameters. It is readily apparent from Figure 4.8 that the hybrid method is certainly no slower than the LNV method and in this case it is indeed faster. This figure also demonstrates how much faster the PM method is compared to either of the other methods. However, this gain in speed is at the expense of the exact forces, which renders the PM method unsuitable for applications requiring knowledge of the forces, such as the geometry optimisation of our capped carbon nanotubes.

4.6 Applicability of density matrix methods

Density matrix methods are only one type of many possible linear scaling methods. The different ways of achieving linear scaling make the resulting methods most appropriate for different classes of materials. We shall concentrate on the materials for which density matrix methods are most suited; a number of reviews of linear scaling methods are available which discuss this topic more generally.[92,93,98,99]

Since the central approximation in density matrix methods is truncating the density matrix, the most important consideration in determining the suitability of these methods for a particular application is the range of the density matrix. Section 4.1.4 stated that the range of the density matrix is greater for metals than insulators, which implies that density matrix methods are most suitable for insulators. This is supported by other studies[68,93,99] which have

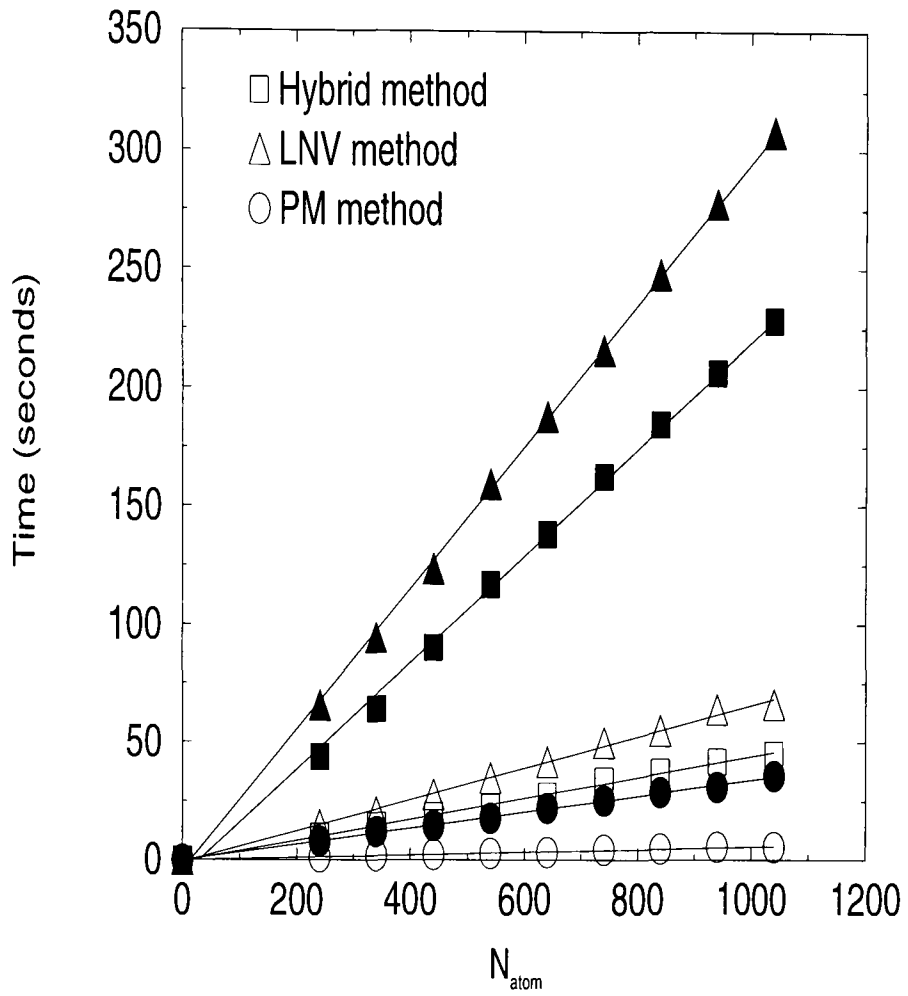


Figure 4.8: Time taken for a single geometry calculation with different methods as a function of system size with a $R_c = 6$Å (solid symbols) and $R_c = 4$Å (open symbols).

concluded that a density matrix method of some sort is an excellent choice for molecules with band gaps, but is not the optimum choice for metallic systems.

Once it has been decided that a method based on the density matrix is appropriate, the PM, LNV or hybrid method can be chosen depending on the application. When forces are not required, the PM method is the method of choice because it is fast. If forces are required, either the hybrid method or the LNV method could be used, but the hybrid method is to be preferred since it is faster and more robust.

4.7 Applications of density matrix methods

Linear scaling methods open the door to tight-binding studies of systems which were previously not accessible to this type of calculation and were therefore mainly studied (if at all) by less accurate methods. Linear scaling methods are also very well suited to molecular dynamics studies. The types of materials that can be examined include large molecules such as RNA[100] and DNA molecules, water clusters,[100] proteins, capped carbon and BN nanotubes. Molecular dynamics simulations which have used density matrix methods include studies of crystalline, amorphous and liquid carbon[68] and defects in ordered systems (defects have also been studied using another linear scaling method[98]). In fact, larger fullerenes are excellent candidates for density matrix methods and several studies using various linear scaling methods have investigated them.[67,98] Indeed, a study using the LNV method has been performed on fullerenes with up to 8640 atoms.[67] However, it should

be noted that in order to study this exceptional number of atoms, the authors applied a constant cut-off distance to all matrices ($\boldsymbol{\rho}$, $\boldsymbol{H\rho}$, *etc.*) and therefore sacrificed the variational property of the LNV method.

Before leaving this topic, it is worth mentioning in passing that these linear scaling methods are not limited to the tight-binding formalism. They have also been implemented in DFT[97,100] and Hartee Fock[100] schemes.

4.8 Summary

This chapter has presented three linear scaling methods based on the density matrix: the LNV, PM and hybrid methods. The only approximation made to exact tight-binding theory in all these methods is the introduction of a cut-off parameter R_c for the density matrix. Obviously, as $R_c \to \infty$ the methods all become exact. Methods based on the density matrix have been applied to a wide range of different materials. Density matrix methods are an excellent choice of linear scaling method for non-metals, however, other linear scaling methods are more appropriate for metals. As we will see in the next chapter, all the (finite) capped nanotubes to be examined possess a band gap, so density matrix methods are a sensible way to investigate these molecules.

Although the three methods discussed in this chapter are all based on the density matrix, they have different properties. The LNV method is variational which means the forces can be determined accurately, however, the method is rather slow and there is a danger of producing unphysical solu-

tions. The PM method is extremely fast, specifies the initial guess for the density matrix and does not produce unphysical solutions, but is not variational. The hybrid method uses the PM method first and then refines the density matrix using the LNV method such that the most desirable features of both methods are preserved. The hybrid method is variational, has a recipe for the initial guess for the density matrix and should prevent unphysical solutions. Therefore, the hybrid method was chosen for the survey of capped nanotubes described in the next chapter.

Chapter 5

Nanotube caps

Initially, theoretical interest in carbon nanotubes focussed on treating them as infinite objects as discussed in Chapter 3. In reality, however, carbon nanotubes are clearly not infinite. Furthermore, electron microscope images of nanotubes such as the one shown in Figure 5.1 have revealed that many nanotubes are terminated by caps made of carbon atoms. In this chapter the relative stabilities of different caps for three selected nanotubes are investigated using the hybrid linear scaling method introduced in the previous chapter.

The chapter starts with a discussion of why capped carbon nanotubes merit attention, before reviewing the census of nanotube caps by Brinkmann *et al.*[39] Next the results of our survey of nanotube caps are presented, followed by a discussion of how the nanotubes used in this study relate to those observed experimentally. The density of states of both infinite and capped nanotubes are examined and compared to STM measurements, and

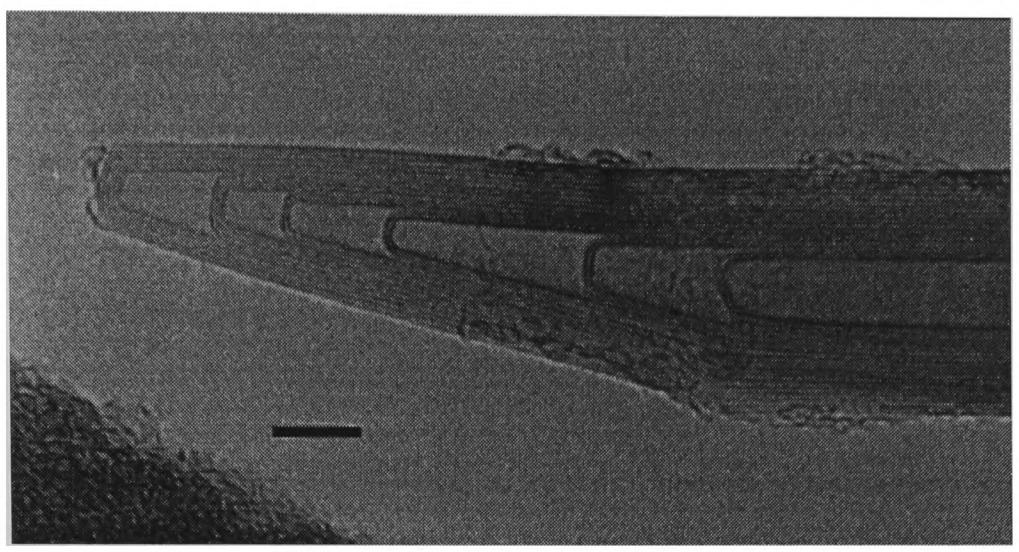


Figure 5.1: An electron micrograph of a multi-walled nanotube in which each constituent tube is capped (from "Carbon nanotubes and related structures: New materials for the twenty-first century"[108]).

the chapter concludes with a summary.

5.1 Motivation

Infinite nanotubes (without defects) are fairly well understood theoretically. However, such studies cannot say anything about the properties of a cap or whether the properties of the remainder of the tube are affected by the presence of the cap, or whether different caps would have different effects. It is therefore crucial to study nanotube caps in order to obtain a proper understanding of the properties of nanotubes. In particular, there are a number of phenomena where nanotube caps are expected to be especially influential.

5.1.1 Relevant phenomena

Nanotube growth

The growth mechanism for carbon nanotubes is still a controversial subject. It is accepted that the growth takes place at an open end of a nanotube,[109,110] even though capped ends are more stable. Further, the mechanism probably depends on the synthesis method[20] and is different for single-walled and multi-walled nanotubes. Multi-walled nanotubes may be grown without catalysts, and it has been postulated, on the basis of molecular dynamics simulations, that the open ends of multi-walled nanotubes are stabilised by "lip-lip" interactions between layers.[111,112] In contrast to multi-walled nanotubes, the growth of single-walled nanotubes does require the presence of a catalyst.[19] The role of the catalyst in the formation of single-walled nanotubes seems to be preventing closure of the tube either by stabilising the open end or by forming part of a closed structure where the metal-carbon bonds are easily broken hence allowing nanotube growth.[19] Therefore, it is likely that caps are important in inhibiting the growth of single-walled carbon nanotubes.

Scanning probe microscopies

As discussed in the introduction to this thesis, the use of carbon nanotubes as probes in scanning probe microscope techniques is perhaps the most important, and certainly one of the most developed, applications of nanotubes. Both multi-walled[29] and single-walled[30] tubes have been used in these applications. The nanotubes can either be attached as probes subsequent to

their manufacture, or generated by chemical vapour deposition (CVD) already attached. In this second case, the nanotubes are often shortened (*i.e.* uncapped).[30]

An early paper on the use of (multi-walled) nanotubes as STM probes states that

> 'While good progress has been made in controlling the position of the macroscopic probe of such devices to sub-ångström accuracy, and in designing sensitive detection schemes, less has been done to improve the probe tip itself. Ideally the tip should be as precisely defined as the object under investigation ... it is rarely clear what the atomic configuration of the tip is during imaging.'[27]

Therefore, this work aims to provide some insight into the atomic configuration of capped (unshortened) single-walled nanotube tips.

Field emission

Many potential nanotube applications make use of the field emission properties of carbon nanotubes (see Chapter 1 for more details). Field emission depends on the end of the nanotube[113] and, therefore, understanding this property relies on knowledge of the structure of the nanotube tip. It has been postulated that the electrons are actually emitted not from the Fermi level, but rather from localised states near the Fermi level.[22,114] The existence and position of such localised states are highly sensitive to the cap considered.[115] In addition, a theoretical study of a (5,5) nanotube showed

that the isolated pentagon C_{5v} cap was one of the best structures for field emission.[113] The motivation for studying capped nanotubes in this context is best summarised by a review of the progress made in understanding field emission from carbon nanotubes which states that

> 'Supplementary information on the electronic structure of the nanotube cap and on the influence of adsorbates or bonded groups are required for a better comprehension of this emission.'[22]

This thesis aims to address the first of these two issues with regard to capped single-walled carbon nanotubes.

5.1.2 Previous studies

In light of the above issues, it is not surprising that there has been a great deal of theoretical interest in nanotube caps.[16,39,91,104,105,113,115–121] However, these studies have focussed almost exclusively on the effect a cap may have on the properties of a nanotube and have not determined which caps are the most probable. This is an unfortunate omission because different caps may influence a property of interest in different ways. Most studies consider only a few of the many possible structures. For example, a comparison of experimental and theoretical results for the (11,2) tube[16] considered just one isolated pentagon cap and one adjacent pentagon cap originally, subsequently adding another isolated pentagon cap.[91] Now, there are 224 possible isolated pentagon caps and 2254 caps in total for the (11,2) tube,[39] which suggests that a study of three caps is somewhat inadequate.

In addition, almost all studies make use of only high symmetry caps (where these are possible). As will be discussed below, these are statistically unlikely. This does not mean they will not occur: there are 1812 possible isomers of C_{60}, but the only one encountered experimentally is the single I_h isomer. However, when the other fullerenes were examined, it was discovered that high symmetry isomers are not necessarily the ground state isomers. For example, in Chapter 3 we saw that the two experimentally observed isomers of C_{84} possess D_2 and D_{2d} symmetry and are not the higher symmetry isomers such as T_d or D_{6h}. The key consideration is which isomers are most stable energetically and in the absence of any proven rules or general trends for nanotube caps, this can only be established by detailed calculations.

Furthermore, previous studies have tended to assume that the isolated pentagon rule which is so reliable for fullerenes carries over to nanotube caps. This is a reasonable supposition, but to the best of my knowledge, it has never been tested before now.

Actually, an attempt has been made to identify the cap present on an experimentally observed multi-walled nanotube using a combination of HRTEM studies on the early stages of the selective structure-related nanotube oxidation process and computer simulations.[116] However, it is not clear whether the experimental data is compatible with the cap suggested in the paper and only that cap, or whether some of the other possible caps are also consistent with the observations. Also, as the oxidation process is destructive and as no synthesis method produces homogeneous nanotubes - there is always a diameter range - this approach to determining the cap is of limited use.

It is apparent, then, that in order to examine the electronic structure of a capped nanotube satisfactorily, a systematic study of all possible caps for the nanotube in question is desirable. As we saw in Chapter 3, the experimentally observed isomer(s) of the fullerenes C_{76} and C_{84} are the isomers which calculations find to be most stable. This is generally true for the fullerenes[40] suggesting that it will also be true for nanotube caps. Therefore, the results of our study should determine which caps are experimentally likely. This in turn allows the determination of the electronic structures of the capped nanotubes which most probably occur in experiments. Obviously, this a more satisfactory solution than selecting a cap at random.

These, then, are the motivations for a new, systematic tight-binding study of capped nanotubes. The first stage is to determine what caps are possible and it is to this that we shall turn our attention next.

5.2 A census of nanotube caps

A census of nanotube caps was carried out by Brinkmann *et al.*[39] for all nanotubes with $n_1 + n_2 \leq 25$ (given infinite computer resources, it would be feasible to use this method for any nanotube). This census identified all the caps consistent with all the possible combinations of n_1 and n_2 in this range (subject to the constraint $n_1 \geq n_2$), and categorised the caps according to n_1 and n_2, by symmetry and by whether or not they obeyed the isolated pentagon rule. Thus this survey clearly provides sufficient information for an investigation of cap stability. This section reviews some of the underly-

ing ideas behind the census method and its results. Interested readers are referred to the original paper for more details.[39]

A capped nanotube can be thought of as consisting of two regions: the nanotube body and the cap. In Chapter 3, it was assumed that infinite nanotubes are "perfect", *i.e.* only contain hexagons, and the corresponding assumption that the nanotube body (as distinct from the cap) consists solely of hexagons is made for capped nanotubes. According to Euler's theorem (see Chapter 3), six pentagons are required to close the tube and the cap is defined such that all of these lie in the cap.

To facilitate the enumeration of the nanotube caps, a less ambiguous definition of a cap is required. As discussed in Chapter 3, an (n_1, n_2) infinite nanotube can be regarded as a section of a graphene lattice rolled-up into a cylinder. Such a situation is illustrated for the (5,5) tube in Figure 5.2. Imagine that the tube body continues along the dotted lines, but that the end shown is attached to a cap. A unique boundary marking the end of the tube body can be drawn on the graphene lattice along the carbon bonds (and, in fact, the infinite tube can be regarded as a series of rings with this structure). The boundary for a (5,5) nanotube is illustrated in Figure 5.2. This boundary follows the bonds directly below the vector $n_1\boldsymbol{a}_1$, then follows those to the right of $n_2\boldsymbol{a}_2$. Since a nanotube cap for a given tube must have the same boundary structure this can be used as a criterion for finding the possible caps for a given tube. It is, however, necessary to further refine the definition of a cap such that a pentagon appears in the boundary to prevent otherwise identical caps which have different numbers of hexagons attached

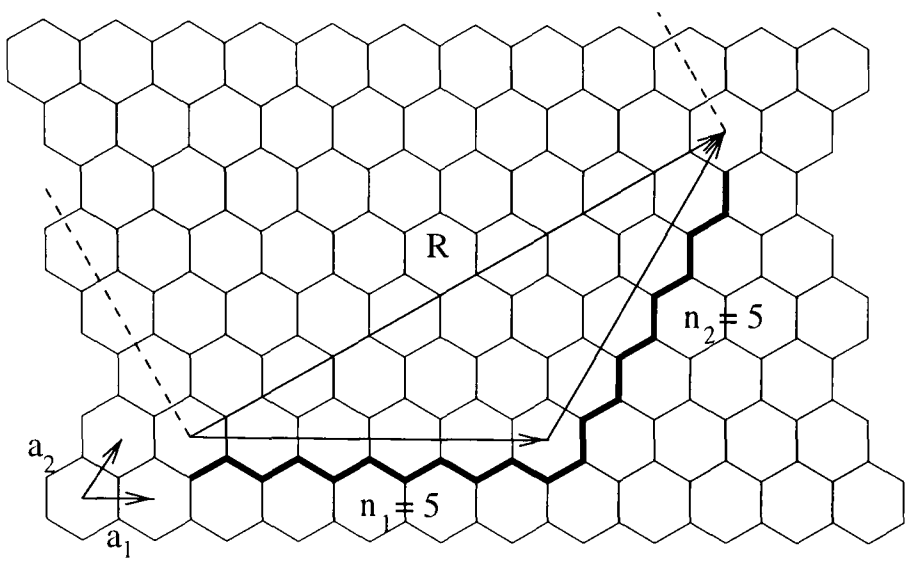


Figure 5.2: The boundary for the (5,5) nanotube.

from being counted separately.

Having established an unambiguous definition for the boundary structure of a cap, it follows that for a nanotube to be capped by a given cap, it too must have this boundary structure. The census of nanotube caps used this idea.[39] However, the method is rather involved and the details are beyond the scope of this thesis. Nevertheless, before we move on to discuss the results, the following points arising from the method of enumeration are noteworthy:[39]

1. The method does not count enantiomeric caps.

2. The method can be shown to be mathematically complete, that is, it generates all the possible caps for a given tube within the assumptions listed above and without repetition.[115]

$n_1+n_2 \backslash\ n_2$	0	1	2	3	4	5	6	7	8	9	10
5	1	0	0								
6	5	1	1	1							
7	13	7	4	3							
8	42	30	22	14	12						
9	106	99	66	51	42						
10	258	273	199	143	127	73					
11	552	663	476	364	301	273					
12	1153	1463	1101	840	708	612	348				
13	2199	2990	2254	1771	1463	1281	1197				
14	4083	5742	4459	3510	2946	2530	2354	1223			
15	7157	10472	8184	6589	5481	4751	4311	4095			
16	12208	18278	14577	11781	9928	8536	7736	7192	3731		
17	19984	30723	24682	20254	17043	14763	13209	12243	11781		
18	31998	49980	40756	33636	28567	24682	22130	20254	19342	9787	
19	49632	79002	64872	54145	46060	40044	35673	32637	30723	29799	
20	75558	121737	101123	84854	72747	63251	56485	51340	48140	46060	23316

Table 5.1: The number of semi-infinite nanotubes.[39]

Tables 5.1 and 5.2 list the number of possible caps for nanotubes between (5,0) and (10,10), and the number of isolated pentagon (IPR) caps for nanotubes between (9,0) and (10,10).[39] Figure 5.3 shows the total number of caps and the number of IPR caps plotted as a function of a quantity proportional to the nanotube diameter (see Chapter 3).[39] It can be seen that as the diameter of the tube increases, the number of possible caps increases very rapidly. If the possible caps are further classified by symmetry,[39] it becomes clear that as the tube radius increases, low symmetry caps tend to predominate which means that high symmetry caps are statistically unlikely.

The results of this census prompt the question: which caps out of the vast number of possible caps are likely to be found experimentally for a given nanotube? This is the subject I shall address next.

5.3 Extrapolating the energy of a cap

The work presented in the following sections comprises a study of a few representative nanotubes: the (5,5) tube (an example of a metal); the (10,0) tube (an example of a semi-conductor) and the (11,2) tube (for which experimental results are available for comparison[16,91] - see below).

$n_1+n_2 \setminus n_2$	0	1	2	3	4	5	6	7	8	9	10
9	1	0	0	0	0						
10	7	1	1	0	0	1					
11	31	18	7	3	2	1					
12	124	99	58	33	22	14	18				
13	347	364	224	143	99	77	66				
14	889	1043	708	476	362	273	245	145			
15	1963	2530	1771	1288	969	776	669	612			
16	4032	5481	4039	2990	2354	1886	1639	1463	805		
17	7617	10899	8184	6293	4988	4095	3510	3159	2990		
18	13754	20254	15675	12255	9928	8184	7104	6293	5929	3047	
19	23473	35673	28017	22386	18278	15314	13209	11781	10899	10472	
20	38777	60102	48140	38916	32278	27151	23618	20951	19342	18278	9342

Table 5.2: The number of semi-infinite nanotubes with isolated pentagons.[39]

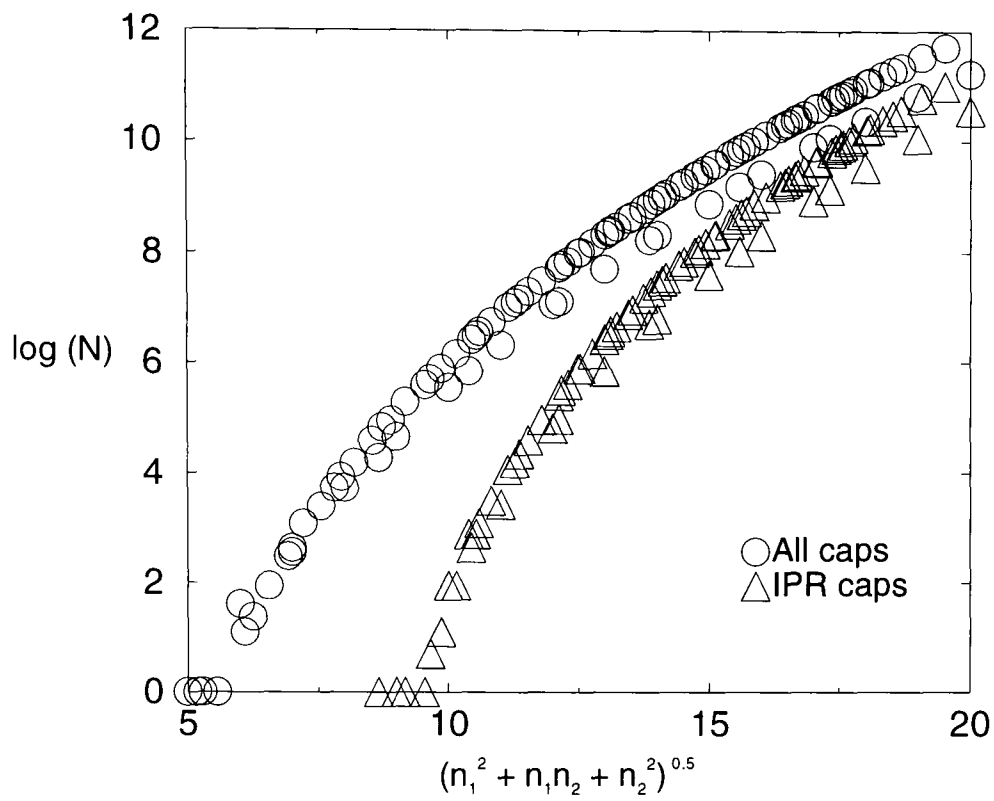


Figure 5.3: The number of caps as a function of diameter.[39]

At the present time state-of-the-art experiments cannot determine the topology of a nanotube cap *i.e.* it is not yet possible to directly establish which cap caps a nanotube. However, the local density of states (LDOS) in the cap region of an experimentally observed nanotube has been measured by STM[16,91] (this will be discussed in more detail in Section 5.6) and it was possible to determine the helicity and diameter of the tube and thus deduce that an (11,2) tube was examined. It was not possible to locate the pentagons in the cap and hence establish unambiguously which cap it was. However, this experiment provides at least some data with which we can compare the results of our study. If the most stable cap(s) for the (11,2) tube can be determined and then the LDOS calculated, the LDOS can be directly compared to the experiment. One of the key aims of the present work is thus to

determine whether STM measurements allow an unambiguous assignment of the cap.

In order to simulate a capped carbon nanotube I have used finite segments from the infinite tube and connected both ends to two identical caps (this idea has also been used in other studies of capped nanotubes[104]). The adjacency matrix of this molecule is used to generate topological coordinates[72] and then the geometry of the doubly capped tube is optimised using the conjugate gradient method,[69] with the grand canonical potential obtained using the hybrid density matrix method described in the previous chapter. The Fermi level is set to that of the infinite tube as this work is actually interested in investigating the properties of capped semi-infinite tubes. The segment from the infinite tube was gradually increased to obtain results for different numbers of atoms. The calculations were performed with segments of 7 to 47 rings, in increments of 5 rings. Each ring consists of $2(n_1 + n_2)$ atoms and has the boundary structure described in Section 5.2. The calculations were carried out using the full valence tight-binding Hamiltonian of Ho *et al.*[41] which was introduced in Chapter 2. This Hamiltonian yields realistic results for both the fullerenes and infinite nanotubes (see Chapter 3), and it is therefore reasonable to expect realistic results for capped nanotubes too.

The cut-off parameter for the density matrix in the hybrid density matrix method must also be chosen. Figure 5.4 shows the energy per atom for the (5,5) and (10,0) tubes, both with seven rings from the respective infinite tubes and doubly capped with C_{5v} caps. The energy per atom converges rapidly with increasing R_c, showing that the use of a method employing a cut-off

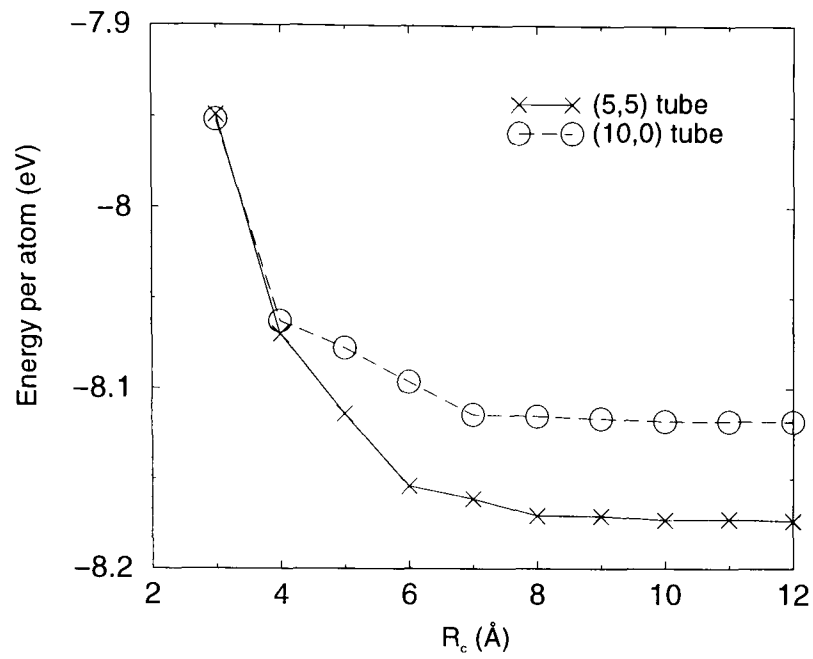


Figure 5.4: Energies per atom as a function of R_c for the (5,5) and (10,0) tubes. Both contain seven rings from the respective infinite nanotubes and are doubly capped with C_{5v} caps. The (5,5) tube consists of 240 atoms, while the (10,0) tube contains 220 atoms.

parameter is justified. This trend is expected to be generally observed for doubly capped nanotubes. The final value chosen for R_c depends not only on the accuracy of the resulting energies, but also on the available computational resources since increasing R_c results in a more demanding calculation in terms of CPU time and memory requirements. With these considerations in mind, we chose $R_c = 4$ Å and $R_c = 6$ Å for this work.

The next stage is to plot the resulting grand canonical potentials $\Omega(N_{\text{atom}})$ against the number of atoms N_{atom} in the tube. This is illustrated for the $C_1(1)$ cap of the (5,5) tube in Figure 5.5, and it is seen to yield a straight line:

$$\Omega(N_{\text{atom}}) = mN_{\text{atom}} + c. \tag{5.1}$$

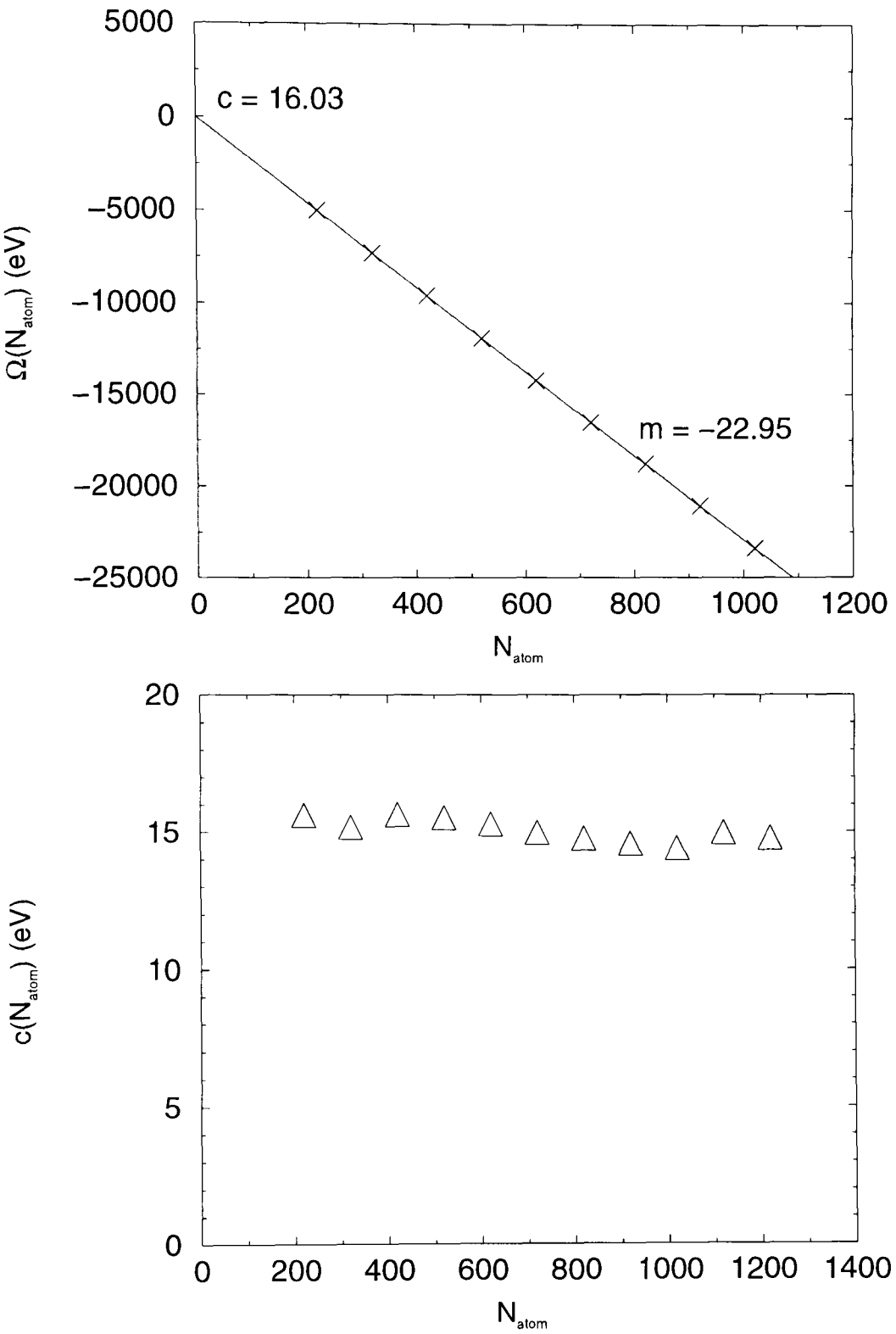


Figure 5.5: $\Omega(N_{\text{atom}})$ (top) and $c(N_{\text{atom}})$ (below) plotted against N_{atom} for the $C_1(1)$ cap of the (5,5) nanotube (top). $c(N_{\text{atom}})$ is calculated via Eqn. 5.3.

As $N_{\text{atom}} \to \infty$, $m(= \partial\Omega/\partial N_{\text{atom}}) \to \Omega_\infty$, where Ω_∞ is the energy per atom of an infinite tube. Since this work is concerned with mimicking semi-infinite nanotubes, we assume that for the lengths of tube we are examining $m = \Omega_\infty$. In fact, the gradient calculated in Figure 5.5 is -22.95 eV/atom and Ω_∞ calculated for an infinite (5,5) nanotube is -23.05 eV/atom. This excellent agreement justifies this assumption which in turn allows us to determine the y-intercept using Eqn. 5.1:

$$c = \lim_{N_{\text{atom}} \to \infty} c(N_{\text{atom}}), \tag{5.2}$$

where

$$c(N_{\text{atom}}) = \Omega(N_{\text{atom}}) - \Omega_\infty N_{\text{atom}} \tag{5.3}$$

i.e. the y-intercept is equal to the energy of the doubly capped tube minus the energy of N_{atom} atoms in the infinite tube. This difference in energy is attributable to the strain energy of the caps and thus the cap energy is defined as half the y-intercept. The lower panel in Figure 5.5 illustrates the variation of $c(N_{\text{atom}})$ in Eqn. 5.3 as a function of N_{atom}. As $c(N_{\text{atom}})$ is approximately constant for all values of N_{atom}, this suggests that the extrapolation described above is indeed a reliable method for obtaining the cap energy.

The extrapolation to the $N_{\text{atom}} = \infty$ limit was investigated. Figure 5.6 shows the results for four (5,5) C_1 caps. The cap energies obtained from larger molecules are very similar to those obtained with the range of molecule size used in this work. Certainly, the difference is much less than that between the results for $R_c = 4$ Å and $R_c = 6$ Å. The largest N_{atom} to be used in this study was chosen with reference to these results whilst bearing in mind that the calculation of the density of states requires optimised coordinates

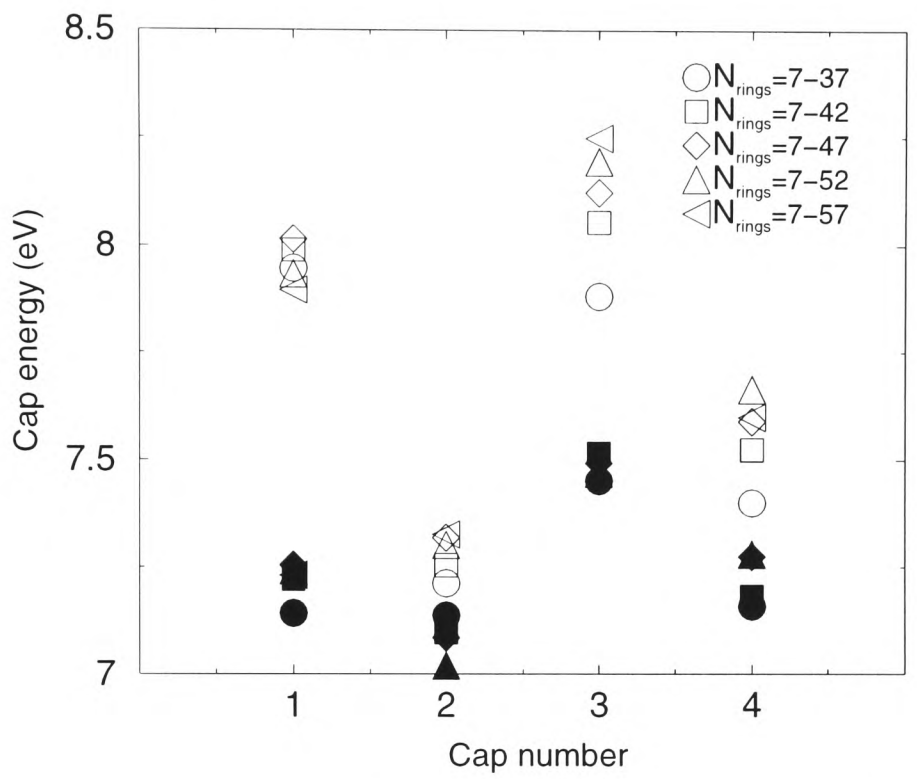


Figure 5.6: Cap energies for the (5,5) nanotube with $C_1(1)$, $C_1(2)$, $C_1(3)$ and $C_1(4)$ caps where different ranges of N_{atom} have been used to extrapolate. Filled symbols correspond to results obtained with $R_c = 6$ Å, open symbols to results obtained with $R_c = 4$ Å.

for a reasonably long tube. There was no discernible difference between the extrapolation for a metallic tube and for a semi-conducting tube.

At first glance, this study may seem to be simply a study of elongated fullerenes. However, fundamentally, capped nanotubes are *not* fullerenes. First, the energy levels are filled to the Fermi level of the corresponding infinite tube rather than merely occupying the lowest $2N_{\mathrm{atom}}$ orbitals. This will, in general, lead to an excess of electrons because the pentagons in the cap region give rise to more bonding than anti-bonding orbitals. Secondly, studies of fullerenes are concerned with the relative stabilities of all isomers of $\mathrm{C}_{N_{\mathrm{atom}}}$. This study compares caps which fit onto the same bulk tube structure, which can be achieved by caps with differing numbers of atoms.

Before the results are discussed, the notation used to identify the caps must be explained. When there is more than one cap of a given symmetry for a given tube, these are distinguished by a number showing the position in the output from the census code.[39] In the case of the (5,5) and (10,0) tubes this number relates to the list of all caps of that symmetry; for the (11,2) tube, the number refers only to the isolated pentagon caps of a given symmetry. Finally, (5,5):C_1(1) refers to a (5,5) tube doubly capped with the first C_1 cap output by the census code.

5.4 Results

5.4.1 (5,5) tube

The (5,5) tube is metallic and has 73 possible caps: 40 C_1 caps; 32 C_s caps, and 1 C_{5v} cap which is the only isolated pentagon cap. Figure 5.7 shows the energy of these caps as a function of the number of bonds shared by pentagons. The calculations were performed with $R_c = 4$ Å and $R_c = 6$ Å and Figure 5.7 shows that the general trend is the same whichever cut-off distance is chosen. A larger cut-off distance means a more accurate calculation, at the cost of a more expensive calculation. It is unrealistic to expect to carry out these calculations for a very large value of R_c, or for many different values of R_c. However, the fact that the results for $R_c = 4$ Å and $R_c = 6$ Å are in excellent qualitative agreement suggests that the trends seen are those that would be seen with an infinite cut-off distance, and that the conclusions of this survey should therefore be applicable to experimental observations.

Figure 5.7 shows that the energy of the cap increases (on average) in proportion to the number of bonds shared by pentagons. As the cut-off distance is increased from $R_c = 4$ Å to $R_c = 6$ Å the calculated cap energies do change (in most cases lessening), but the overall trend remains the same. Adjacent pentagons destabilise the cap and the greater the number of adjacent pentagons, the less stable the cap is. This is exact the same situation as for the fullerenes[40,75,76] (see Section 3.1.2). In particular, if the data in Figure 5.7 ($R_c = 6$ Å) is fitted to a straight line, the energy penalty associated with a pair of adjacent pentagons is found to be approximately 0.8 eV, a result

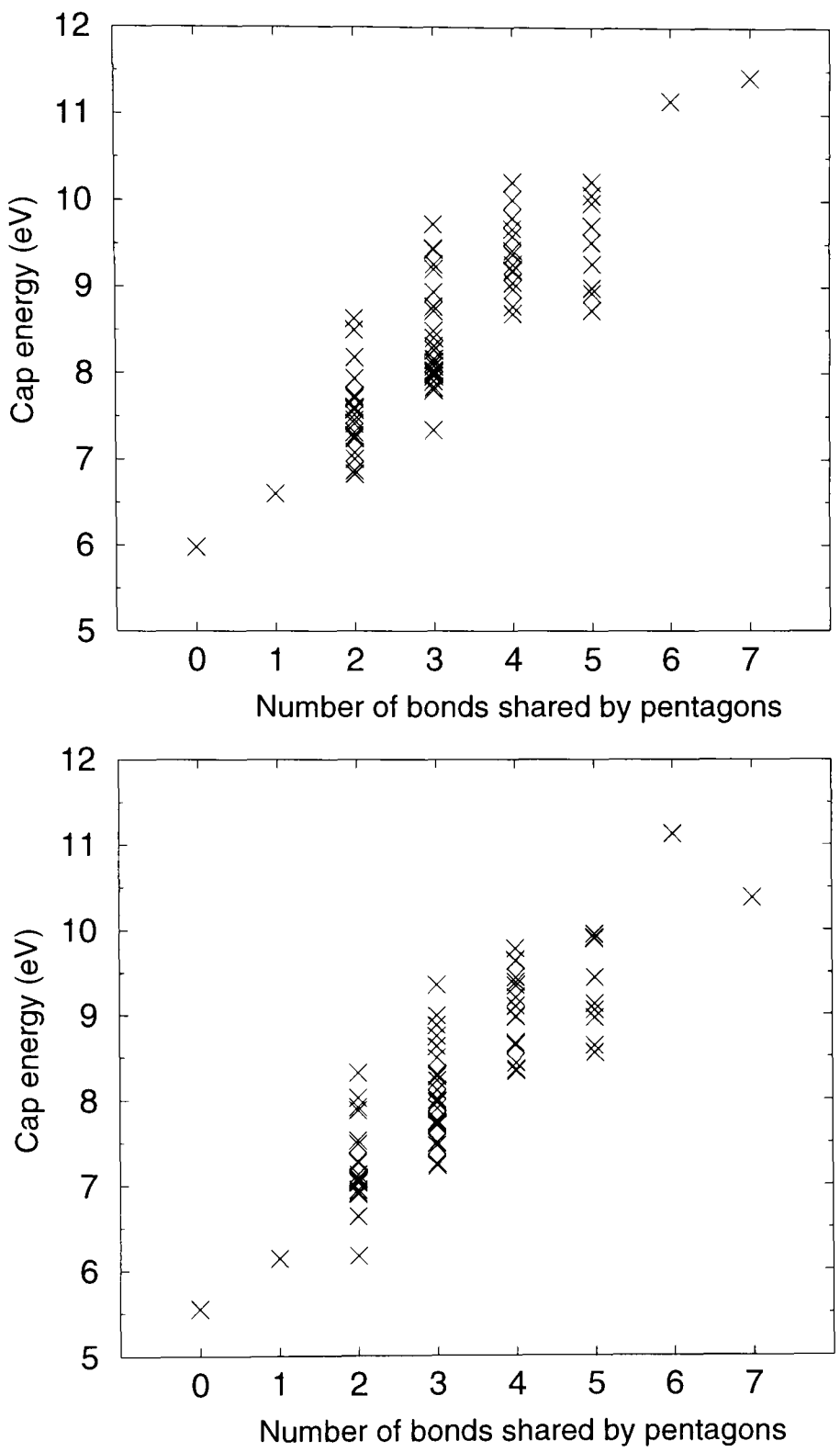


Figure 5.7: The energy of (5,5) tube caps where $R_c = 4$ Å (top) and $R_c = 6$ Å (bottom).

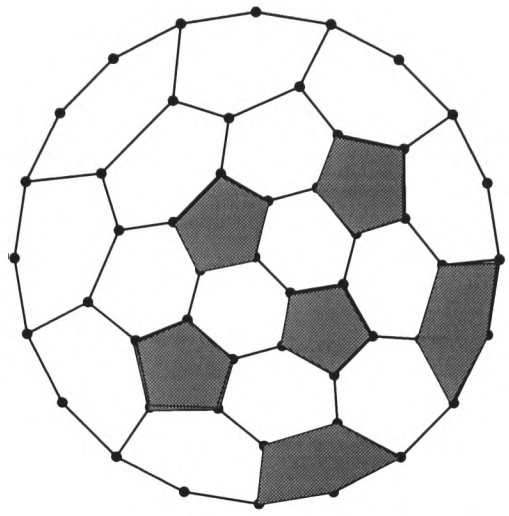


Figure 5.8: The C_{5v} cap for the (5,5) tube. This is the only isolated pentagon cap for this nanotube and is predicted to be the lowest in energy.

which is similar to previous theoretical estimates of 0.7 eV[66] and 0.9 eV[40,122] for the fullerenes themselves. Therefore, the IPR seems to carry over from the fullerenes to capped nanotubes. In line with these findings, the C_{5v} cap, the single isolated pentagon cap, is found to be the lowest in energy (for both density truncation parameters R_c = 4 and 6 Å). The C_{5v} cap is shown as a 2-D Schlegel diagram (bordered by the boundary path in Figure 5.2) in Figure 5.8.

5.4.2 (10,0) tube

The (10,0) tube is an insulator with 258 possible caps, including 7 isolated pentagon caps. These comprise 166 (0 IPR) caps with C_1 symmetry; 71 (3) caps with C_s symmetry; 12 (3) caps with C_{2v} symmetry; 8 (0) caps with C_2 symmetry and 1 (1) cap with C_{5v} symmetry. As in the case of the (5,5)

tube, these calculations were performed with two separate cut-off distances: $R_c = 4$ Å and $R_c = 6$ Å. The results of these calculations are summarised in Figure 5.9. The general trend for $R_c = 4$ Å is once again replicated with $R_c = 6$ Å. Increasing R_c results in some changes in the ordering in cap energy, but it does not effect the lowest energy cap. The stability of a cap decreases as the number of adjacent pentagons increases, as one would expect on the basis of the isolated pentagon rule. The energy penalty for a pair of adjacent pentagons can be estimated (from the $R_c = 6$ Å data) to be about 0.9 eV, which is in good agreement with both the (5,5) tube results and previous studies on fullerenes.[40,66,122]

I shall now discuss the seven isolated pentagon caps in more detail. The energies of these caps are plotted in Figure 5.10 where it can be seen that the C_{5v} cap is lowest in energy. The discussion of the IPR caps will be aided by considering modified hexagon indices. Hexagon indices were first applied to fullerenes[40,77] in an attempt to rationalise trends in stability beyond the isolated pentagon rule. Each index h_n is the number of hexagons surrounded by n hexagons where $n = 0, 1, \cdots, 6$. However, for isolated pentagon isomers $h_0 = h_1 = h_2 = 0$ which leaves us with a hexagon index signature (h_3, h_4, h_5, h_6). Raghavachari proposed that the most stable fullerene isomer of $C_{N_{\text{atom}}}$ should have all its hexagons in as similar environment as possible so that the strain energy is equally spread over the whole molecule.[77] These signatures have been used to rationalise ground state fullerene isomers with mixed success.[40]

These indices must be further modified before being applied to nanotube

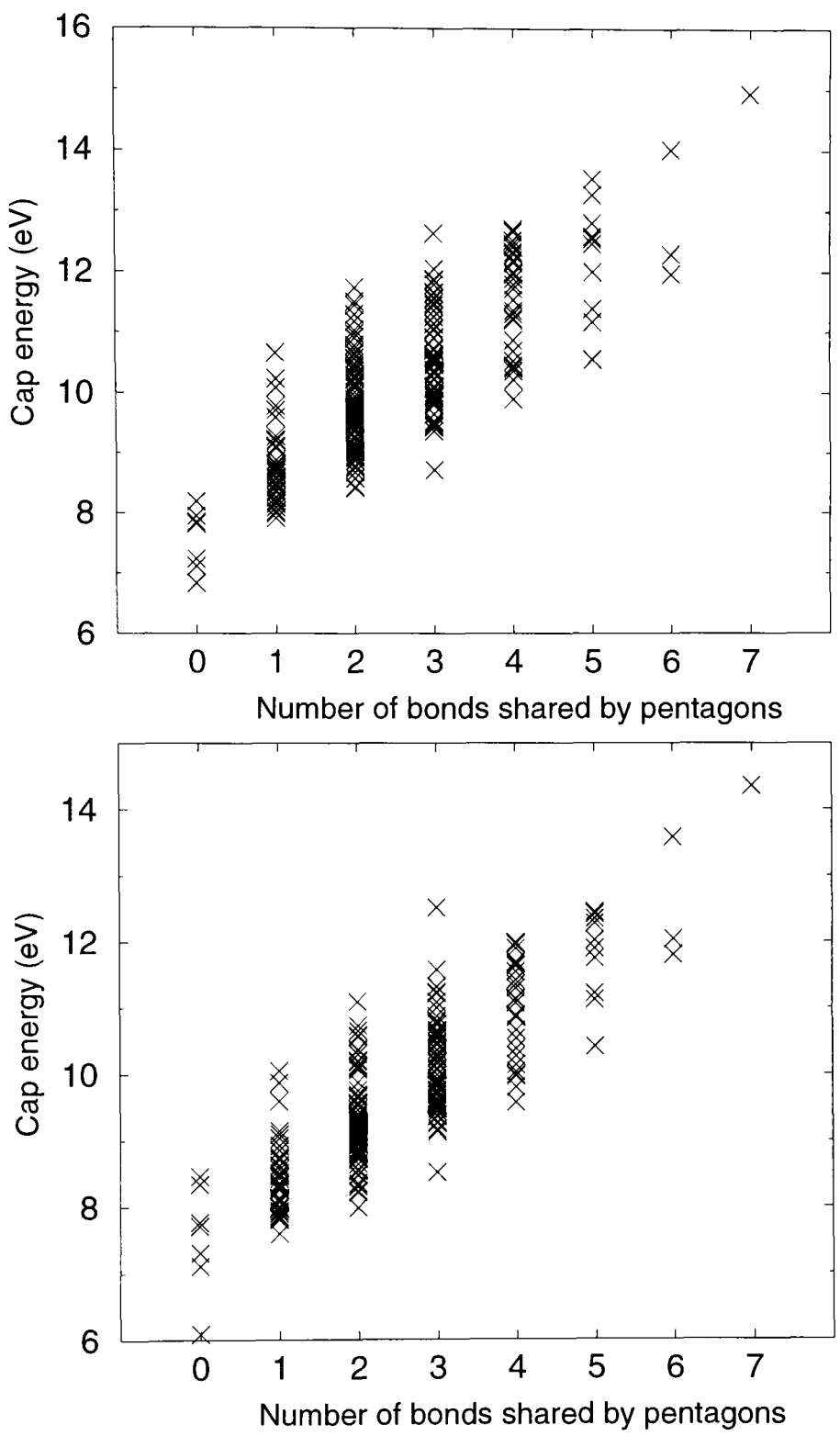


Figure 5.9: The energy of (10,0) tube caps with $R_c = 4$ Å (top) and $R_c = 6$ Å (bottom).

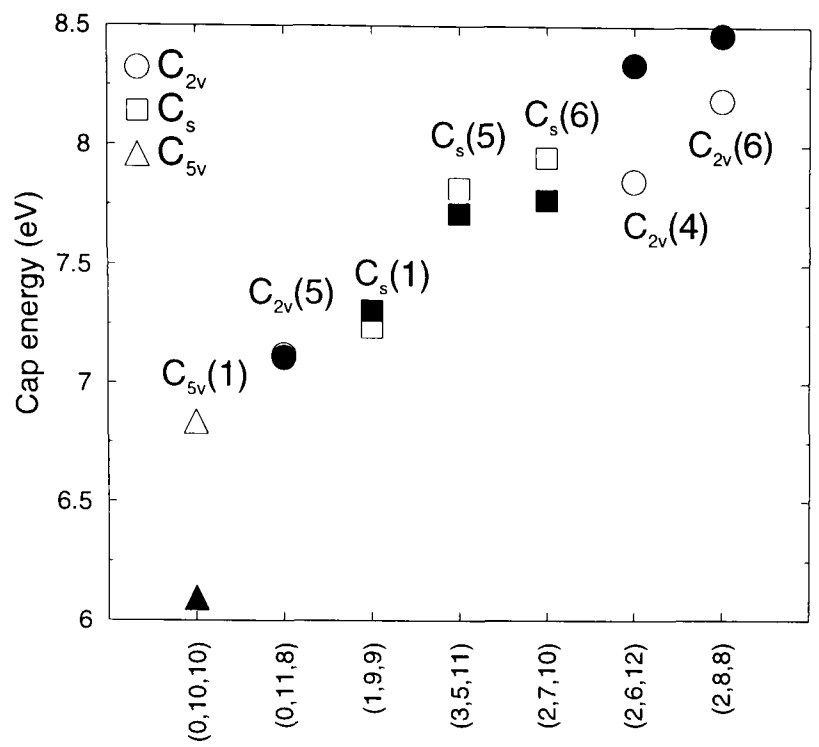


Figure 5.10: Energies of the (10,0) IPR caps with $R_c = 4$ Å (open symbols) and $R_c = 6$ Å (filled symbols).

caps. This is because the main body of the tube only consists of hexagons and hence all the hexagons in this section of the tube will be surrounded by six hexagons. Accordingly, this index depends on the length of the tube considered rather than on the cap. Therefore h_6 is omitted from the index signature.

It is possible to reduce the hexagon indices to just one parameter σ_h[40] which is the standard deviation of the hexagon neighbour index distribution:

$$\sigma_h = \sqrt{\langle k^2 \rangle - \langle k \rangle^2}, \tag{5.4}$$

where

$$\langle k \rangle = \sum_{k=3}^{5} k h_k / \sum_{k=3}^{5} h_k, \tag{5.5}$$

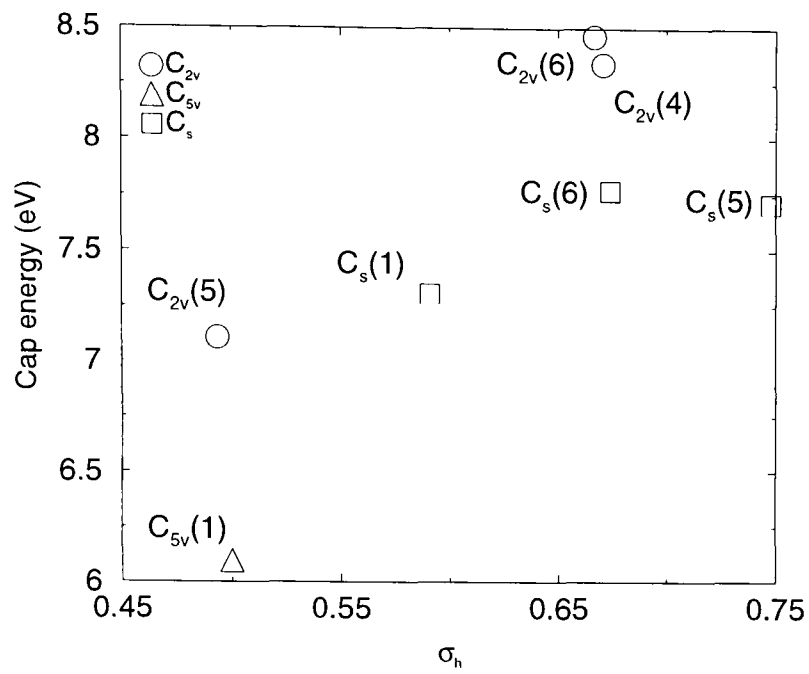


Figure 5.11: Energies of the (10,0) isolated pentagon caps with $R_c = 6$ Å plotted as a function of σ_h in Eqn. 5.4.

and

$$\langle k^2 \rangle = \sum_{k=3}^{5} k^2 h_k / \sum_{k=3}^{5} h_k. \tag{5.6}$$

The definition of σ_h has been altered as per the modification of the hexagon indices discussed above. Note also that in this work σ_h refers to a single cap and not the doubly capped nanotube. σ_h should be a measure of the strain for a particular cap and is plotted against the calculated cap energy for the (10,0) IPR tubes in Figure 5.11. There is some correlation with cap energy, but on the basis of the value of σ_h alone one would predict a different cap to be the most stable: $C_{2v}(5)$ with hexagon index signature (0,11,8). Thus the strain parameter is not useful in this case. Actually, the difference between σ_h for the $C_{2v}(5)$ cap ($\sigma_h = 0.49$) and the $C_{5v}(1)$ cap ($\sigma_h = 0.5$) is small. This may be because the definition employed for the caps ignores any hexagons surrounded by hexagons in the tube body. This ensures σ_h is independent

of tube length, but, in so doing, the standard deviation becomes merely the standard deviation of hexagons next to pentagons. The fact that most of the hexagons are in the tube body and are only connected to other hexagons is disregarded. It is difficult, owing to the minimal data available, to see any general trends in IPR cap stability.

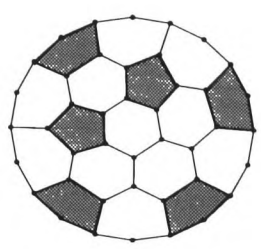
Finally, Figure 5.12 shows Schlegel diagrams of the (10,0) isolated pentagon caps and the corresponding hexagon indices (note that the hexagon indices given are for a single cap and not the doubly capped nanotube).

5.4.3 (11,2) tube

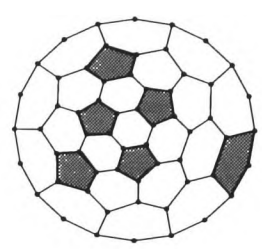
The (11,2) nanotube has 2254 possible caps, all of which possess C_1 symmetry. It is clearly not feasible to undertake a systematic tight-binding study on all these caps. However, it is desirable to do so because of the availability of experimental data for this tube. Fortunately, the above studies of the (5,5) and the (10,0) tubes allow us to conclude that the isolated pentagon rule does indeed apply for nanotube caps and therefore, in this study we shall only consider the 224 isolated pentagon caps.

As before, the calculations were carried out for $R_c = 4$ Å and $R_c = 6$ Å. However, in this case, the calculations were not repeated with $R_c = 6$ Å for all the caps, merely the twenty lowest energy caps in the $R_c = 4$ Å calculation which are likely to be the most stable caps. This was necessary owing to computational constraints.

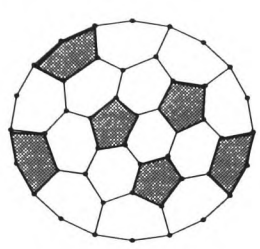
Figure 5.13 shows the results for $R_c = 4$ Å and $R_c = 6$ Å. The caps are



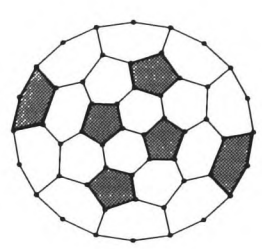
C_s (1) (1,9,9)



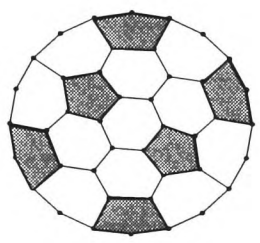
C_s (5) (3,5,11)



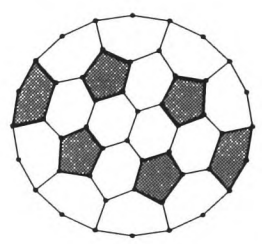
C_s (6) (2,7,10)



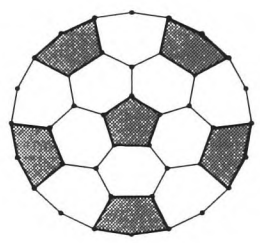
C_{2v} (4) (2,6,12)



C_{2v} (5) (0,11,8)



C_{2v} (6) (2,8,8)



C_{5v} (0,10,10)

Figure 5.12: The (10,0) isolated pentagon caps shown as Schlegel diagrams with hexagon indices.

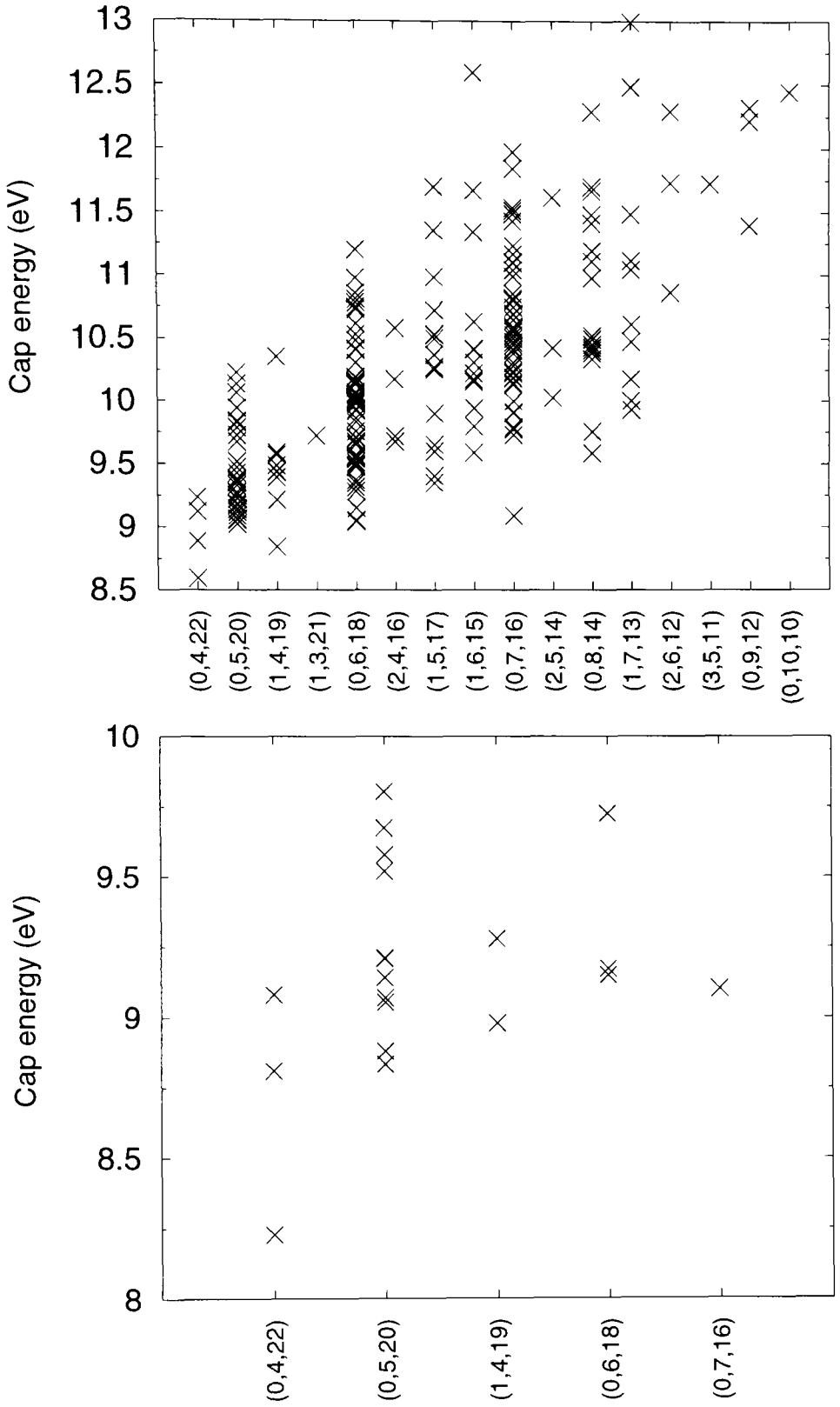


Figure 5.13: Calculated cap energies for the (11,2) tube with $R_c = 4$ Å (top) and for the twenty most stable caps with $R_c = 6$ Å (bottom). The caps are classified according to their hexagon indices. These are ordered according to the mean cap energy for a given hexagon index signature.

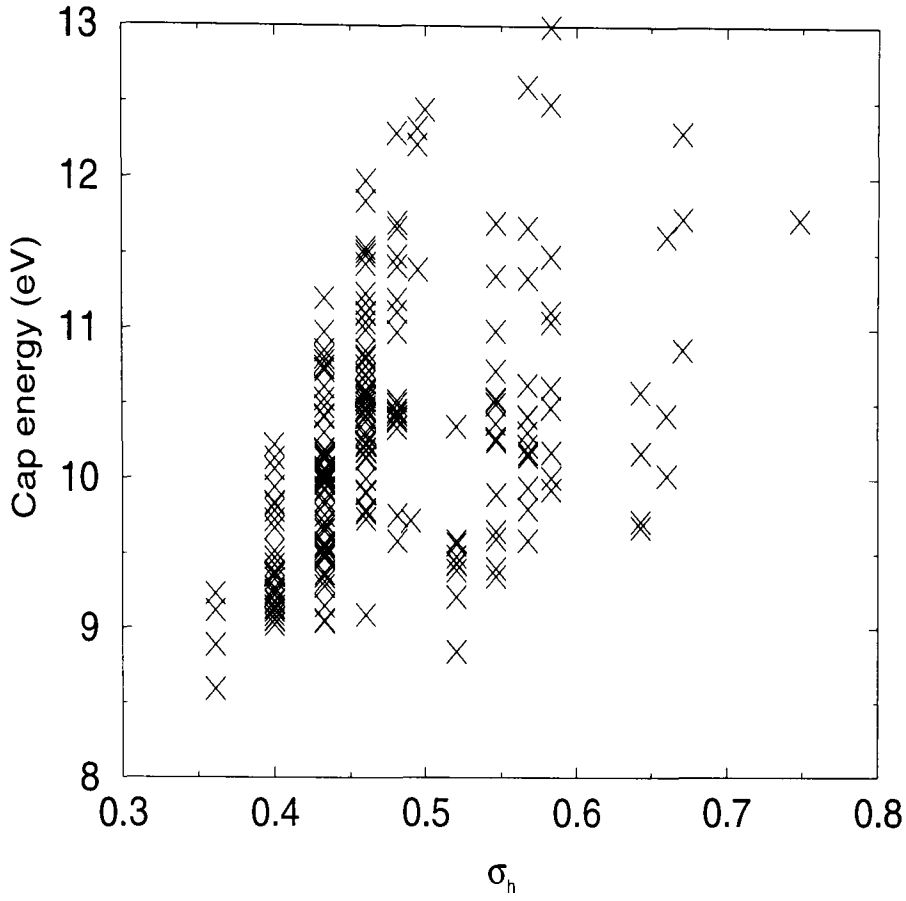


Figure 5.14: Calculated cap energies for the (11,2) tube with $R_c = 4$ Å as a function of σ_h in Eqn. 5.4.

arranged by hexagon index signature (for the cap, not the doubly capped nanotube) and are ordered according to the mean cap energy for each signature. There does seem to be some correlation between cap energy and the hexagon index signature which can be seen most clearly in the $R_c = 4$ Å case since more data is available.

In Figure 5.14 the cap energies for $R_c = 4$ Å are plotted as a function of σ_h. In contrast with the (10,0) tube, the cap with the lowest σ_h is predicted to be the lowest energy cap for the (11,2) tube. However, it is clear that σ_h still does not correlate especially well with cap energy for the other caps.

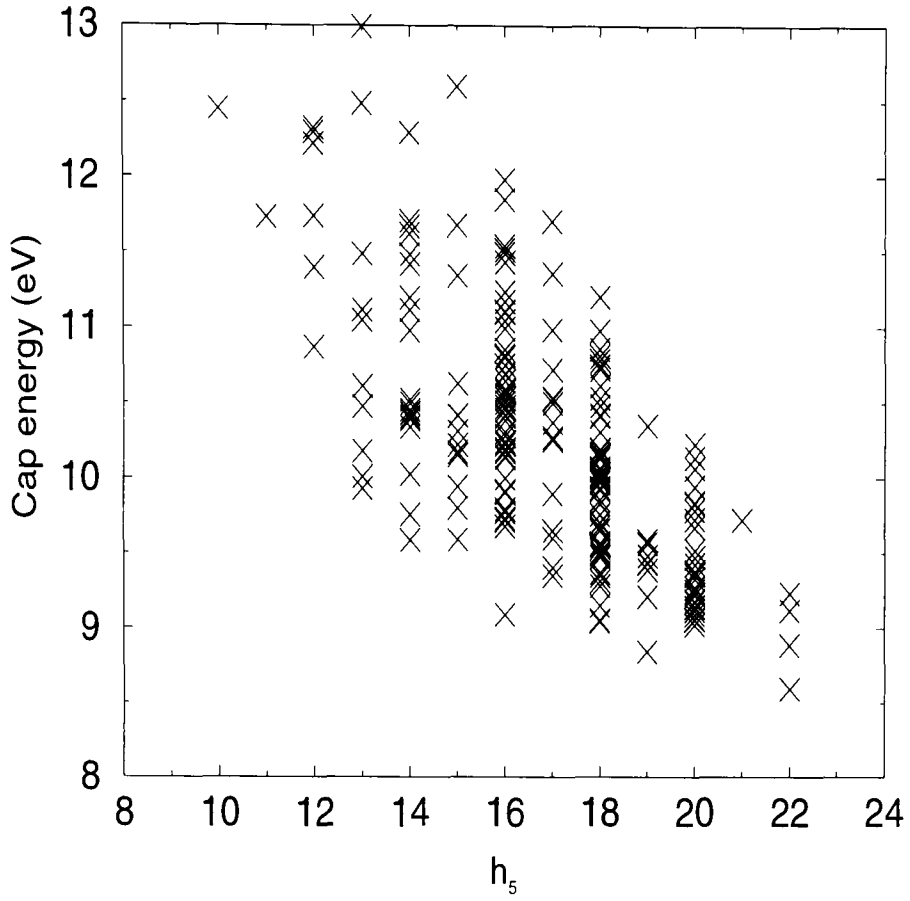


Figure 5.15: Calculated cap energies for the (11,2) tube with $R_c = 4$ Å plotted against h_5.

Closer inspection of Figure 5.13 reveals that as h_5 increases, the caps become more stable. Figure 5.15 shows the cap energies against h_5. This correlation is interesting because, in one sense, h_5 is a measure of the pentagon distribution since h_5 is the number of hexagons adjacent to five hexagons and just one pentagon. A large value of h_5 corresponds to pentagons which are widely separated, the maximum possible value being $6 \times 5 = 30$ (which cannot be realised for the (11,2) tube because of its small diameter). The lowest energy cap $C_1(78)$, which has a hexagon neighbour index signature of (0,4,22), is shown in Figure 5.16.

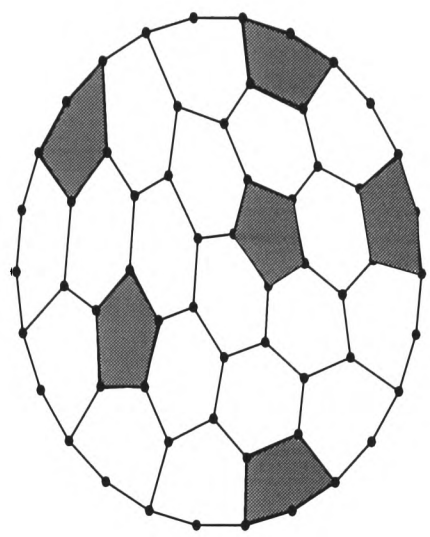


Figure 5.16: The $C_1(78)$ cap for the (11,2) tube; our calculations predict this to be the most stable cap for this particular nanotube.

It is evident that, for the (11,2) tube at least, there is a clear correlation of the cap energy with h_5. This trend is *not* seen for the (10,0) tube, although, in that case there is rather limited data (there are only seven IPR caps) with which to compare. In any case, any energetic trends beyond the isolated pentagon rule are likely to be subtle and more elusive as experience with the fullerenes has shown. For example, the ground state isomers of C_{84} are amongst those the hexagon index signatures predict, but this is not the case for C_{76} or C_{78}.[40] As with the fullerenes, it would seem that the hexagon neighbour index signatures *may* be useful for explaining known results (and are certainly useful for labelling purposes), but they are no substitute for detailed calculations.

5.5 Relevance of this survey to experiment

This chapter raises the question of whether the nanotubes studied here are representative of the nanotubes synthesised in a laboratory. I shall therefore now move on to discuss how the nanotubes investigated in this thesis relate to experimentally observed tubes.

First, how do the diameters of these tubes compare to the diameters of experimentally observed single-walled nanotubes? To some extent, the diameters of nanotubes depend on the synthesis method.[9,10,23,123] For example, a carbon arc discharge method using an iron catalyst and a mixture of methane and argon gases yielded a diameter range of 0.7 nm to 1.6 nm, with the most common diameters being around 0.8 nm and 1.05 nm.[9] However, another carbon arc discharge experiment with different experimental conditions (including a cobalt catalyst and helium gas) found that the nanotubes produced almost all had a diameter of 1.2 ± 0.1 nm.[10] A recent CVD method found a diameter distribution ranging from 0.7 nm to 5 nm.[23] The diameters of the tubes studied in this work are 0.68 nm [(5,5)]; 0.78 nm [(10,0)] and 0.95 nm [(11,2)]. Thus the tubes studied are at the low end of the range of the experimentally observed diameters.

It is worth noting that the innermost nanotube in a multi-walled nanotube may have a much smaller diameter.[124,125] The smallest diameter tube so far reported has a diameter of 0.4 nm and has been characterised as the (3,3) nanotube.[125] Of course, the present study has no direct relevance to multi-walled tubes because it neglects the dispersion interactions between different tubes.[12] However, exceptional synthesis techniques have also produced

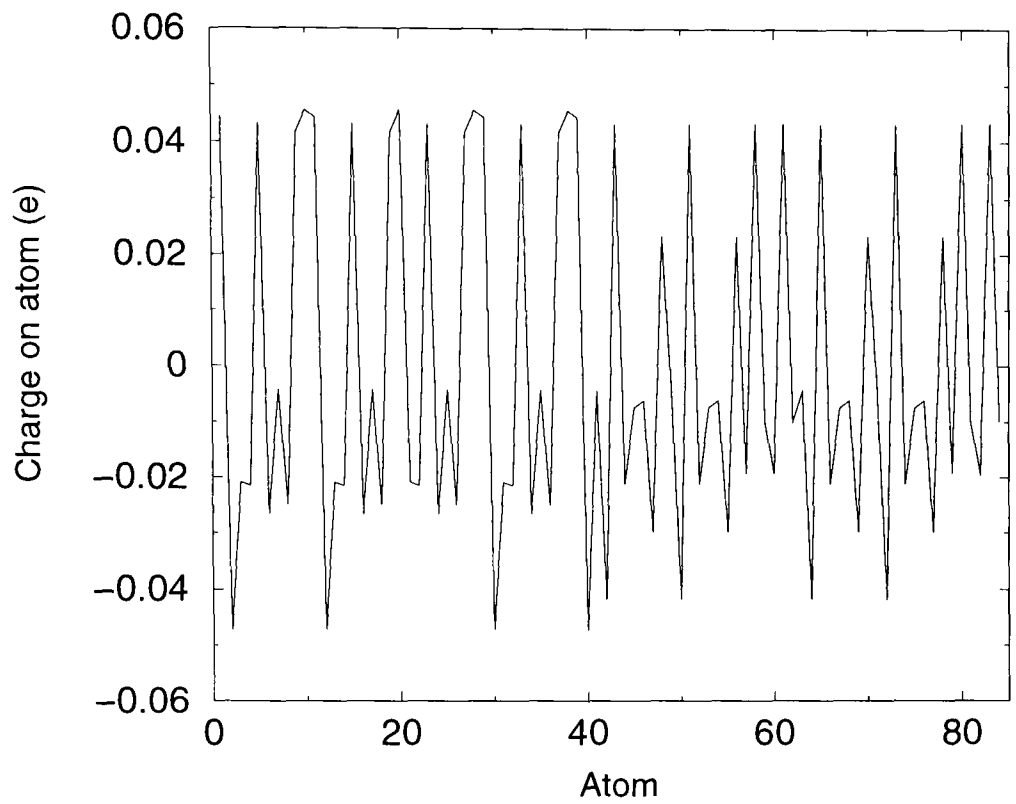
single-walled nanotubes with a diameter of 0.4 nm.[126] These tubes have not been fully characterised, but the diameter is consistent with a (3,3), a (4,2) or a (5,0) tube,[126] each of which admit one possible (non-isolated pentagon) cap (see Table 5.1).

Secondly, the length of these nanotubes relative to experimentally observed tubes must be considered. The length of the experimental tubes varies and depends on the method of production. A CVD synthesis can produce nanotubes of between 300 nm and 10000 nm in length.[23] In this study, the smallest tubes (with 7 rings of infinite tube between the capped ends) have a minimum of around 200 atoms and are about 2.5 nm in length. The largest tubes (with 47 rings) have a minimum of around 1000 atoms and are about 12.5 nm in length. It is therefore clear that the tubes in this work are considerably shorter than those seen experimentally. This is not uncommon in theoretical work on nanotubes as it is simply impossible at the present time to perform an adequate level of calculation with very large numbers of atoms. However, I do not believe this issue to be of crucial importance, at least as far as determining the lowest energy cap is concerned, since the results of the study are based on an extrapolation, rather than on one finite tube. A fuller discussion of the effects of finite length will be reserved for Section 5.6 where this issue is much more important.

Thirdly, the capped nanotubes studied here are charged molecules, owing to the fact that the orbitals have been filled up to the Fermi level of the corresponding infinite tube. This survey has not explicitly considered electron-electron interactions. Unfortunately, any adequate investigation into these

effects is likely to require a more sophisticated approach and certainly more computational resources than were available for this work. However, an examination of the number of excess electrons for each of the caps studied in this thesis shows that the build up of charge is typically small. From Eqn. 4.6, the number of electrons is given by the trace of the density matrix; in every case the number of excess electrons is six or less, which even for the smallest tubes would only increase the number of valence electrons by around 0.75 %.

A more detailed study of the charge distribution was carried out on two selected tubes using direct diagonalisation. As a preliminary, Figure 5.17 shows the predicted charge distribution for C_{84}:22(D_2), which is one of the two experimentally observed isomers of C_{84}. Figure 5.18 shows the charge distribution for the optimum cap for the (11,2) tube, (11,2):C_1(78), and Figure 5.19 shows a cap for the (11,2) tube which contains adjacent pentagons (the first isomer generated by the method of Brinkmann *et al.*[39]). As C_{84}:22 (D_2) is a neutral molecule, the charge on each atom fluctuates between a small positive charge and a small negative charge. For the capped nanotubes, the increased charge is expected to affect the caps rather than the main body of the tubes. This is because the caps contain all the pentagons and a hypothetical isolated pentagon of carbon atoms would have more bonding orbitals than anti-bonding ones and would therefore more easily be able to accommodate the excess electrons. Indeed Figures 5.18 and 5.19 justify this assumption as the charge fluctuations soon die away as one moves down the tube.



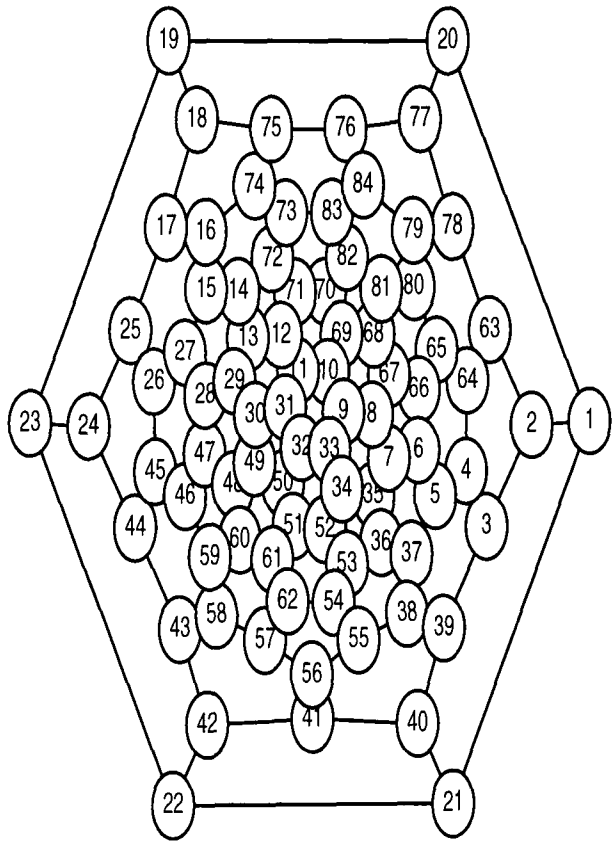


Figure 5.17: Charge distribution for C_{84}:22 (D_2), which is one of the two experimentally observed isomers of C_{84}. The Schlegel diagram shows the numbering scheme for the atoms.

Figure 5.20 confirms the expectation that the excess charge will be associated with the capping pentagons. It shows all the atoms in the $C_1(78)$ cap for the (11,2) tube which have a charge in excess of -0.05 e. These atoms correspond to all the vertices of all the pentagons.

A comparison of Figures 5.18 and 5.19 shows that the magnitude of the atomic charges is greater for the cap with adjacent pentagons. The important point arising from these charge distributions is that the charges per atom for the IPR cap are of the same order of magnitude as those for the atoms in C_{84}:22 (D_2), which is an IPR fullerene.

Obviously, more detailed calculations are necessary to fully understand the effect of the excess electrons but, nevertheless, since the charge fluctuations for the IPR capped nanotube are similar to those for the IPR fullerene isomer (where the Ho Hamiltonian results are known to agree with experiment), this work suggests that these considerations may not be as important as one might think. Therefore, there is good reason to believe the results of this survey of nanotube caps are reliable.

Finally, since STM measurements are available for a capped (11,2) tube,[16,91] the results of this survey can be compared directly with experiments in this case. This will be dealt with next.

5.6 Density of states

The primary reason why the (11,2) tube was included in this study, and indeed, one of the major motivations for this survey of nanotube caps, was

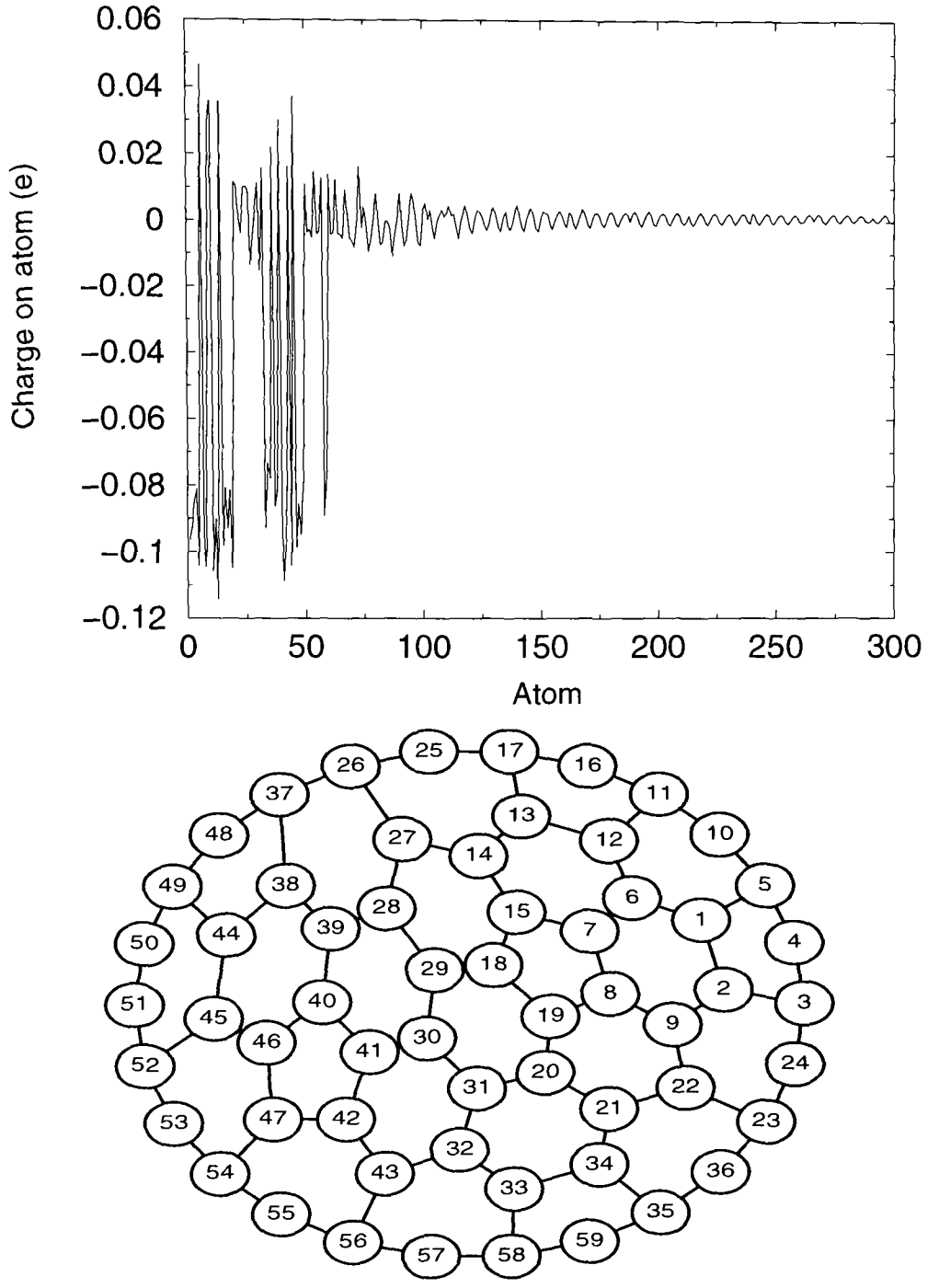
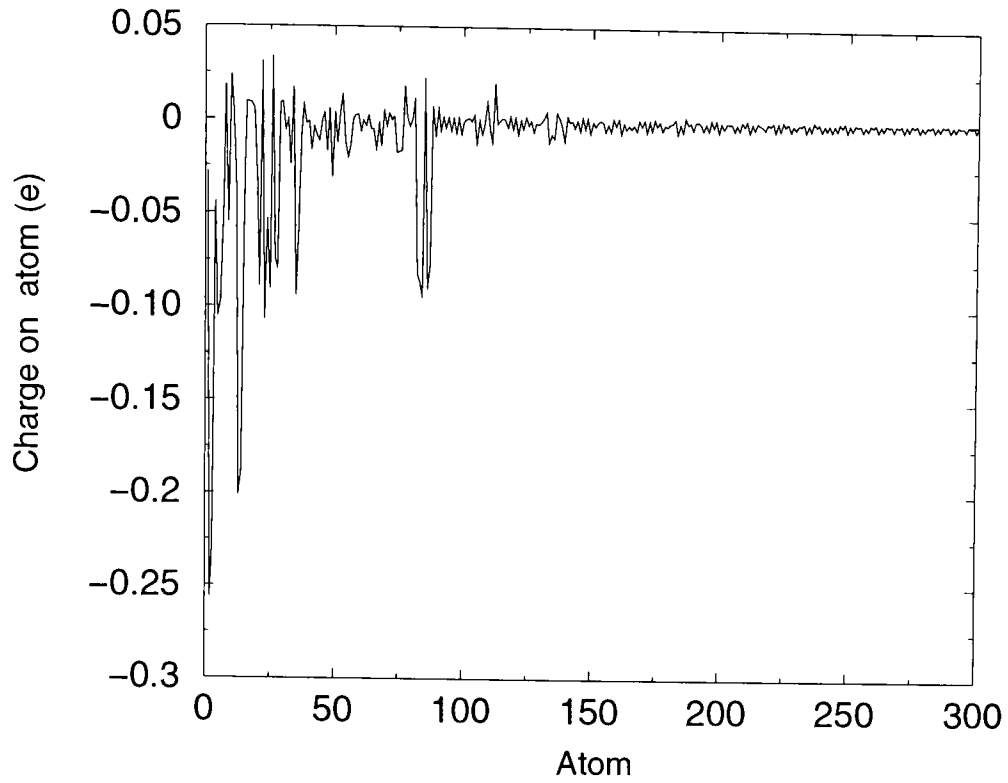


Figure 5.18: Charge distribution for the first 300 atoms of an (11,2) tube with the optimum cap, $C_1(78)$. The nanotube has 47 rings and a total of 1340 atoms. The Schlegel diagram depicts the cap and shows the numbers of the atoms in the cap region.



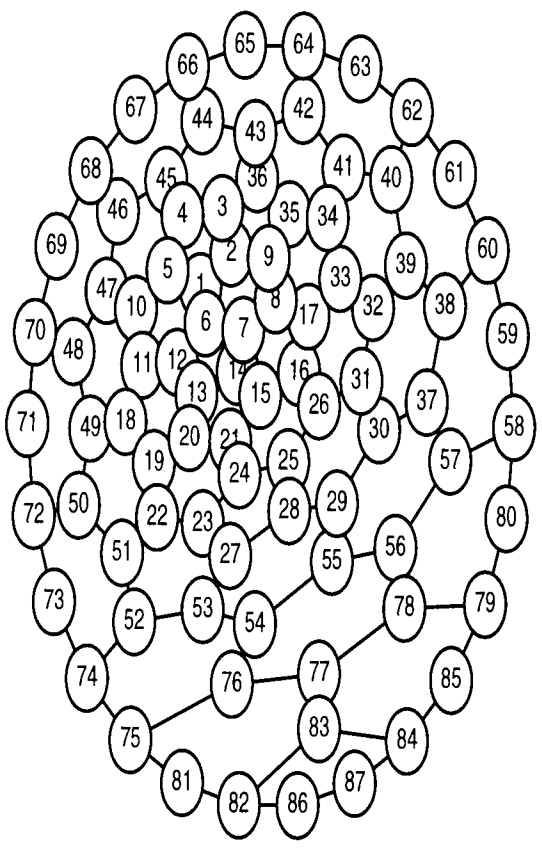


Figure 5.19: Charge distribution for the first 300 atoms of an (11,2) tube with a cap with adjacent pentagons. The nanotube has 47 rings and a total of 1396 atoms. The Schlegel diagram depicts the cap and shows the numbers of the atoms in the cap region.

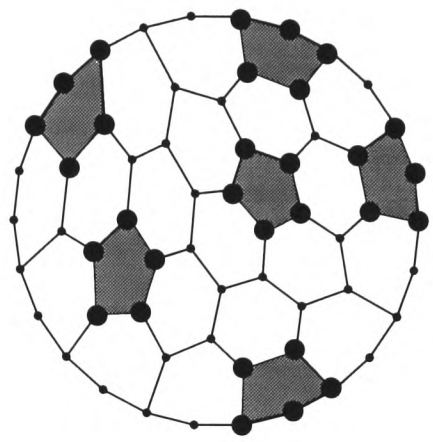


Figure 5.20: Atoms in the $C_1(78)$ cap of the (11,2) tube with a charge of less than -0.05 e (according to the orthogonal tight-binding Hamiltonian of Ho *et al.*).

the existence of STM measurements along a capped (11,2) tube. These STM measurements can be compared with calculated densities of states. Therefore, I will now discuss the predicted densities of states for capped nanotubes, particularly the (11,2) tube. However, I shall start by reviewing the densities of states for infinite nanotubes.

5.6.1 Infinite nanotubes

The density of states (DOS) for an infinite tube is simply the number of electronic states with an energy between E and $E + \mathrm{d}E$. The band structure of the nanotube was determined (see Chapter 3) for a large number of values of κ and hence the DOS was deduced via a histogram. Figure 5.21 shows the DOS for the (5,5), (10,0) and (11,2) tubes. The metallic and semi-conducting natures of the (5,5) and (10,0) tubes respectively are clearly apparent from the DOS. The band gap for the (11,2) tube (= 0.04 eV), which arises from

$\sigma - \pi$ mixing, is much smaller than the (10,0) band gap (= 0.59 eV), thus a finer grid of points and a much increased scale is necessary to observe it. Nevertheless, the Ho Hamiltonian does predict a band gap, as can be seen from the inset in Figure 5.21. Each of the DOS for the infinite tubes also shows van Hove singularities.[16,91] These are the sharp peaks in the DOS plots that correspond to energies at which the slopes of the dispersion curves $\mathrm{d}E(\kappa)/\mathrm{d}\kappa$ vanish, and they are characteristic of 1-D systems.

The DOS of a nanotube can be measured using STM.[13–16,91] STM experiments measure the current I as a function of voltage V. From this information the conductance $\mathrm{d}I/\mathrm{d}V$ may be obtained numerically. The significance of this for the present thesis is that the LDOS in the region of the STM tip is proportional to $(V/I)(\mathrm{d}I/\mathrm{d}V)$[13,14,16] and hence STM measurements can be compared directly to the DOS calculated here (although strictly speaking one should also consider the interaction of the tube with the tip and the substrate). Figure 5.22 shows the experimental DOS for the (10,0) tube,[91] the calculated DOS using Hückel theory[91] and the DOS calculated using the Ho Hamiltonian. As STM measurements are only sensitive to the states near the Fermi level, we concentrate on these. The STM measurement of the DOS has broader peaks than either of the calculations, but both theoretical calculations give a reasonable approximation to the observed result below the Fermi level. The agreement between the Hückel theory calculation and the experiment is poorer above the Fermi level. Odom *et al.*[91] suggested this was because $\sigma - \pi$ mixing (which is expected to be more important for states above the Fermi level[127]) was ignored. The fact that there is slightly better agreement between the STM measurements and the Ho Hamiltonian

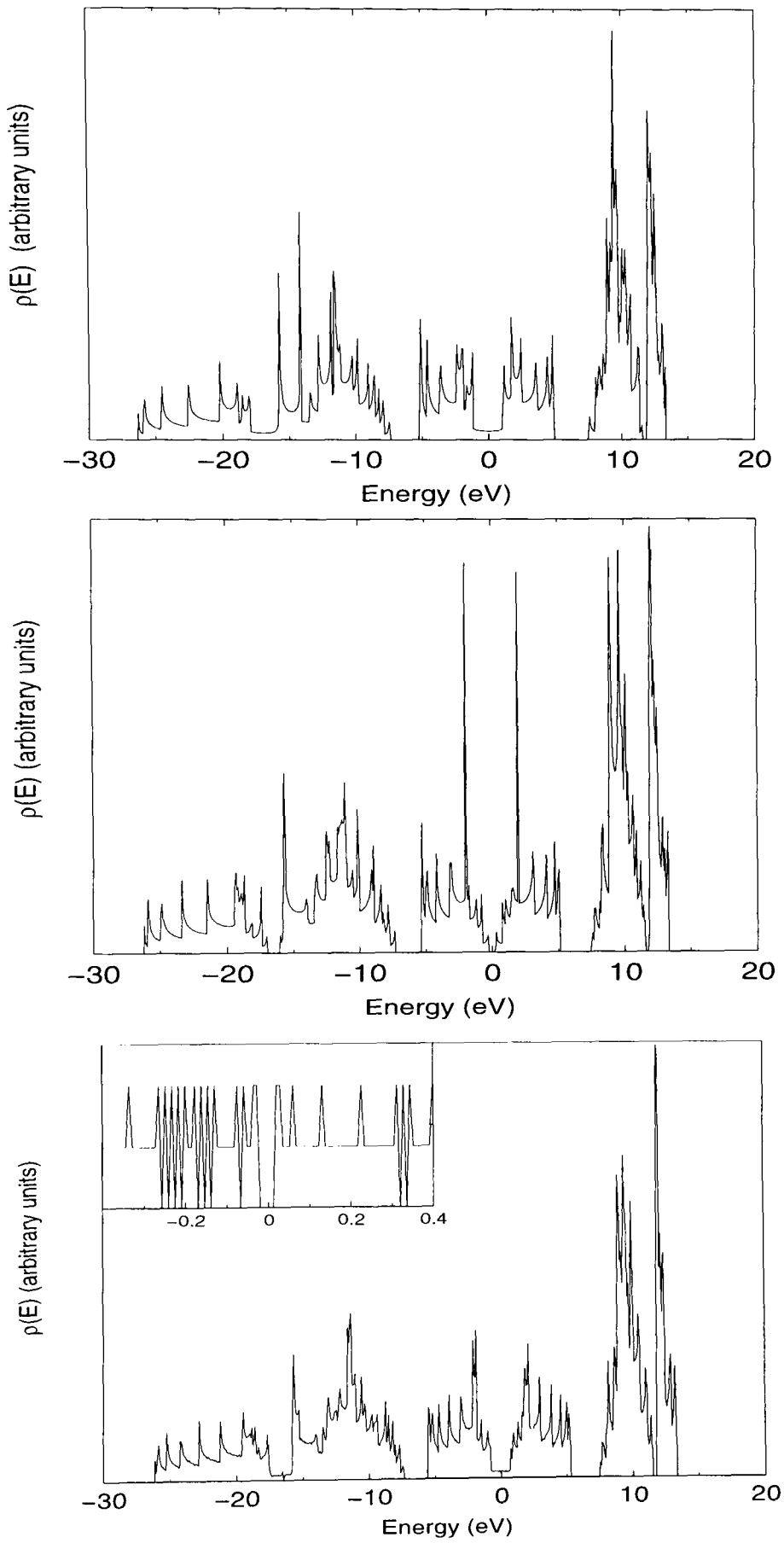


Figure 5.21: DOS for the infinite (5,5) tube (top), the infinite (10,0) tube (middle) and the infinite (11,2) tube (bottom).

calculation confirms this.

5.6.2 Capped nanotubes

In order to simulate the local density of states (LDOS) on atom α for a capped nanotube I have used a doubly capped nanotube (the optimised geometry of the nanotube was obtained using the hybrid density matrix method). Once the eigenvalues and eigenvectors of the Hamiltonian matrix have been obtained by direct diagonalisation, the LDOS is given by

$$\rho_\alpha(E) = \sum_{n=1}^{N_{\text{occ}}} \sum_{i \in \alpha} |c_{in}|^2 \delta(E - E_n), \tag{5.7}$$

where the delta function may be approximated by a narrow Gaussian function:

$$\delta(E - E_n) \approx \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(E - E_n)^2}{2\sigma^2}\right). \tag{5.8}$$

The global DOS is then just

$$\rho(E) = \frac{1}{N_{\text{atom}}} \sum_{\alpha=1}^{N_{\text{atom}}} \rho_\alpha(E). \tag{5.9}$$

This approach assumes that the LDOS and DOS obtained are representative of experimental observations. However, as discussed earlier, even the longest nanotubes considered in this study are many times shorter than experimentally observed tubes. The next section discusses the reasons why the DOS and LDOS presented in this thesis are experimentally relevant.

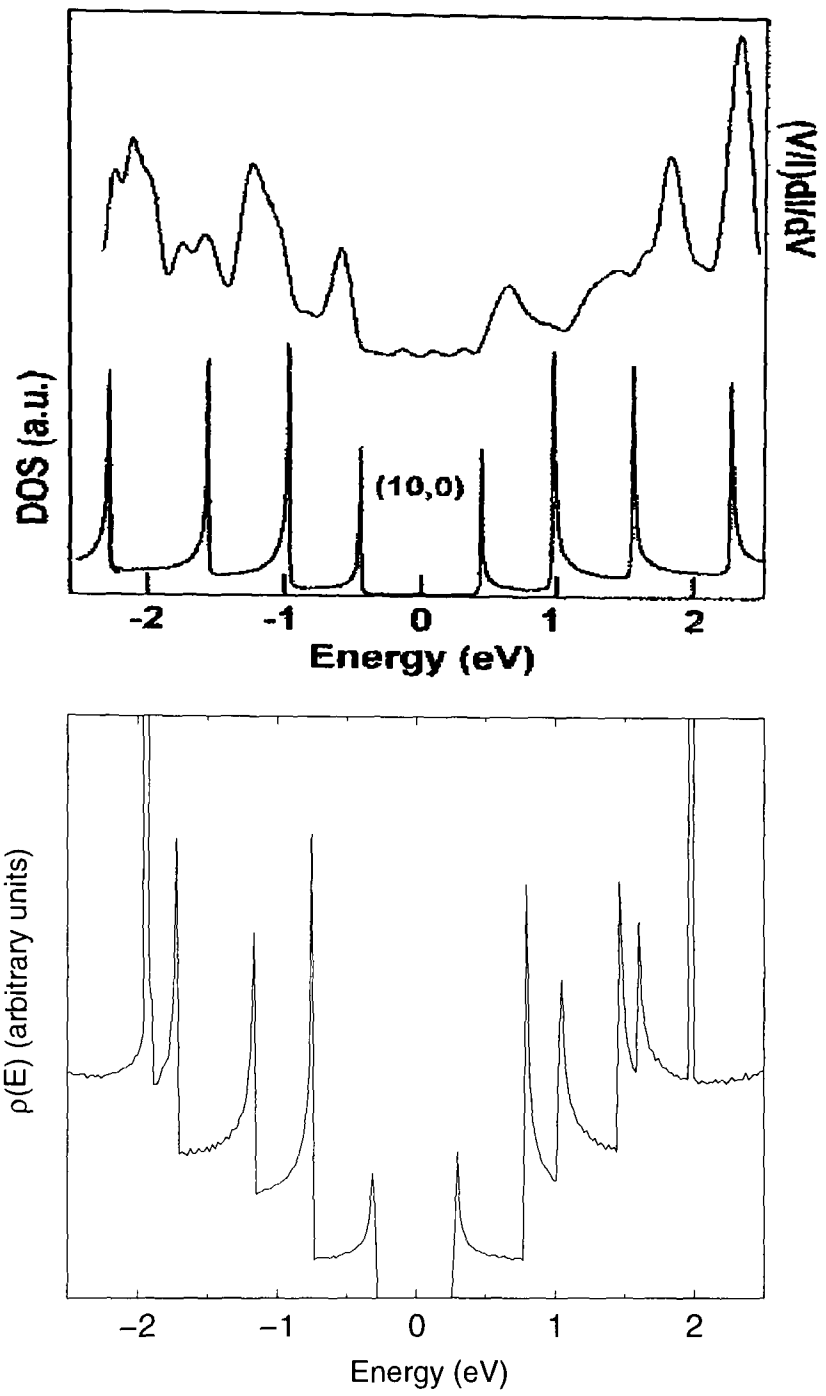


Figure 5.22: STM measurement of the DOS for the (10,0) tube and a Hückel theory simulation[91] of the (10,0) tube (top) and the Ho Hamiltonian DOS for the (10,0) tube (bottom).

Comparison of DOS for finite and infinite nanotubes

There has been some interest in finite length nanotubes,[91,102–105,128,129] mostly owing to the discovery that nanotubes may be broken into shorter segments.[130,131] As we are interested in doubly capped nanotubes, some of the findings of these studies may not be relevant as the theoretical studies often consider only open nanotubes which are usually saturated with hydrogen[102,103,105,128] (although capped nanotubes have also been considered[104,105]). Nonetheless, some of this earlier work has direct relevance to this thesis.

The properties of finite nanotubes obviously differ from infinite nanotubes in some regards. One important example is the band gap (strictly HOMO-LUMO gap for a finite tube, but I shall use the term band gap) which varies with the length of the tube.[102–105,129] In general, the band gap eventually seems to converge to a constant value (the number of atoms involved make calculations on very large nanotubes prohibitively expensive in terms of computational cost). The band gap may oscillate as it converges[102–105] as in Figure 5.23, which shows the variation of the band gap for the (5,5):$C_{5v}(1)$ tube with the number of rings in the finite tube and therefore the length of the tube. Alternatively, some finite tubes show monotonic convergence with respect to length.[102,105] It has been shown that the behaviour of the band gap of an uncapped finite nanotube can be rationalised by examining the band gaps for a series of tubes increased in length by one ring each time and considering the topology of the tubes.[103,129]

Even though the band gap for a finite tube is not necessarily the same as

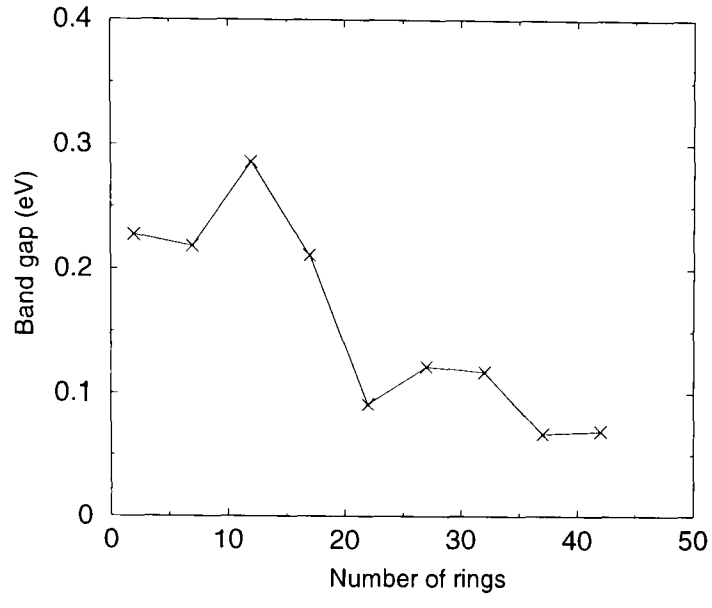


Figure 5.23: Band gaps calculated for varying lengths of (5,5):$C_{5v}(1)$ tube.

for the corresponding infinite tube, these studies have shown that some of the properties of a finite tube are actually not that different to those of an infinite tube of the same helicity. This is especially true as regards the DOS. The general conclusion of these studies is that the overall *shape* of the DOS and position of the peaks are unaffected by the length of the tube.[102,128] This conclusion is supported by the present calculations. The DOS for the middle ring and the total DOS of the doubly capped nanotube (47 rings) are compared to the DOS for the infinite nanotubes for the (5,5):$C_{5v}(1)$, (10,0):$C_{5v}(1)$ and (11,2):$C_1(78)$ nanotubes (the most stable caps in each case) in Figures 5.24, 5.25 and 5.26 respectively. In all three cases, the DOS from the finite tubes are both in good agreement with the DOS for the infinite tube. This is particularly impressive in view of the fact that different methods were used to calculate the DOS of the infinite and finite tubes. Although, the energy range over which the DOS is non-zero is slightly less in the case

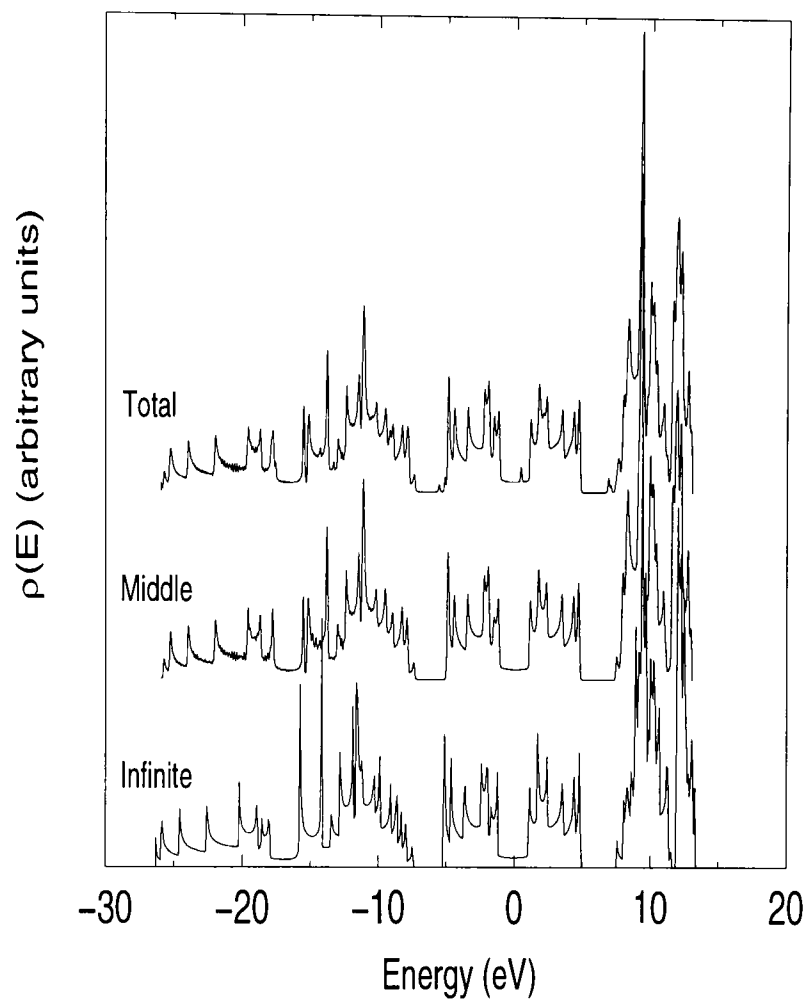


Figure 5.24: DOS of the whole capped tube (total DOS), central ring of the capped tube (middle DOS) and the infinite (uncapped) tube for the (5,5):$C_{5v}(1)$ tube.

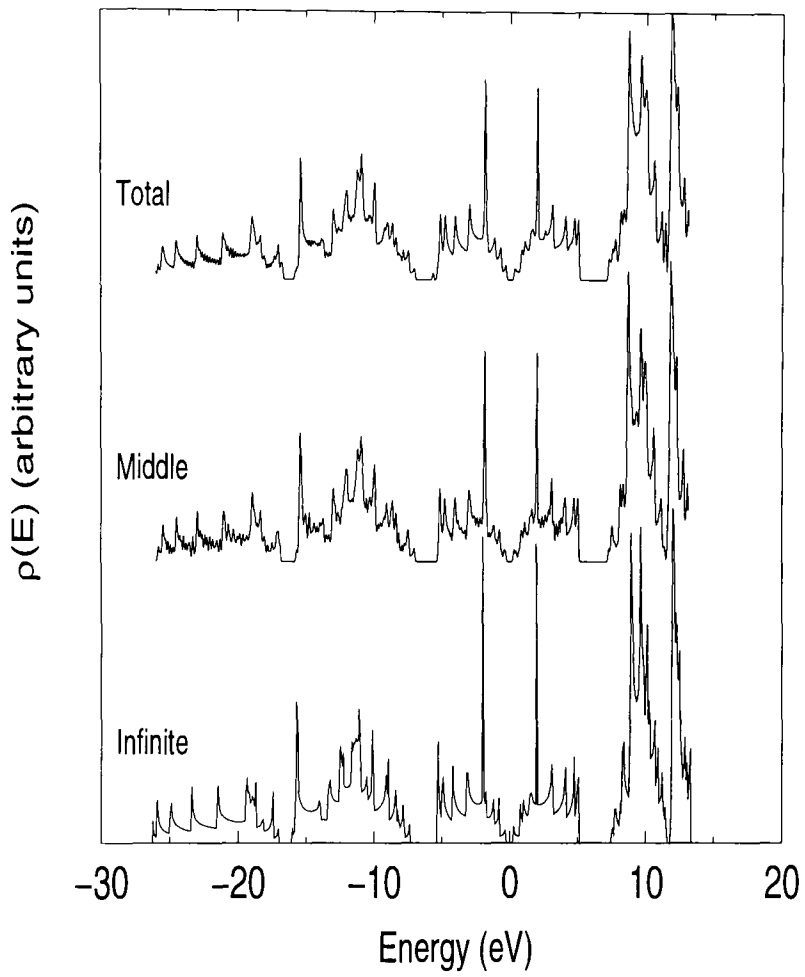


Figure 5.25: DOS of the whole capped tube (total DOS), central ring of the capped tube (middle DOS) and the infinite (uncapped) tube for the (10,0):$C_{5v}(1)$ tube.

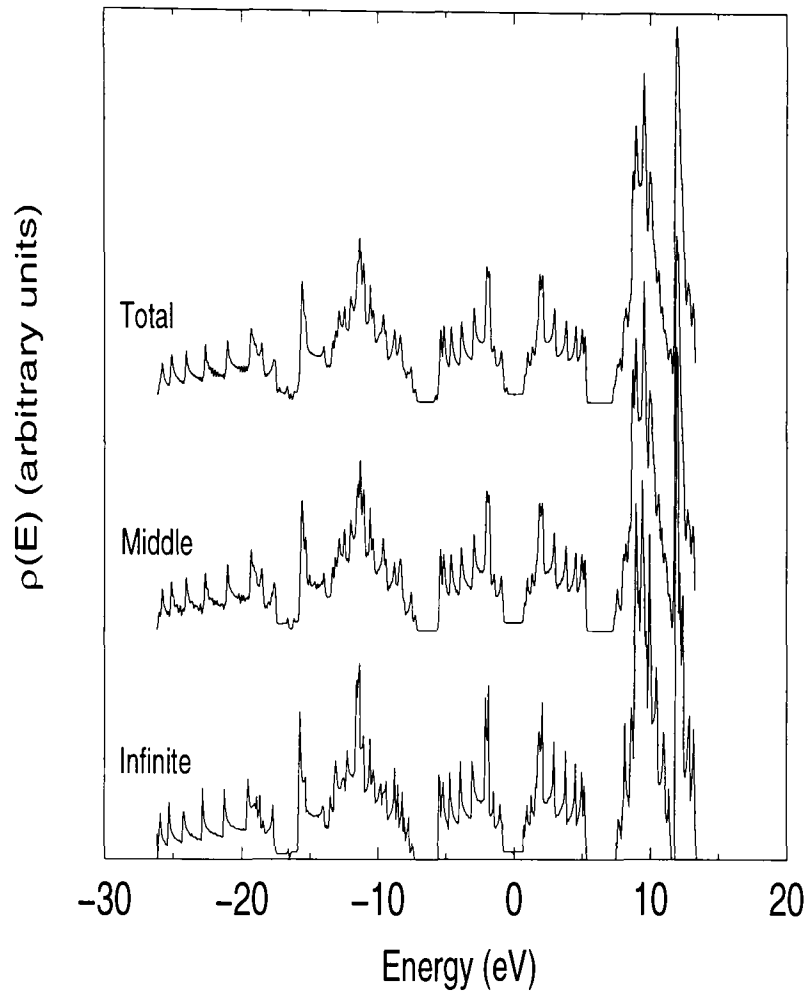


Figure 5.26: DOS of the whole capped tube (total DOS), central ring of the capped tube (middle DOS) and the infinite (uncapped) tube for the (11,2):$C_1(78)$ tube.

of finite tubes, this is not a serious problem as the states near the Fermi level are the most relevant to STM and field emission experiments.

The effect of length on the total DOS was investigated further for the (5,5):$C_{5v}(1)$ tube and the results are shown in Figure 5.27. The overall shape and position of the peaks are remarkably similar for all the tubes, regardless of length. However, Figure 5.28 shows an expanded plot of the DOS close to the Fermi level. The agreement between the infinite tube and the finite tubes is less good in this region.

Cap-induced states

Interestingly, a closer inspection of the total DOS of the (5,5):$C_{5v}(1)$ tube in Figures 5.24, 5.27 and 5.28 reveals the presence of a cap-induced state approximately 0.43 eV above the Fermi level which is not present in either the middle DOS or the infinite DOS (note that the size of this peak decreases with the length of the tube owing to the normalisation of the density of states). Within Hückel theory, four cap-induced states are expected for the (5,5):$C_{5v}(1)$ tube, one below the Fermi level (E_2 symmetry) and three above the Fermi level (A_1, E_1 and E_2 symmetries).[115]

Cap-induced end states can be divided into two distinct categories: localised and resonant states.[91] Resonant states are end states which have the correct symmetry and energy to connect with a bulk state. This results in an increased amplitude for this state at the cap, but a finite amplitude in the bulk. Localised states, however, do not have the correct symmetry and energy to match to a bulk state and therefore the amplitude of this kind of

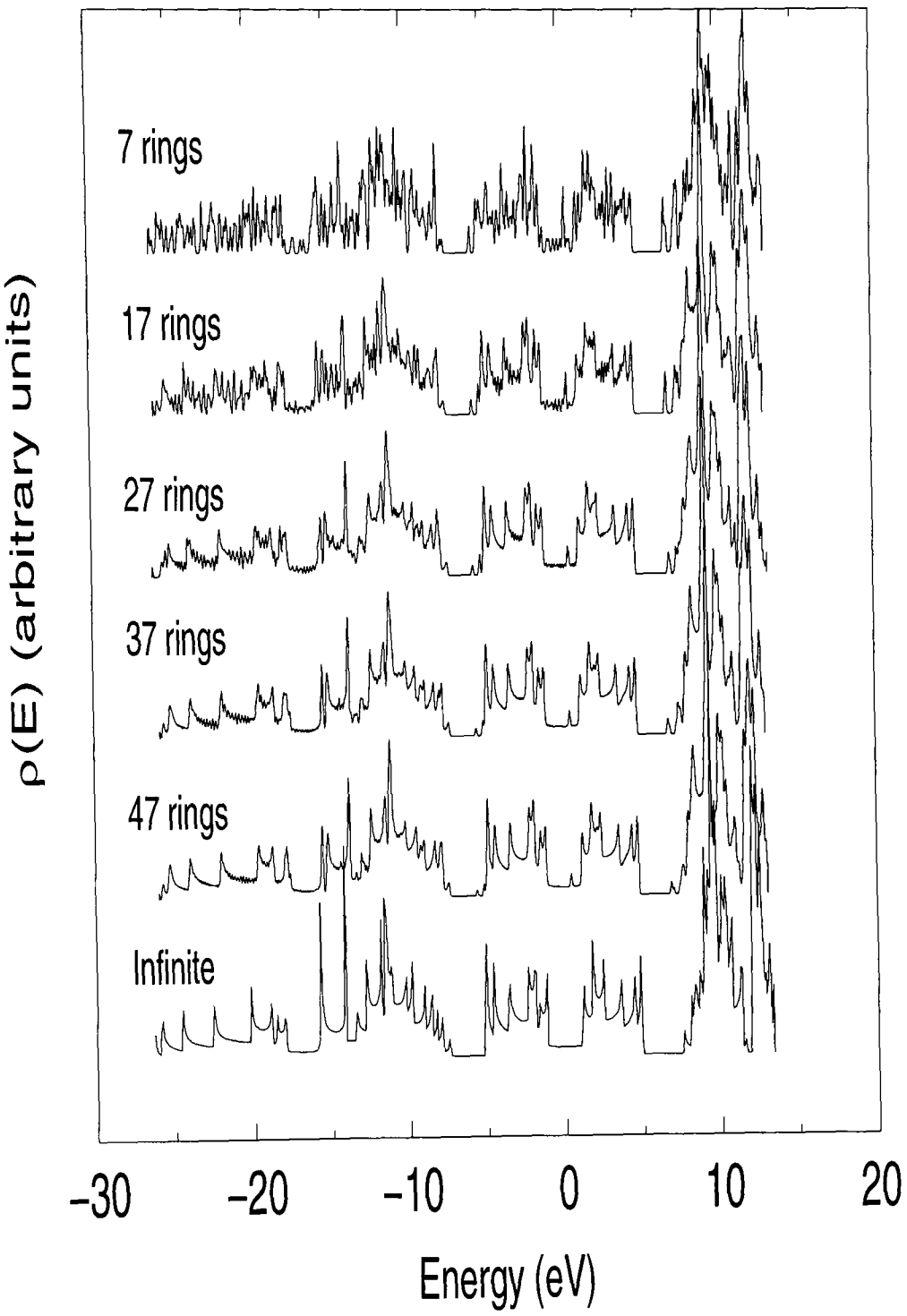


Figure 5.27: Global DOS of the (5,5):$C_{5v}(1)$ tube as a function of tube length.

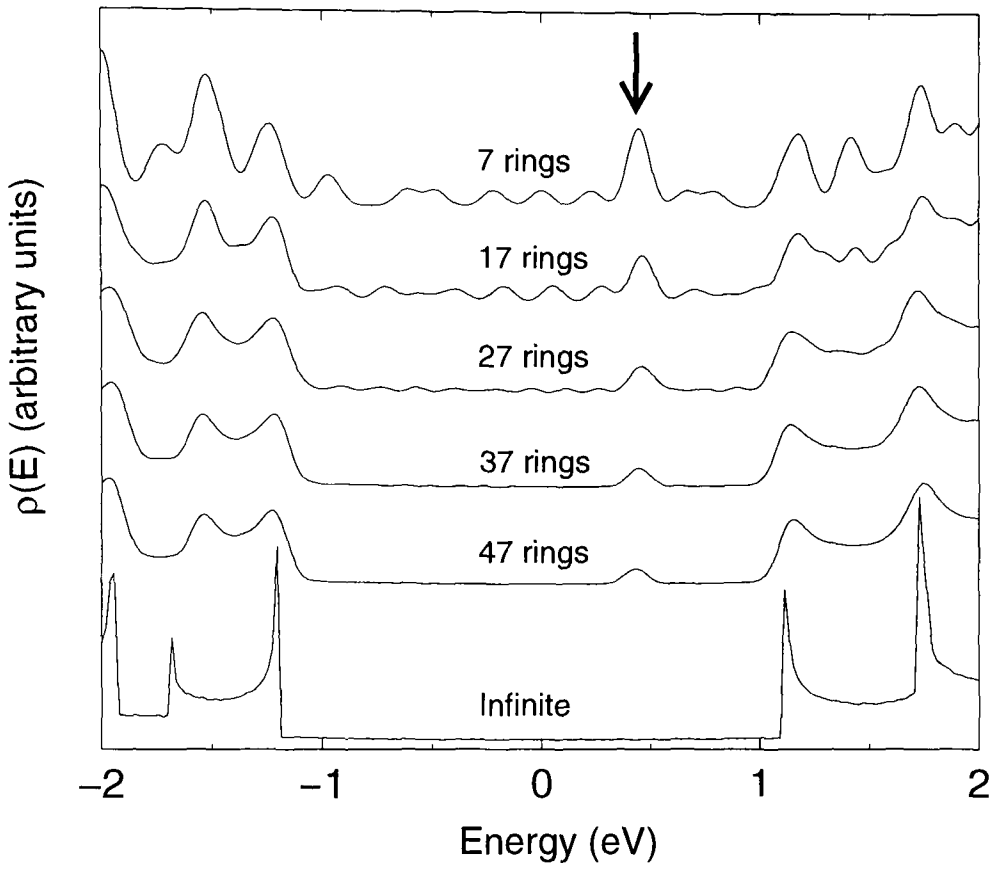


Figure 5.28: Global DOS near the Fermi level for the (5,5):$C_{5v}(1)$ tube as a function of tube length. The cap induced state is marked by an arrow.

state quickly dies away in the bulk.

All of the cap-induced states predicted by Hückel theory for the (5,5):$C_{5v}(1)$ tube are localised.[115] Only one of these states is observed in the present spectra because the Hückel theory study considered different symmetries separately, whereas we have considered all symmetries together and broadened the peaks; thus the other cap-induced states may be hidden by the rest of the spectrum. Therefore, these states will only be observed (with our method) where the DOS is flat (or if the energy of the end state occurs in a band gap). A more detailed analysis would be necessary to assign the symmetry of the visible cap-induced state, but, in view of its proximity to the Fermi level, Hückel theory calculations[115] suggest it is the E_1 state. The presence of localised cap-induced states for the (5,5):$C_{5v}(1)$ tube is interesting in light of the fact that field emission is believed to occur from localised states near the Fermi level.[22,114] In particular, this may partially explain why a study of the (5,5) tube concluded that a capped tube was favourable for field emission.[113]

A cap-induced state is also evident in the total DOS for the (11,2):$C_1(78)$ nanotube shown in Figure 5.26. Figure 5.29 shows a close-up of the region around the Fermi level and the cap-induced peak at about 0.55 eV below the Fermi level is more obvious. Hückel theory calculations have not been carried out for this cap, but given that the symmetry of the cap is C_1, the state must have A symmetry. As all the states of the bulk tube also have A symmetry and the peak does not occur in the band gap, this implies that the cap-induced state for the (11,2):$C_1(78)$ is a resonant state.

The appearance of additional peaks in the DOS for the capped tubes com-

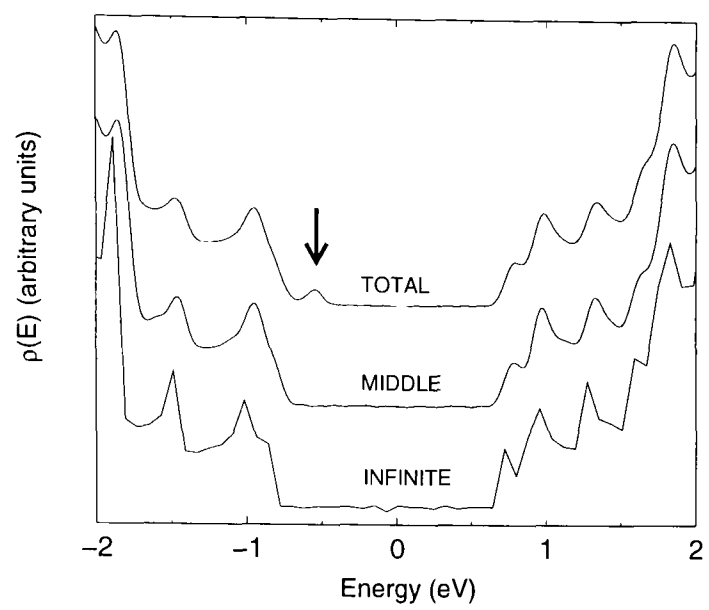


Figure 5.29: DOS near the Fermi level for the (11,2):$C_1(78)$ tube. The cap-induced state in the total DOS is marked by an arrow.

pared with the infinite tubes can be attributed to the presence of the cap and, more specifically, the pentagons.[16,91,121] This has led to the suggestion that STM measurements of the LDOS of the cap may be able to identify the cap. However, this assumes first that different arrangements of pentagons give rise to different LDOS patterns, and second that these differences are sufficiently pronounced to be able to distinguish between the caps in theoretical calculations and unambiguously match the theoretical calculation to the experimental STM signature. A following section comparing the experimental results for the capped (11,2) tube with theoretical predictions arising from our study addresses this. First, however, the influence of $\sigma - \pi$ interactions on the DOS will be examined.

Importance of $\sigma - \pi$ interactions

In common with many theoretical studies of capped nanotubes,[115,117–119,121] Kim *et al.*[16,91] use Hückel theory to obtain their predicted LDOS. It is often said that Hückel theory generates a reasonable approximation to the DOS near the Fermi level because these are mostly π states. Figure 5.30 tests this idea by comparing the DOS for the (11,2) tube calculated using the Ho Hamiltonian and Hückel theory with $\beta = -2.5$ eV (the value used by Kim *et al.* which was obtained from band gap measurements[14,16]) and $\beta = -1.84$ eV (the value obtained from the calculated band gaps shown in Figure 3.10[14]). As one might expect, there is a marked difference between the Hückel calculation with $\beta = -2.5$ eV and the Ho Hamiltonian. However, the Ho Hamiltonian predictions are in good agreement with experiment and we wish, in this instance, to compare the Hückel calculation with the Ho Hamiltonian rather than an experimental observation. Therefore, I shall concentrate on the Hückel model with $\beta = -1.84$ eV. Near the Fermi level, say ± 2 eV, the agreement between the Hückel calculation with $\beta = -1.84$ eV and the Ho Hamiltonian results is fairly good. Outside this range the agreement is less satisfactory. This suggests that it is reasonable to use Hückel theory to calculate the DOS of an infinite nanotube, at least in the region of the Fermi level.

However, the $\sigma - \pi$ interactions are a result of the curvature of the molecule and, since the cap contains the pentagons that induce curvature and hence is more curved, it is not so clear whether this conclusion will hold in the cap region. The lower panel of Figure 5.30 therefore compares the LDOS for the

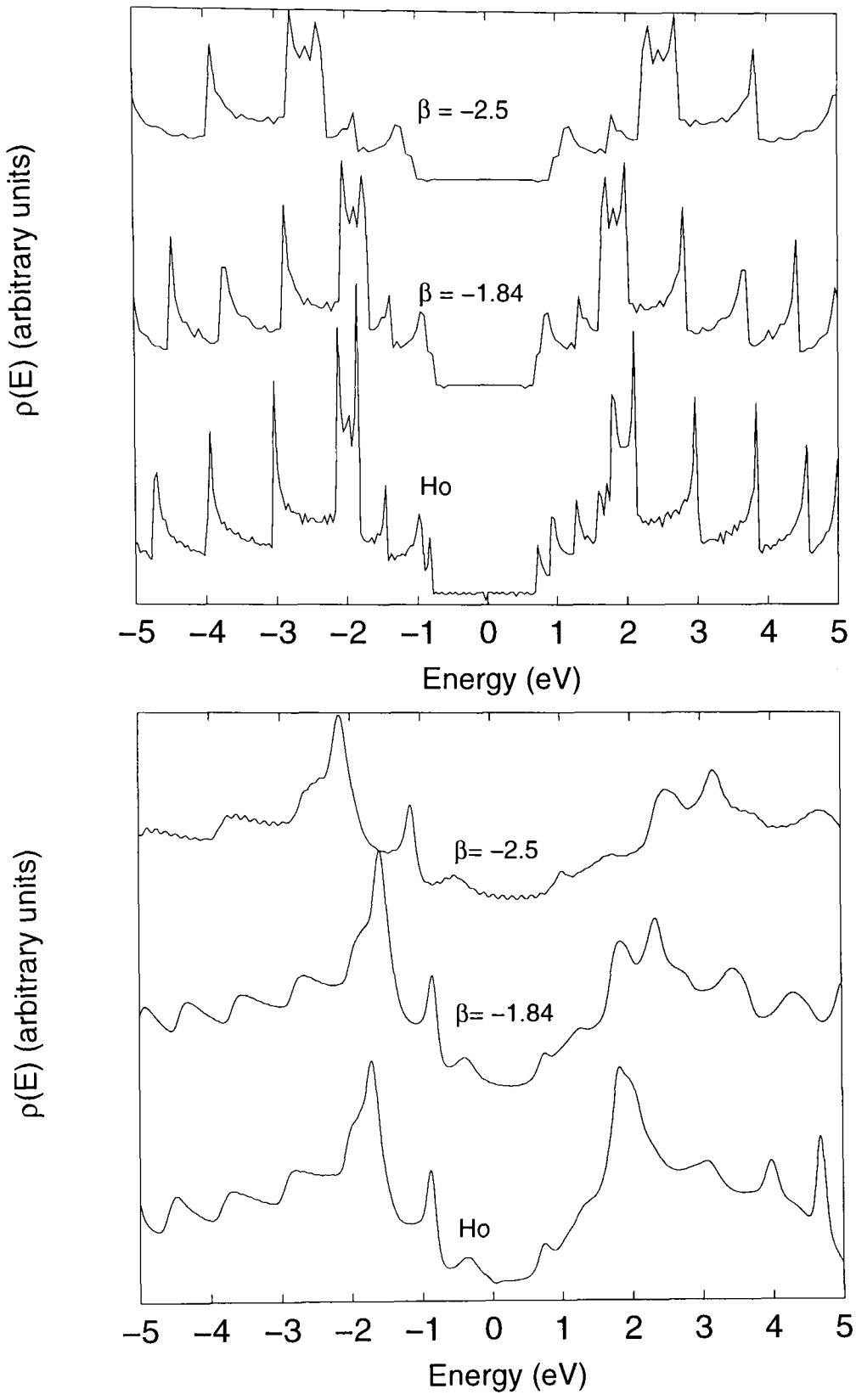


Figure 5.30: Comparison of the global DOS for the infinite (11,2) tube (top) and the LDOS in the cap region for the finite (11,2):$C_1(121)$ tube (bottom). These calculatons use the Ho Hamiltonian and Hückel theory (with $\beta = -1.84$ and -2.5 eV).

(11,2):$C_1(121)$ tube calculated using the Ho Hamiltonian and Hückel theory. The LDOS for the capped tube is the LDOS of the capped region where the capped region is defined as those atoms in the cap output by the census code of Brinkmann *et al.*[39] For the capped tube, the Hückel LDOS ($\beta = -1.84$ eV) is still in good agreement with the Ho Hamiltonian between -2 eV and the Fermi level. Above the Fermi level the two DOS begin to differ significantly around 2 eV, providing evidence that $\sigma - \pi$ interactions are important.[127] Finally, Figure 5.30 also shows the total DOS for the (11,2):$C_1(121)$ tube calculated using Hückel theory with $\beta = -2.5$ eV, for which the agreement is not as good.

In conclusion, $\sigma - \pi$ interactions do affect the DOS of nanotubes, especially in the cap region, but the effect is relatively small near the Fermi level. Nevertheless, in view of the fact that it exists, this thesis will use the Ho Hamiltonian DOS to compare with experiment.

Comparison with measured DOS

Carroll *et al.*,[121] Klusek *et al.*[132] and Kim *et al.*[16,91] have all reported experimental LDOS for capped carbon nanotubes. Unfortunately, neither Carroll *et al.* nor Klusek *et al.* were able to characterise the tube that they were observing and thus their results are of limited value for this thesis (additionally, the tube observed by Carroll *et al.* was multi-walled). However, Kim *et al.* characterised the tube they examined as an (11,2) tube (more precisely, it was the chiral enantiomer of an (11,2) tube, however, this does not affect the DOS). A STM image of the carbon nanotube they examined is shown

in Figure 5.31 along with their experimental measurements and theoretical calculations of the LDOS for three possible caps for the (11,2) tube. One of these caps (model cap III) has adjacent pentagons, the others (model caps I and II) have isolated pentagons. In view of the large number of caps that are possible for this nanotube we were not able to identify precisely which adjacent pentagon cap was used in their study. However, a comparison with the output from the census code revealed that model cap I corresponds to the (11,2):C_1(121) tube while model cap II corresponds to the (11,2):C_1(106) tube. The LDOS calculated by Kim *et al.*[16,91] was obtained using Hückel theory and was for an unspecified region of atoms in each cap. The Hückel theory result obtained for the (11,2):C_1(121) tube in this work is shown in Figure 5.30. The agreement between the LDOS obtained in our study and their study is not perfect. It is difficult to reproduce the results of Kim *et al.* because their papers do not make clear how many atoms were included in the LDOS calculations, nor the method used to broaden the calculated LDOS peaks.

Since this thesis predicts neither of their suggested IPR caps to be the lowest in energy, we also calculated the LDOS for some of the most stable caps for the (11,2) tube using the Ho Hamiltonian. These calculations used $\sigma = 0.05$ and the number of atoms used was the number of cap atoms output by the Brinkmann *et al.* code. Decreasing σ in an attempt to improve the resolution did not significantly change the result (although clearly some of the features narrowed) and introduced spurious oscillations due to the narrower width of the Gaussian function as shown in Figure 5.32.

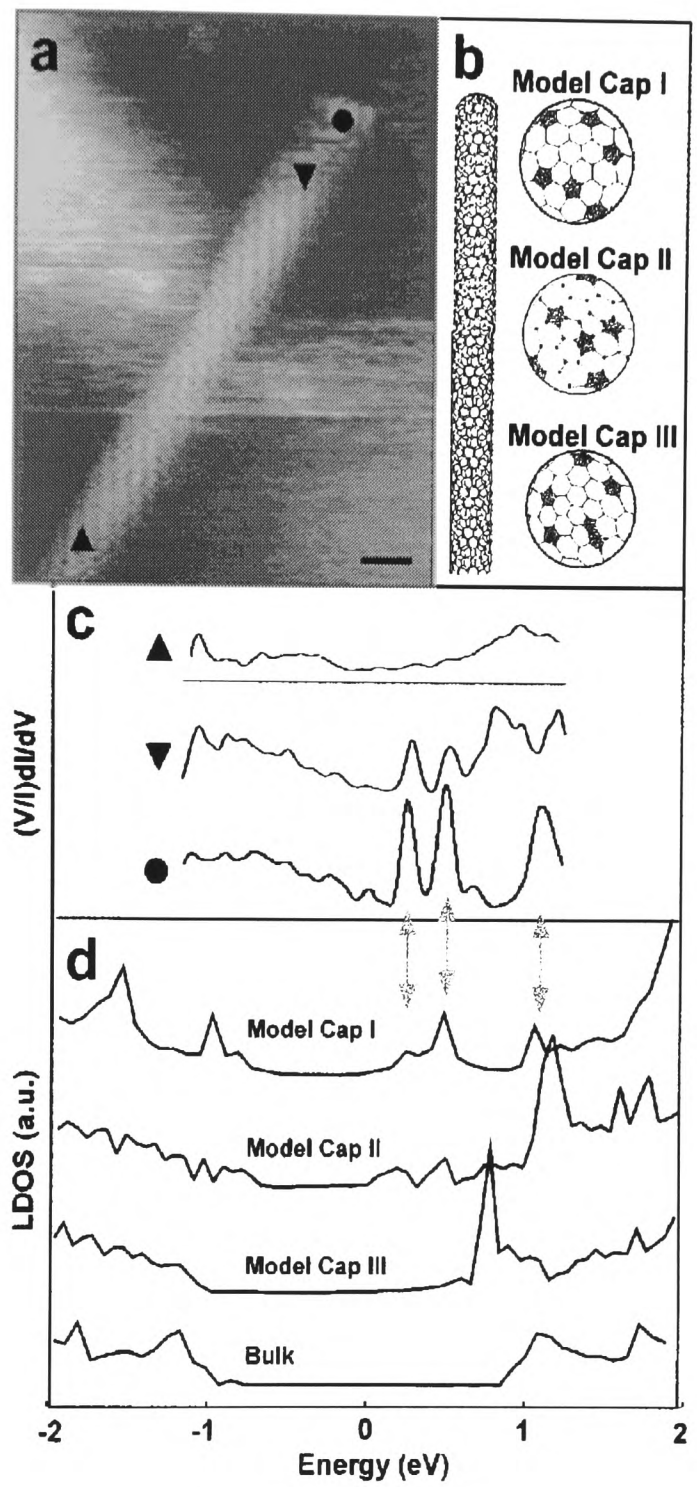


Figure 5.31: Experimental STM measurements for the (11,2) tube and cap (taken from a paper by Odom *et al.*[91]).

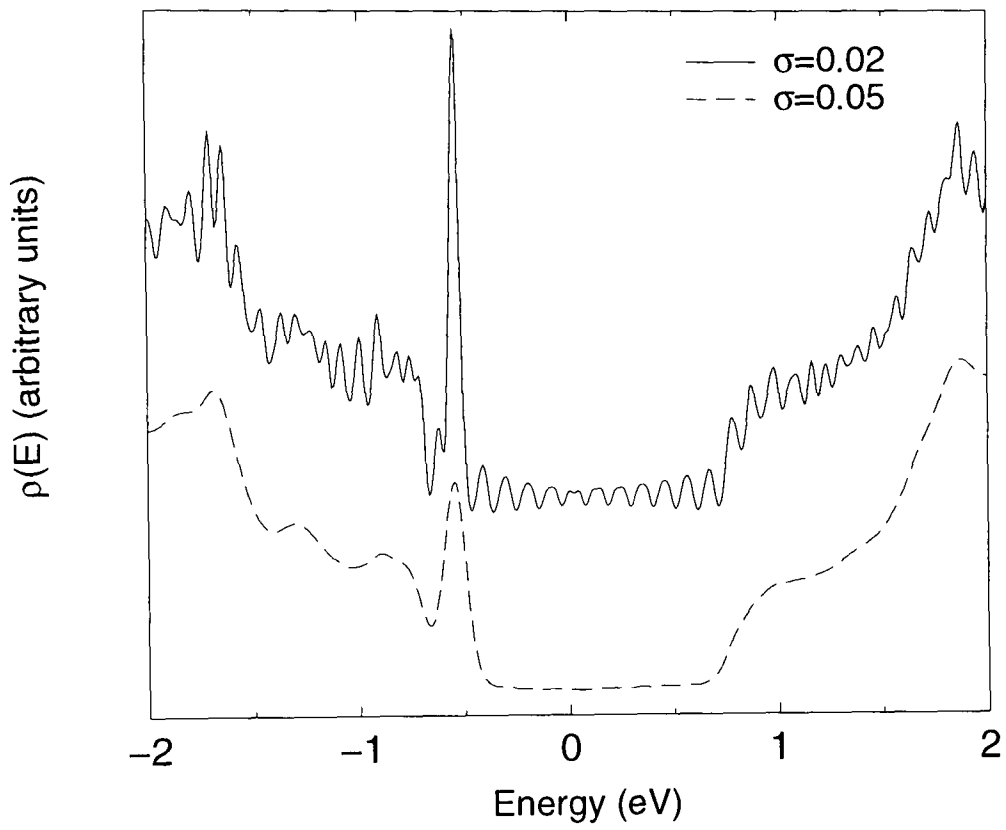


Figure 5.32: LDOS in the cap region for the (11,2):$C_1(78)$ tube, with $\sigma = 0.05$ and $\sigma = 0.02$, as predicted by the Ho Hamiltonian.

Figure 5.33 shows the LDOS for the caps atoms of the five most stable capped tubes (in order of stability): (11,2):C_1(78); (11,2):C_1(110); (11,2):C_1(112); (11,2):C_1(200) and (11,2):C_1(174). The LDOS for the (11,2):C_1(121) tube and for a tube doubly capped with an adjacent pentagon cap (the first cap output by the Brinkmann *et al.* code[39]) are also included. It is clear from Figure 5.33 that the LDOS does change for different caps, justifying previous assertions which have suggested the position of the pentagons determines the LDOS of the cap.[16,121] However a comparison between the calculated LDOS and the measured LDOS shows that none of the suggested caps is really a good match for the experimental data. This leaves us with four possibilities:

1. The differences between the LDOS for various caps are not pronounced enough, nor the STM measurements precise enough, to identify the correct cap.

2. The experimental tube is not capped with any of the caps we (or Kim *et al.*) have considered.

3. The parameters n_1 and n_2 were incorrectly assigned in the experiment.

4. The STM measurements relate to an uncapped nanotube rather than a capped tube.

The third of these possibilities is unlikely. Many of the experimental measurements against which my results for infinite nanotubes have been benchmarked are taken from papers by the same group who carried out this experiment.[14,16,17,91] Unfortunately, the bulk DOS they observed using STM[16,91] (reproduced in Figure 5.31) is not well resolved and it is therefore fruitless to

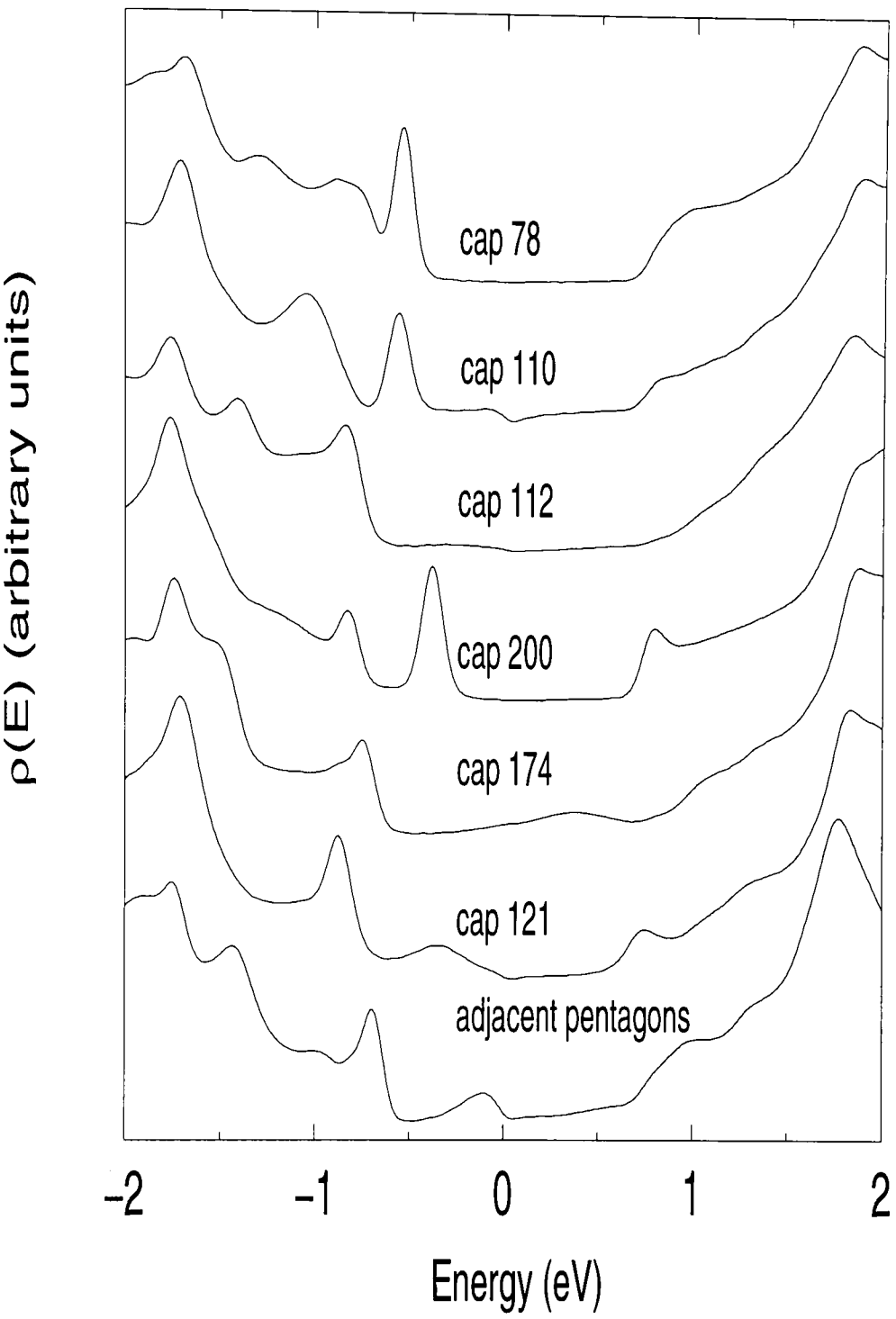


Figure 5.33: LDOS in the cap region for the (11,2) tube with a number of different caps, obtained using the Ho Hamiltonian. In this calculation $\sigma = 0.05$ and the number of atoms used in each case is the number in the cap as output by the census code.[39]

attempt to compare it with the calculated DOS for an infinite (11,2) tube. However, in view of this agreement between the theoretical predictions and the experimental observations for the bulk (see Sections 3.2.5 and 5.6.1) of other tubes and the fact that the bulk of the tube is atomically resolved,[16,133] it is very unlikely that the Hamada indices assigned are incorrect.

The final possibility, that the nanotube may in fact be uncapped, was investigated. It is obviously impossible to consider all possible open nanotube structures. However, the STM image in Figure 5.31 suggests that the nanotube may be terminated with a slanted cut at an angle of about $60 - 70°$ to the tube axis. Three of the four open tubes with LDOS shown in Figure 5.34 are terminated with a slanted cut with an angle roughly in the range $55 - 75°$, and the remaining one is terminated by a straight cut across the tube. The end structures for each of these open (11,2) tubes are shown in Figure 5.35. Each of these open tubes has two strong peaks near the Fermi level, and any of these would be a better match to the measured STM LDOS in Figure 5.31 than the capped nanotubes considered in Figure 5.33.

These results suggest that the (11,2) nanotube observed by Odom *et al.*[16,91] may not have been capped after all. The STM image is not atomically resolved at the end of tube, making it impossible to know from the image alone whether it is capped or not. Odom *et al.* believe that the nanotube is capped because of the rounded shape of the tube end.[16] However, the end could correspond to a slanted cut of the tube. The nanotubes used in the experimental study were produced using laser vaporisation[16,18,91,133,134] and subsequently purified.[16,91,133] It is quite possible that the nanotube

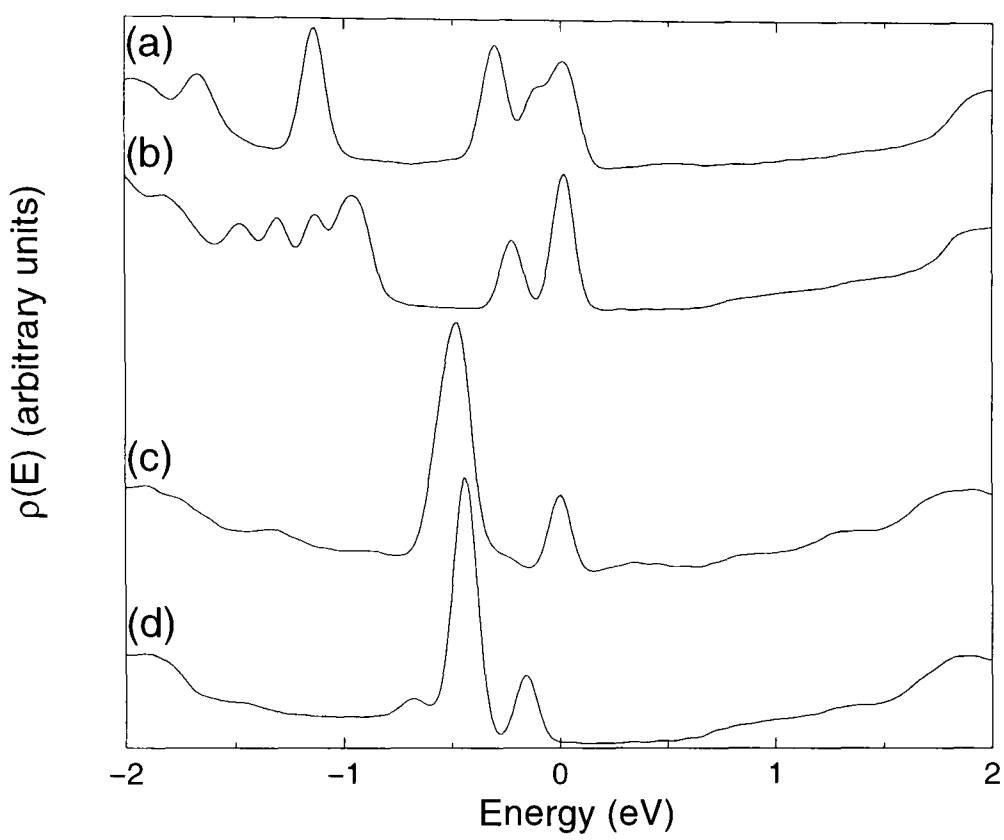


Figure 5.34: LDOS for four open (11,2) nanotubes, each with a different end structure. The lower three curves correspond to tubes terminated by a slanted cut of with an angle roughly in the range $55 - 75°$. The upper curve corresponds to a tube with a flat termination. Figure 5.35 shows the cuts corresponding to each of these tubes. The calculations used $\sigma = 0.05$ and the 50 atoms nearest the end.

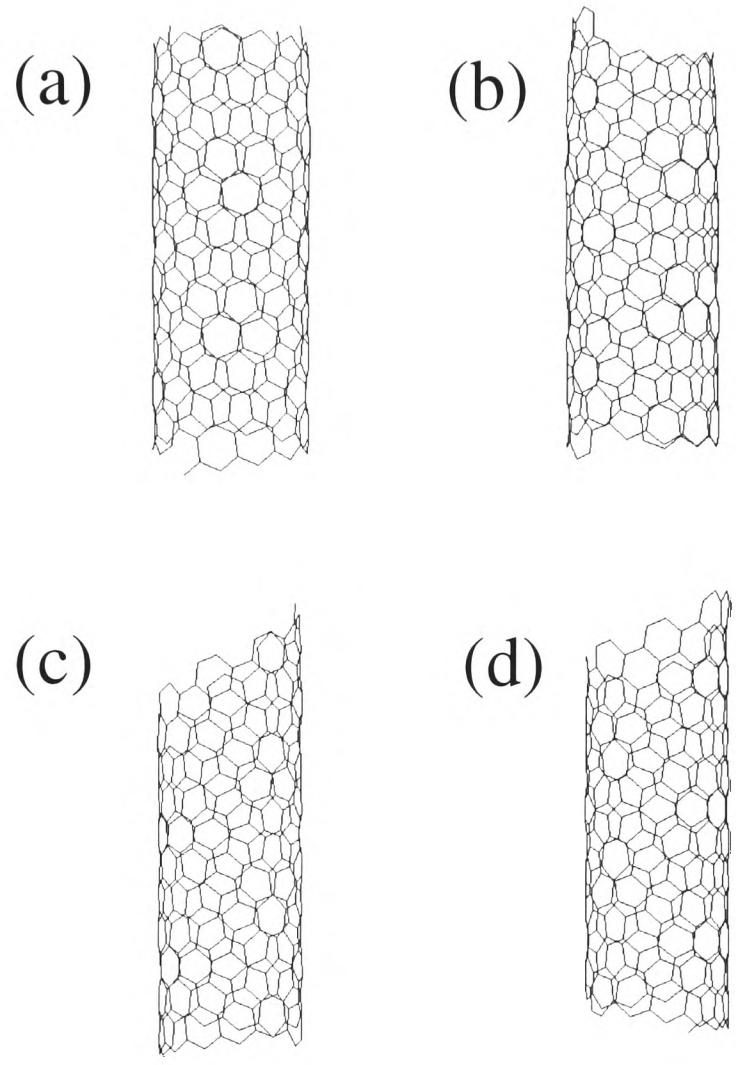


Figure 5.35: Fragments of the open (11,2) tubes used to calculate the LDOS shown in Figure 5.34. The top of the tubes correspond to the open end.

would have been capped when formed and then damaged by the purification procedure,[130] if, indeed, it is uncapped.

It is, however, important to remember that the Ho Hamiltonian is only parameterised for carbon so the open nanotubes we have considered have "dangling" bonds (rather than being saturated by hydrogen). Charge effects have also been neglected. Figure 5.36 shows the charge distribution for two of the open-ended tubes (similar results are obtained for the other two). The magnitude of the charge on the atoms at the open end of the tube is much greater than for the capped nanotubes and therefore charge effects (which are ignored within the Ho Hamiltonian) are likely to be more important than for capped nanotubes.

5.7 Summary

This chapter began by highlighting the need for a systematic study of nanotube caps and then presented such a study for three nanotubes. It was seen that the celebrated isolated pentagon rule still applies to nanotube caps. However beyond the IPR, it is not possible to predict the most stable cap using simple topological criteria such as modified hexagon neighbour index signatures (although there is a definite correlation between the stability of the cap and h_5 for the (11,2) tube). Future studies of capped nanotubes will, perforce, have to rely on full valence tight-binding studies of all possible isolated pentagon caps for the nanotube in question (as is the case for the fullerenes).

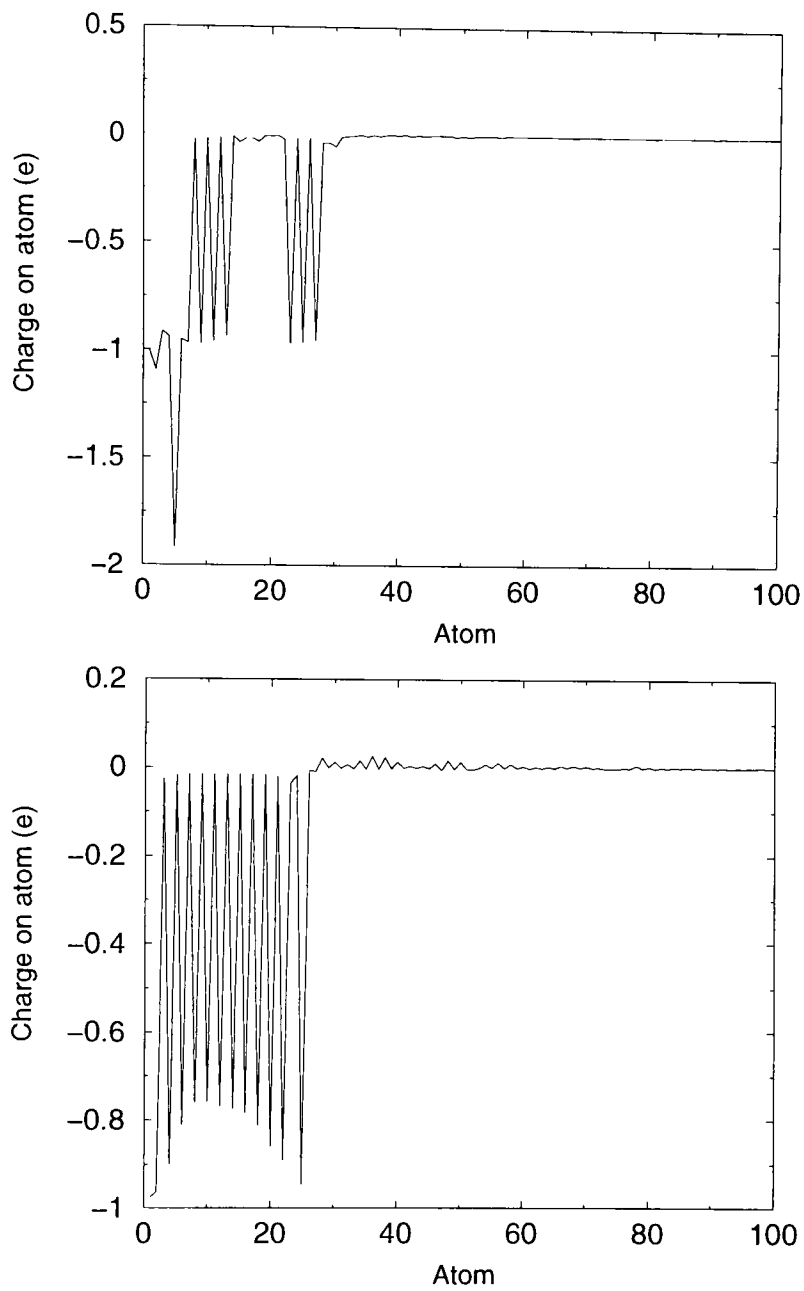


Figure 5.36: Charge distribution for the open ended (11,2) tube with cuts (a) (top) and (b) (bottom) shown in Figure 5.35.

This chapter also discussed the DOS for the nanotubes studied. The experimental and theoretical results were compared for the infinite (10,0) tube and found to be in good agreement. Cap-induced states were observed in the calculated DOS for the (5,5):$C_{5v}(1)$ and the (11,2):$C_1(78)$ tubes. The predicted LDOS for the five most stable caps for the (11,2) tube were compared with the experimental results. The STM measurements were not well reproduced by any of the caps considered by us or Kim *et al.* In fact the best match was obtained for an open nanotube. It may be that, although the LDOS is sensitive to changes in the cap structure, the differences are not pronounced enough to be able to reliably assign the identity of the cap from an STM measurement of the LDOS or even to rule out the possibility that the tube studied in the experiment was in fact open-ended. Alternatively, the observed nanotube cap may not be one of the most stable caps for the tube.

Chapter 6

Conclusions

This thesis began by discussing some of the applications which hope to exploit the interesting properties of carbon nanotubes. It was pointed out that nanotubes are often capped and that the caps may influence some of the properties of nanotubes, especially those such as field emission where the end of the tube is known to be important. However, there exists a large number of possible caps for a given nanotube, and no previous study of capped nanotubes has considered which caps are most likely to be found experimentally. These considerations motivated the study of capped nanotubes carried out in this thesis.

The tight-binding method was chosen to investigate the electronic structure of the nanotubes. This was introduced in Chapter 2, and two parameterisations for carbon were reviewed. It was argued that the Ho Hamiltonian was the most appropriate choice for this study. The suitability of the Ho Hamiltonian was further demonstrated by applying it to fullerenes and infi-

nite nanotubes in Chapter 3. Both fullerenes and infinite nanotubes can be viewed as limiting cases of capped carbon nanotubes and the Ho Hamiltonian results for both classes of material were found to be in good agreement with both other (more sophisticated) theoretical calculations and experimental observations.

The large number of atoms in any but the smallest of capped carbon nanotubes makes traditional diagonalisation techniques, which scale as $O(N^3)$, prohibitively expensive. Therefore, Chapter 4 concentrated on three linear scaling $[O(N)]$ methods based on the density matrix: the LNV, PM and hybrid methods. The underlying ideas were examined and tight-binding calculations for capped carbon nanotubes were shown to scale linearly. The hybrid method was chosen for this thesis because it is robust, fast and variational (allowing the forces for the geometry optimisation to be obtained using the Hellmann-Feynman theorem).

Chapter 5 applied the hybrid density matrix method to capped carbon nanotubes. Three nanotubes were studied: the (5,5) tube (an example of a metal), the (10,0) tube (an example of an insulator) and the (11,2) tube (for which experimental data is available). The lowest energy cap for each nanotube was found to be as follows: the $C_{5v}(1)$ cap for the (5,5) tube, the $C_{5v}(1)$ cap for the (10,0) tube and the $C_1(78)$ cap for the (11,2) tube. The results for the (5,5) and (10,0) nanotubes clearly demonstrated that the isolated pentagon rule is valid for capped carbon nanotubes. The predicted energy penalty per fused pentagon pair was also similar to that found for fullerenes. However, no equally general and reliable rule was found for the

stability of IPR caps. The energy differences in this case are much more subtle and therefore more difficult to predict, something which is also true of the fullerenes. However, in the case of the (11,2) tube there was a clear correlation between h_5 (the number of hexagons adjacent to five hexagons and one pentagon) and the stability of the IPR caps. The larger the value of h_5 for a given cap, the more stable that cap is, suggesting that caps where the pentagons are further apart are more stable. This trend was not observed for the (10,0) caps, but the data in that case may be insufficient to demonstrate a correlation (because there are only seven IPR caps).

Having established the lowest energy caps for each of the three nanotubes, Chapter 5 went on to compare the density of states of infinite and capped nanotubes with other calculations and experimental data. A localised cap state was found for the (5,5):$C_{5v}(1)$ tube and a resonant state was observed for the (11,2):$C_1(78)$ tube. The experimental LDOS in the cap region of what was believed to be a capped (11,2) tube was compared with several LDOS calculated for the five lowest energy caps predicted by our study. The LDOS did indeed vary with the structure of the cap, however, no good match was found between the experimental data and any of the capped tubes considered. In fact, the LDOS obtained via STM bore some resemblance to the LDOS for the open (*i.e.* uncapped) (11,2) nanotubes which were also considered. We believe that this suggests that the experimental STM measurements are not yet powerful enough to unambiguously identify the atomic structure of a nanotube tip.

Appendix A

Hybrid method

In this appendix, the statements made about the purification path and the LNV gradient in Chapter 4 are justified.

A.1 Displacements from the idempotency surface

Recall that Chapter 4 stated that tangent planes to the idempotency surface could be defined as follows. For any point $\boldsymbol{P}$ on the idempotency surface, one can construct a series of lines $\boldsymbol{P}' = \boldsymbol{P} + \alpha\boldsymbol{B}$ where $\boldsymbol{B}$ is a Hermitian matrix and α is a (scalar) variable. While $\boldsymbol{P}'$ is not, in general, idempotent, some matrices $\boldsymbol{B}$ will result in $\boldsymbol{P}'$ being idempotent to linear order *i.e.* $\boldsymbol{P}'^2 - \boldsymbol{P}' = \alpha^2\boldsymbol{B}^2$. When this is the case, the shortest distance between the idempotency surface and $\boldsymbol{P}'$ is of order α^2 and the straight line is considered to lie in the

tangent plane to the idempotency surface at $\boldsymbol{P}$. This concept of a tangent plane will help us to discuss the rationale behind the hybrid method.

Let us assume that $\boldsymbol{P}$ and $\boldsymbol{Q}$ are points on the idempotency surface (from which it follows that $\boldsymbol{P} = \boldsymbol{P}^2$ and $\boldsymbol{Q} = \boldsymbol{Q}^2$) where $\boldsymbol{Q} = \boldsymbol{I} - \boldsymbol{P}$. It follows that

$$\boldsymbol{QP} = \boldsymbol{PQ} = 0. \tag{A.1}$$

Any displacement from the idempotency surface can be written as the sum of a displacement parallel to the idempotency surface and a displacement perpendicular to it:

$$\boldsymbol{\delta P} = \boldsymbol{\delta P}_{||} + \boldsymbol{\delta P}_{\perp}, \tag{A.2}$$

where

$$\boldsymbol{\delta P}_{||} = \boldsymbol{P\delta PQ} + \boldsymbol{Q\delta PP}, \tag{A.3}$$

and

$$\boldsymbol{\delta P}_{\perp} = \boldsymbol{P\delta PP} + \boldsymbol{Q\delta PQ}. \tag{A.4}$$

In order to demonstrate that the forms given for $\boldsymbol{\delta P}_{||}$ and $\boldsymbol{\delta P}_{\perp}$ in Eqns. A.3 and A.4 really do correspond to a displacement from the idempotency surface parallel to the tangent plane and perpendicular to it, we shall consider displacements in these directions explicitly.

First, we consider a displacement $\boldsymbol{P'} = \boldsymbol{P} + \alpha\boldsymbol{\delta P}_{||}$,

$$\begin{aligned} \boldsymbol{P'}^2 &= \left(\boldsymbol{P} + \alpha\boldsymbol{\delta P}_{||}\right)^2 \\ &= \boldsymbol{P}^2 + \alpha\left(\boldsymbol{P\delta P}_{||} + \boldsymbol{\delta P}_{||}\boldsymbol{P}\right) + \alpha^2\boldsymbol{\delta P}_{||}^2, \end{aligned} \tag{A.5}$$

which using Eqn. A.1 and the definition for $\boldsymbol{\delta P}_{||}$ given in Eqn. A.3 becomes

$$\begin{aligned} \boldsymbol{P}'^2 &= \boldsymbol{P} + \alpha\boldsymbol{\delta P}_{||} + \alpha^2\boldsymbol{\delta P}_{||}^2 \\ &= \boldsymbol{P}' + \alpha^2\boldsymbol{\delta P}_{||}^2, \end{aligned} \tag{A.6}$$

and hence such a displacement is in the tangent plane.

Next we consider a displacement perpendicular to the idempotency surface $\boldsymbol{P}' = \boldsymbol{P} + \alpha\boldsymbol{\delta P}_{\perp}$,

$$\begin{aligned} \boldsymbol{P}'^2 &= (\boldsymbol{P} + \alpha\boldsymbol{\delta P}_{\perp})^2 \\ &= \boldsymbol{P}^2 + \alpha\left(\boldsymbol{P\delta P}_{\perp} + \boldsymbol{\delta P}_{\perp}\boldsymbol{P}\right) + \alpha^2\boldsymbol{\delta P}_{\perp}^2. \end{aligned} \tag{A.7}$$

Inserting the definition of $\boldsymbol{\delta P}_{\perp}$ given in Eqn. A.4 and using Eqn. A.1 yields

$$\begin{aligned} \boldsymbol{P}'^2 &= \boldsymbol{P} + 2\alpha\boldsymbol{P\delta PP} + \alpha^2\boldsymbol{\delta P}_{\perp}^2 \\ &= \boldsymbol{P} + \alpha\left[(\boldsymbol{P\delta PP} + \boldsymbol{Q\delta PQ}) + (\boldsymbol{P\delta PP} - \boldsymbol{Q\delta PQ})\right] + \alpha^2\boldsymbol{\delta P}_{\perp}^2 \\ &= \boldsymbol{P} + \alpha\left[\boldsymbol{\delta P}_{\perp} + (\boldsymbol{P\delta PP} - \boldsymbol{Q\delta PQ})\right] + \alpha^2\boldsymbol{\delta P}_{\perp}^2 \\ &= \boldsymbol{P}' + \alpha\left(\boldsymbol{P\delta PP} - \boldsymbol{Q\delta PQ}\right) + \alpha^2\boldsymbol{\delta P}_{\perp}^2, \end{aligned} \tag{A.8}$$

which can only be on the tangent plane if $\boldsymbol{P\delta PP} = \boldsymbol{Q\delta PQ} = 0$. However, from Eqn. A.4 this implies that $\boldsymbol{\delta P}_{\perp} = 0$. Therefore, no displacements of the form $\boldsymbol{P}' = \boldsymbol{P} + \alpha\boldsymbol{\delta P}_{\perp}$ lie in the tangent plane. Furthermore, a general displacement of the form $\boldsymbol{P}' = \boldsymbol{P} + \alpha\left(\boldsymbol{\delta P}_{||} + \boldsymbol{\delta P}_{\perp}\right)$ can only lie in the tangent plane if $\boldsymbol{\delta P}_{\perp} = 0$.

A.2 PM path

A point on the PM path that leads to $\boldsymbol{P}$ has the form $\boldsymbol{P}+\alpha\boldsymbol{\delta P}$. The matrix $\boldsymbol{P}'$ is obtained by performing a single McWeeny iteration on $\boldsymbol{P}+\alpha\boldsymbol{\delta P}$,

$$\boldsymbol{P}' = 3\left(\boldsymbol{P}+\alpha\boldsymbol{\delta P}\right)^2 - 2\left(\boldsymbol{P}+\alpha\boldsymbol{\delta P}\right)^3. \tag{A.9}$$

However, the convergence of the PM method is quadratic and therefore the linear term in Eqn. A.9 must vanish *i.e.* $\boldsymbol{P\delta P}+\boldsymbol{\delta P P}-2\boldsymbol{P\delta P P}=0$. Expressing $\boldsymbol{\delta P}$ in terms of its tangential and perpendicular components and then substituting in Eqns. A.3 and A.4 yields the condition $\boldsymbol{\delta P}_{||}=0$. This implies that the PM path leading to $\boldsymbol{P}$ approaches along a direction of the form $\boldsymbol{\delta P}_{\perp}$ which means that it is orthogonal to the tangent plane.

A.3 LNV gradient

The LNV gradient is given by Eqn. 4.18,

$$\frac{\mathrm{d}\Omega}{\mathrm{d}\sigma_{ij}} = 3(\boldsymbol{\sigma H}+\boldsymbol{H\sigma})_{ji} - 2(\boldsymbol{\sigma}^2\boldsymbol{H}+\boldsymbol{\sigma H\sigma}+\boldsymbol{H\sigma}^2)_{ji}. \tag{A.10}$$

If $\boldsymbol{F}_{ji}$ is defined as $\mathrm{d}\Omega/\mathrm{d}\sigma_{ij}$, then when $\boldsymbol{\sigma}$ is idempotent (*i.e.* when $\boldsymbol{\sigma}=\boldsymbol{P}$),

$$\boldsymbol{F} = \boldsymbol{PH}+\boldsymbol{HP}-2\boldsymbol{PHP}. \tag{A.11}$$

Consider a point $\boldsymbol{P}'=\boldsymbol{P}+\alpha\boldsymbol{F}$,

$$\begin{aligned}\boldsymbol{P}'^2 &= \left(\boldsymbol{P}+\alpha\boldsymbol{F}\right)^2 \\ &= \boldsymbol{P}^2+\alpha\left(\boldsymbol{PF}+\boldsymbol{FP}\right)+\alpha^2\boldsymbol{F}^2\end{aligned}$$

$$
\begin{aligned}
&= \boldsymbol{P}^2 + \alpha\left(\boldsymbol{PH} + \boldsymbol{HP} - 2\boldsymbol{PHP}\right) + \alpha^2\boldsymbol{F}^2 \\
&= \boldsymbol{P} + \alpha\boldsymbol{F} + \alpha^2\boldsymbol{F}^2 \\
&= \boldsymbol{P}' + \alpha^2\boldsymbol{F}^2, \qquad \text{(A.12)}
\end{aligned}
$$

and thus $\boldsymbol{P}'$ is a tangent line. Therefore, the LNV gradient (when $\boldsymbol{\sigma} = \boldsymbol{P}$) lies in the tangent plane to the idempotency surface.

Bibliography

[1] H.W. Kroto, J.R. Heath, S.C. O'Brien, R.F. Curl, and R.E. Smalley. C_{60}: Buckminsterfullerene. *Nature*, 318:162, 1985.

[2] R.C. Haddon. Electronic structure, conductivity, and superconductivity of alkali metal doped C_{60}. *Acc. Chem. Res.*, 25:127, 1992.

[3] W. Krätschmer, L.D. Lamb, K. Fostiropoulos, and D.R. Huffman. Solid C_{60}: A new form of carbon. *Nature*, 347:354, 1990.

[4] R. Taylor, J.P. Hare, A.K. Abdul-Sada, and H.W. Kroto. Isolation, separation and characterisation of the fullerenes C_{60} and C_{70}: The third form of carbon. *J. Chem. Soc., Chem. Comm.*, pages 1423–1425, 1990.

[5] R. Taylor. The pattern of additions to fullerenes. *Phil. Trans. R. Soc. Lond. A*, 343:87, 1993.

[6] Y. Chai, T. Guo, C. Jin, R.E. Haufler, L.P.F. Chibante, J. Fure, L. Wang, J.M. Alford, and R.E. Smalley. Fullerenes with metals inside. *J. Phys. Chem.*, 95:7564, 1991.

[7] T. Guo, M.D. Diener, Y. Chai, M.J. Alford, R.E. Haufler, S.M. McClure, T. Ohno, J.H. Weaver, G.E. Scuseria, and R.E. Smalley. Uranium stabilisation of C_{28}: A tetravalent fullerene. *Science*, 257:1661, 1992.

[8] S. Iijima. Helical microtubules of graphitic carbon. *Nature*, 354:56, 1991.

[9] S. Iijima and T. Ichihashi. Single-shell carbon nanotubes of 1-nm diameter. *Nature*, 363:603, 1993.

[10] D.S. Bethune, C.H. Kiang, M.S. de Vries, G. Gorman, R. Savoy, J Vazquez, and R. Beyers. Cobalt-catalysed growth of carbon nanotubes with single-atomic-layer walls. *Nature*, 363:605, 1993.

[11] J.-M. Bonard, J.-P. Salvetat, T. Stöckli, L. Forró, and A. Châtelain. Field emission from carbon nanotubes: Perspectives for applications and clues to the emission mechanism. *Appl. Phys. A*, 69:245, 1999.

[12] A.H.R. Palser. Interlayer interactions in graphite and carbon nanotubes. *Phys. Chem. Chem. Phys.*, 1:4459, 1999.

[13] J.W.G. Wildöer, L.C. Venema, A.G. Rinzler, R.E. Smalley, and C. Dekker. Electronic structure of atomically resolved carbon nanotubes. *Nature*, 391:59, 1998.

[14] T.W. Odom, J.-L. Huang, P. Kim, and C.M. Lieber. Atomic structure and electronic properties of single-walled carbon nanotubes. *Nature*, 391:62, 1998.

[15] S.G. Lemay, J.W. Janssen, M. van den Hout, M. Mooij, M.J. Bronikowski, P.A. Willis, R.E. Smalley, L.P. Kouwenhoven, and C. Dekker. Two-dimensional imaging of electronic wavefunctions in carbon nanotubes. *Nature*, 412:617, 2001.

[16] P. Kim, T.W. Odom, J.-L. Huang, and C.M. Lieber. Electronic density of states of atomically resolved single-walled carbon nanotubes: Van Hove singularities and end states. *Phys. Rev. Lett.*, 82:1225, 1999.

[17] M. Ouyang, J.-L. Huang, and C.M. Lieber. Scanning tunneling microscopy studies of the one-dimensional electronic properties of single-walled carbon nanotubes. *Annu. Rev. Phys. Chem.*, 53:201, 2002.

[18] T. Guo, P. Nikolaev, A. Thess, D.T. Colbert, and R.E. Smalley. Catalytic growth of single-walled nanotubes by laser vaporization. *Chem. Phys. Lett.*, 243:49, 1995.

[19] J.-C. Charlier and S. Iijima. Growth mechanisms of carbon nanotubes. *Topics Appl. Phys.*, 80:55, 2001.

[20] L. Thiên-Nga, J.-M. Bonard, R. Gáal, and L. Forró. Comparison of catalytically grown and arc-discharge carbon nanotube tips. *Appl. Phys. Lett.*, 80:850, 2002.

[21] Y. Saito and S. Uemura. Field emission from carbon nanotubes and its application to electron sources. *Carbon*, 38:169, 2000.

[22] J.-M. Bonard, H. Kind, T. Stöckli, and L.-O. Nilsson. Field emission from carbon nanotubes: The first five years. *Solid-State Electronics*, 45:893, 2001.

[23] J. Kong, C. Zhou, A. Morpurgo, H.T. Soh, C.F. Quate, C. Marcus, and H. Dai. Synthesis, integration, and electrical properities of individual single-walled carbon nanotubes. *Appl. Phys. A*, 69:305, 1999.

[24] R.H. Baughman, A.A. Zakhidov, and W.A. de Heer. Carbon nanotubes – the route toward applications. *Science*, 297:787, 2002.

[25] G. Che, B.B. Lakshmi, E.R. Fisher, and C.R. Martin. Carbon nanotubule membranes for electrochemical energy storage and production. *Nature*, 393:346, 1998.

[26] P. Calvert. A recipe for strength. *Nature*, 399:210, 1999.

[27] H. Dai, J.H. Hafner, A.G. Rinzler, D.T. Colbert, and R.E. Smalley. Nanotubes as nanoprobes in scanning probe microscopy. *Nature*, 384:147, 1996.

[28] G. Binning, H. Rohrer, Ch. Gerber, and E. Weibel. Surface studies by scanning tunneling microscopy. *Phys. Rev. Lett.*, 49:57, 1982.

[29] J.H. Hafner, C.L. Cheung, and C.M. Lieber. Growth of nanotubes for probe microscopy tips. *Nature*, 398:761, 1999.

[30] J.H. Hafner, C.L. Cheung, and C.M. Lieber. Direct growth of single-walled carbon nanotube scanning probe microscopy tips. *J. Am. Chem. Soc.*, 121:9750, 1999.

[31] S.S. Wong, A.T. Woolley, T.W. Odom, J.-L. Huang, P. Kim, D.V. Vezenov, and C.M. Lieber. Single-walled carbon nanotube probes

for high-resolution nanostructure imaging. *Appl. Phys. Lett.*, 73:3465, 1998.

[32] N.S. Lee, D.S. Chung, I.T. Han, J.H. Kang, Y.S. Choi, H.Y. Kim, S.H. Park, Y.W. Jin, W.K. Yi, M.J. Yun, J.E. Jung, C.J. Lee, J.H. You, S.H. Jo, C.G. Lee, and J.M. Kim. Applications of carbon nanotubes to field emission displays. *Diamonds and Related Materials*, 10:265, 2001.

[33] C. Dekker. Carbon nanotubes as molecular quantum wires. *Physics Today*, pages 22–28, May 1999.

[34] Z. Yao, H.W.Ch. Postma, L. Balents, and C. Dekker. Carbon nanotube intramolecular junctions. *Nature*, 402:273, 1999.

[35] P.M. Ajayan and S. Iijima. Capillarity-induced filling of carbon nanotubes. *Nature*, 361:333, 1993.

[36] R.R. Meyer, J. Sloan, R.E. Dunin-Borkowski, A.I. Kirkland, M.C. Novotny, S.R. Bailey, J.L. Hutchison, and M.L.H. Green. Discrete atom imaging of one-dimensional crystals formed within single-walled carbon nanotubes. *Science*, 289:1324, 2000.

[37] M. Wilson and P.A. Madden. Growth of ionic crystals in carbon nanotubes. *J. Am. Chem. Soc.*, 123:2101, 2001.

[38] C.K.W. Adu, G.U. Sumanasekera, B.K. Pradhan, H.E. Romero, and P.C. Eklund. Carbon nanotubes: A thermoelectric nano-nose. *Chem. Phys. Lett.*, 337:31, 2001.

[39] G. Brinkmann, P.W. Fowler, D.E. Manolopoulos, and A.H.R. Palser. A census of nanotube caps. *Chem. Phys. Lett.*, 315:335, 1999.

[40] P.W. Fowler and D.E. Manolopoulos. *An atlas of fullerenes.* International series of monographs on chemistry. OUP, 1995.

[41] C.H. Xu, C.Z. Wang, C.T. Chan, and K.M. Ho. A transferable tight-binding potential for carbon. *J. Phys.: Condens. Matter*, 4:6047, 1992.

[42] J. Tersoff. New empirical model for the structural properties of silicon. *Phys. Rev. Lett.*, 56:632, 1986.

[43] J. Tersoff. New empirical approach for the structure and energy of covalent systems. *Phys. Rev. B*, 37:6991, 1988.

[44] J. Tersoff. Empirical interatomic potential for carbon, with applications to amorphous carbon. *Phys. Rev. Lett.*, 61:2879, 1988.

[45] D.W. Brenner. Empirical potential for hydrocarbons for use in simulating the chemical vapour deposition of diamond films. *Phys. Rev. B*, 42:9458, 1990.

[46] J. Tersoff. Modelling solid-state chemistry: Interatomic potentials for multicomponent systems. *Phys. Rev. B*, 39:5566, 1989.

[47] D.W. Brenner. The art and science of an analytic potential. *Phys. Stat. Sol. (b)*, 217:23, 2000.

[48] R.C. Mowrey, D.W. Brenner, B.I. Dunlap, J.W. Mintmire, and C.T. White. Simulations of C_{60} collisions with a hydrogen-terminated diamond {111} surface. *J. Phys. Chem.*, 95:7138, 1991.

[49] D.H. Robertson, D.W. Brenner, and J.W. Mintmire. Energetics of nanoscale graphitic tubules. *Phys. Rev. B*, 45:12592, 1992.

[50] B.I. Dunlap. Energetics and fullerene fractionation. *Phys. Rev. B*, 47:4018, 1993.

[51] B.I. Dunlap, D.W. Brenner, and G.W. Schriver. Symmetric isomers of $C_{60}H_{36}$. *J. Phys. Chem.*, 98:1756, 1994.

[52] D.G. Pettifor and I.I. Oleinik. Interatomic bond-order potentials and structural prediction. *Submitted to Progress in Materials Science.*

[53] D. Porezag, Th. Frauenheim, Th. Köhler, G. Seifert, and R. Kaschner. Construction of tight-binding-like potentials on the basis of density-functional theory: Application to carbon. *Phys. Rev. B*, 51:12947, 1995.

[54] J.C. Slater and G.F. Koster. Simplified LCAO method for periodic potential problem. *Phys. Rev.*, 94:1498, 1954.

[55] C.M. Goringe, D.R. Bowler, and E. Hernández. Tight-binding modelling of materials. *Rep. Prog. Phys.*, 60:1447, 1997.

[56] P.-O. Löwdin. On the non-orthogonality problem connected with the use of atomic wave functions in the theory of molecules and crystals. *J. Chem. Phys.*, 18:365, 1950.

[57] A.P. Sutton, M.W. Finnis, D.G. Pettifor, and Y. Ohta. The tight-binding bond model. *J. Phys. C: Solid State Phys.*, 21:35, 1988.

[58] W.M.C. Foulkes and R. Haydock. Tight-binding models and density-functional theory. *Phys. Rev. B*, 39:12520, 1989.

[59] H. Hellmann. *Einfuhrung in die quantenchemie.* Franz Deuticke, 1937.

[60] R.P. Feynman. Forces in molecules. *Phys. Rev.*, 56:340, 1939.

[61] C.Z. Wang and K.M Ho. Tight-binding molecular dynamics studies of covalent systems. *Adv. Chem. Phys.*, 93:651, 1996.

[62] L. Goodwin, A.J. Skinner, and D.G. Pettifor. Generating transferable tight-binding parameters: Application to silicon. *Europhys. Lett.*, 9:701, 1989.

[63] L.M. Canel, A.E. Carlsson, and P.A. Fedders. Effect of overlap on semiempirical potentials derived from tight binding. *Phys. Rev. B*, 48:10739, 1993.

[64] B.L. Zhang, C.Z. Wang, and K.M. Ho. Search for the ground-state structure of C_{84}. *J. Chem. Phys.*, 96:7183, 1992.

[65] B.L. Zhang, C.Z. Wang, and K.M. Ho. Structures of large fullerenes: C_{60} to C_{94}. *Chem. Phys. Lett.*, 193:225, 1992.

[66] B.L. Zhang, C.Z. Wang, K.M. Ho, C.H. Xu, and C.T. Chan. The geometry of small fullerene cages: C_{20} to C_{70}. *J. Chem. Phys.*, 97:5007, 1992.

[67] C.H. Xu and G.E. Scuseria. An *O(N)* tight-binding study of carbon clusters up to C_{8640}: The geometrical shape of the giant icosahedral fullerenes. *Chem. Phys. Lett.*, 262:219, 1996.

[68] S.-Y. Qui, C.Z. Wang, K.M. Ho, and C.T. Chan. Tight-binding molecular dynamics with linear system-size scaling. *J. Phys.: Condens. Matter*, 6:9153, 1994.

[69] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical recipes in FORTRAN: The art of scientific computing.* CUP, Second edition, 1992.

[70] C.G. Dare. *Electron repulsion in carbon clusters*, Part II Thesis, Oxford University, 2002.

[71] D.E. Manolopoulos, J.C. May, and S.E. Down. Theoretical studies of the fullerenes: C_{34} to C_{70}. *Chem. Phys. Lett.*, 181:105, 1991.

[72] D.E. Manolopoulos and P.W. Fowler. Molecular graphs, point groups and fullerenes. *J. Chem. Phys.*, 96:7603, 1992.

[73] D.E. Manolopoulos and P.W. Fowler. A fullerene without a spiral. *Chem. Phys. Lett.*, 204:1, 1993.

[74] G. Brinkmann and A.W.M. Dress. PentHex puzzles: A reliable and efficient top-down approach to fullerene-structure enumeration. *Adv. Appl. Maths.*, 21:473, 1998.

[75] H.W. Kroto. The stability of the fullerenes C_n, with n = 24, 28, 32, 36, 50, 60 and 70. *Nature*, 329:529, 1987.

[76] T.G. Schmalz, W.A. Seitz, D.J. Klein, and G.E. Hite. Elemental carbon cages. *J. Am. Chem. Soc.*, 110:1113, 1988.

[77] K. Raghavachari. Ground state of C_{84}: two almost isoenergetic isomers. *Chem. Phys. Lett.*, 190:397, 1992.

[78] G. Brinkmann, O.D. Friedrichs, A. Dress, and T. Harmuth. CaGe – a virtual environment for studying some special classes of large molecules. *Match – Communications in mathematical and in computer chemistry*, 36:233, 1997.

[79] J.R. Colt and G.E. Scuseria. An *ab in initio* study of the C_{76} fullerene isomers. *J. Phys. Chem.*, 96:10265, 1992.

[80] A. Ceulemans, D. Beyens, and L.G. Vanquickenborne. Symmetry aspects of Jahn-Teller activity: Structure and reactivity. *J. Am. Chem. Soc.*, 106:5824, 1984.

[81] D.E. Manolopoulos, P.W. Fowler, R. Taylor, H.W. Kroto, and D.R.M. Walton. An end to the search for the ground state of C_{84}? *J. Chem. Soc., Faraday Trans.*, 88:3117, 1992.

[82] S. Iijima, T. Ichihashi, and Y. Ando. Pentagons, heptagons and negative curvature in graphite microtubule growth. *Nature*, 356:776, 1992.

[83] N. Hamada, S. Sawada, and A. Oshiyama. New one-dimensional conductors: Graphitic microtubules. *Phys. Rev. Lett.*, 68:1579, 1992.

[84] C.T. White, D.H. Robertson, and J.W. Mintmire. Helical and rotational symmetries of nanoscale graphitic tubules. *Phys. Rev. B*, 47:5485, 1993.

[85] J.W. Mintmire and C.T. White. Electronic and structural properties of carbon nanotubes. *Carbon*, 33:893, 1995.

[86] J.K. Burdett. *Chemical bonding in solids.* OUP, 1995.

[87] G.G. Tibbetts. Why are carbon filaments tubular? *J. Crystal Growth*, 66:632, 1984.

[88] J. Molina Molina, S.S. Savinsky, and N.V. Khokhriakov. A tight-binding model for calculations of structures and properties of graphitic nanotubes. *J. Chem. Phys.*, 104:4652, 1996.

[89] R. Saito, M. Fujita, G. Dresselhaus, and M.S. Dresselhaus. Electronic structure of chiral graphene tubules. *Appl. Phys. Lett.*, 60:2204, 1992.

[90] J.W. Mintmire, D.H. Robertson, and C.T. White. Properties of fullerene nanotubules. *J. Phys. Chem. Solids*, 54:1835, 1993.

[91] T.W. Odom, J.-L. Huang, P. Kim, and C.M. Lieber. Structure and electronic properties of carbon nanotubes. *J. Phys. Chem. B*, 104:2794, 2000.

[92] A.D. Daniels and G.E. Scuseria. What is the best alternative to diagonalization of the Hamiltonian in large scale semiempirical calculations? *J. Chem. Phys.*, 110:1321, 1999.

[93] D.R. Bowler, M. Aoki, C.M. Goringe, A.P. Horsfield, and D.G. Pettifor. A comparison of linear scaling tight-binding methods. *Modelling Simul. Mater. Sci. Eng.*, 5:199, 1997.

[94] X.-P. Li, R.W. Nunes, and D. Vanderbilt. Density-matrix electronic-structure method with linear system-size scaling. *Phys. Rev. B*, 47:10891, 1993.

[95] R.W. Nunes and D. Vanderbilt. Generalization of the density-matrix method to a nonorthogonal basis. *Phys. Rev. B*, 50:17611, 1994.

[96] A.H.R. Palser and D.E. Manolopoulos. Canonical purification of the density matrix in electronic-structure theory. *Phys. Rev. B*, 58:12704, 1998.

[97] D.R. Bowler and M.J. Gillan. Density matrices in *O(N)* electronic structure calculations: Theory and applications. *Comp. Phys. Comm.*, 120:95, 1999.

[98] S.Y. Wu and C.S. Jayanthi. Order-N methodologies and their applications. *Physics Reports*, 358:1, 2002.

[99] P. Ordejón. Order-N tight-binding methods for electronic-structure and molecular dynamics. *Computational Materials Science*, 12:157, 1998.

[100] A.D. Daniels, J.M. Millam, and G.E. Scuseria. Semiempirical methods with conjugate gradient density matrix search to replace diagonalization for molecular systems containing thousands of atoms. *J. Chem. Phys.*, 107:425, 1997.

[101] W. Kohn. Analytic properties of Bloch waves and Wannier functions. *Phys. Rev.*, 115:809, 1959.

[102] A. Rochefort, D.R. Salahub, and P. Avouris. Effects of finite length on the electronic structure of carbon nanotubes. *J. Phys. Chem. B*, 103:641, 1999.

[103] L. Liu, C.S. Jayanthi, H. Guo, and S.Y. Wu. Broken symmetry, boundary conditions, and band-gap oscillations in finite single-wall carbon nanotubes. *Phys. Rev. B*, 64:033414, 2001.

[104] J. Cioslowski, N. Rao, and D. Moncrieff. Electronic structures and energetics of (5,5) and (9,0) single-walled carbon nanotubes. *J. Am. Chem. Soc.*, 124:8485, 2002.

[105] H.-Y. Zhu, D.J. Klein, T.G. Schmalz, A. Rubio, and N.H. March. Geometric boundary effects on the electronic properties of finite carbon nanotubes. *J. Phys. Chem. Solids*, 59:417, 1998.

[106] M.S. Daw. Model for energetics of solids based on the density matrix. *Phys. Rev. B*, 47:10895, 1993.

[107] R. McWeeny. Some recent advances in density matrix theory. *Rev. Mod. Phys.*, 32:335, 1960.

[108] P.J.F. Harris. *Carbon nanotubes and related structures: New materials for the twenty-first century.* CUP, 1999.

[109] S. Iijima, P.M. Ajayan, and T. Ichihashi. Growth model for carbon nanotubes. *Phys. Rev. Lett.*, 69:3100, 1992.

[110] C.-H. Kiang and W.A. Goddard. Polyyne ring nucleus growth model for single-layer carbon nanotubes. *Phys. Rev. Lett.*, 76:2515, 1996.

[111] J.-C. Charlier, A. De Vita, X. Blase, and R. Car. Microscopic growth mechanisms for carbon nanotubes. *Science*, 275:46, 1997.

[112] J.-Ch. Charlier, X. Blase, A. De Vita, and R. Car. Microscopic growth mechanisms for carbon and boron-nitride nanotubes. *Appl. Phys. A*, 68:267, 1999.

[113] S. Han and J. Ihm. Role of the localized states in field emission of carbon nanotubes. *Phys. Rev. B*, 61:9986, 2000.

[114] W.A. de Heer, J.-M. Bonard, K. Fauth, A. Châtelain, L. Forró, and D. Ugarte. Electron field emitters based on carbon nanotube films. *Adv. Mater.*, 9:87, 1997.

[115] A.H.R. Palser. *Theoretical properties of carbon nanotubes*, D.Phil. Thesis, Oxford University, 2000.

[116] V. Lordi, S.X.C. Ma, and N. Yao. Towards probing pentagons on carbon nanotube tips. *Surface Science*, 421:L150, 1999.

[117] A. De Vita, J.-Ch. Charlier, X. Blase, and R. Car. Electronic structure at carbon nanotube tips. *Appl. Phys. A*, 68:283, 1999.

[118] R. Tamura and M. Tsukada. Electronic states of the cap structure in the carbon nanotube. *Phys. Rev. B*, 52:6015, 1995.

[119] Ph. Lambin, A.A. Lucas, and J.C. Charlier. Electronic properties of carbon nanotubes containing defects. *J. Phys. Chem. Solids*, 58:1833, 1997.

[120] J.A. Harrison, S.J. Stuart, D.H. Robertson, and C.T. White. Properties of capped carbon nanotubes when used as SPM tips. *J. Phys. Chem. B*, 101:9682, 1997.

[121] D.L. Carroll, P. Redlich, P.M. Ajayan, J.C. Charlier, X. Blase, A. De Vita, and R. Car. Electronic structure and localized states at carbon nanotube tips. *Phys. Rev. Lett.*, 78:2811, 1997.

[122] P.W. Fowler, D.E. Manolopoulos, G. Orlandi, and F. Zerbetto. Energetics and isomerisation pathways of a lower fullerene: The Stone-Wales map for C_{40}. *J. Chem. Soc., Faraday Trans.*, 91:1421, 1995.

[123] H. Kanzow, C. Lenski, and A. Ding. Single-wall carbon nanotube diameter distributions calculated from experimental parameters. *Phys. Rev. B*, 63:125402, 2001.

[124] L.F. Sun, S.S. Xie, W. Liu, W.Y. Zhou, Z.Q. Liu, D.S. Tang, G. Wang, and L.X. Qian. Creating the narrowest carbon nanotubes. *Nature*, 403:384, 2000.

[125] L.-C. Qin, X. Zhao, K. Hirahara, Y. Miyamoto, Y. Ando, and S. Iijima. The smallest carbon nanotube. *Nature*, 408:50, 2000.

[126] N. Wang, Z.K. Tang, G.D. Li, and J.S. Chen. Single-walled 4 Å carbon nanotube arrays. *Nature*, 408:50, 2000.

[127] X. Blase, L.X. Benedict, E.L. Shirley, and S.G. Louie. Hybridization effects and metallicity in small radius carbon nanotubes. *Phys. Rev. Lett.*, 72:1878, 1994.

[128] R.A. Jishi, J. Bragin, and L. Lou. Electronic structure of short and long carbon nanotubes from first principles. *Phys. Rev. B*, 59:9862, 1999.

[129] L. Liu, C.S. Jayanthi, and S.Y. Wu. Manifestation of aromaticity and its effects on the electronic structure of finite single-wall carbon nanotubes. *Chem. Phys. Lett.*, 357:91, 2002.

[130] J. Liu, A.G. Rinzler, H. Dai, J.H. Hafner, R.K. Bradley, P.J. Boul, A. Lu, T. Iverson, K. Shelimov, C.B. Huffman, F. Roderiguez-Macias, Y.-S. Shon, T.R. Lee, D.T. Colbert, and R.E. Smalley. Fullerene pipes. *Science*, 280:1253, 1998.

[131] L.C. Venema, J.W.G. Wildöer, H.L.J. Temminck Tuinstra, C. Dekker, A.G. Rinzler, and R.E. Smalley. Length control of individual carbon nanotubes by nanostructuring with a scanning tunnelling microcope. *Appl. Phys. Lett.*, 71:2629, 1997.

[132] Z. Klusek, P. Kowalczyk, and P. Bysezewski. Investigations of electronic structure of capped carbon nanotubes by scanning tunneling spectroscopy. *Vacuum*, 63:145, 2001.

[133] T.W. Odom, J.-L. Huang, P. Kim, M. Ouyang, and C.M. Lieber. Scanning tunneling microscopy and spectroscopy studies of single wall carbon nanotubes. *J. Mater. Res.*, 13:2380, 1998.

[134] A. Thess, R. Lee, P. Nikolaev, H. Dai, P. Petit, J. Robert, C. Xu, Y.H. Lee, S.G. Kim, A.G. Rinzler, D.T. Colbert, G.E. Scuseria, D. Tománek,

J.E. Fischer, and R.E. Smalley. Crystalline ropes of metallic carbon nanotubes. *Science*, 273:483, 1996.