Electronic Structure and Quantum Conductance of Nanostructures

by

Young-Su Lee

M.S., Materials Science and Engineering
Seoul National University, 2001

B.S., Materials Science and Engineering
Seoul National University, 1999

Submitted to the Department of Materials Science and Engineering in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Materials Science and Engineering at the MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2006

© Massachusetts Institute of Technology 2006. All rights reserved.
Electronic Structure and Quantum Conductance
of Nanostructures
by
Young-Su Lee

Submitted to the Department of Materials Science and Engineering
on July 31, 2006, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Materials Science and Engineering

Abstract

This thesis is dedicated to development and application of a novel large-scale first-principles approach to study the electronic structure and quantum conductance of realistic nanoscale materials. Electron transport at the nanometer scale involves phenomena which are beyond the realm of classical transport theory: the wave character of the electrons becomes central, and the Schrödinger equation needs to be solved explicitly. First-principles calculations can nowadays deal with systems containing hundreds of electrons, but simulations for nanostructures that contain thousands of atoms or more need to rely on parametrized Hamiltonians.

The core of our approach lies in the derivation of exact and chemically-specific Hamiltonians from first-principles calculations, in a basis of maximally-localized Wannier functions, that become explicit tight-binding orbitals. Once this optimal basis is determined, the Hamiltonian matrix becomes short-ranged, diagonally-dominant, and transferable - i.e. a large nanostructure can be constructed by assembling together the Hamiltonians of its constitutive building block.

This approach is first demonstrated for pristine semiconducting and metallic nanotubes, demonstrating perfect agreement with full first-principles calculations in a complete planewave basis. Then, it is applied to study the electronic structure and quantum conductance of functionalized carbon nanotubes.

The first class of functionalizing addends, represented by single-bond covalent ligands (e.g. hydrogens or aryls), turns out to affect very strongly the back-scattering and the conductance, since \( sp^3 \) rehybridization at the sidewall carbon where a group is attached dramatically perturbs the conjugated \( \pi \)-bonding network. Inspection on the shape and the on-site energy of MLWFs before and after functionalizations leads to the conclusion that the effect of \( sp^2 \) rehybridization is essentially identical to removing a "half-filled" \( p_\pi \)-orbital from the \( \pi \)-manifold. In this perspective, the chemical difference between functional groups (e.g. different electronegativity of the residues) is relatively minor, even if, of course, will lead to different doping of the tube. We also find that these single-bond ligands tend to cluster, and are more stable when two groups are located nearby (incidentally, the degree of perturbation at the Fermi level
becomes weaker when such paired configuration is assumed).

The second class of functionalizing addends, represented by cycloaddition functionalizations (e.g. carbenes and nitrenes), demonstrates a radically different behavior. These addends are bonded to two neighboring sidewall carbon atoms, creating a three-membered ring structure. On narrow-diameter tubes, cleaving of the sidewall bond takes place to release the high strain energy of a three-membered ring. In the process, the two sidewall carbons recover their original \( sp^2 \) hybridization. This step is crucial, since the quantum conductance of a metallic nanotube then recovers almost perfectly the ideal limit of a pristine tube: the bond cleavage restores a transparent conduction manifold. Bond cleavage is controlled by the chemistry of the functional groups and the curvature of the nanotubes. High-curvature favors bond opening, whereas in graphene the bond is always closed; in between the two limits, chemistry determines the critical curvature at which the open-to-closed transition takes place. The preference for bond opening or closing has been screened extensively for different classes of functional groups, using initially some molecular homologues of the nanotubes. It is found that a subclass of addends, exemplified by dicyanocarbene, can assume both the open and closed form in the same tube around a narrow range of diameters. While these two forms are very similar in energy, and separated by a small barrier (hence they can be considered "fluxional" tautomers), the quantum conductance in the closed case is found to be significantly lower than that in the open case. Interconversion between the two minima could then be directed by optical or electrochemical means, in turn controlling the conductance of the functionalized tubes. We envision thus that this novel class of functionalization will offer a practical way toward non-destructive chemistry that can either preserve the metallic conductance of the tubes, or modulate it in real-time, with foreseeable applications in memories, sensors, imaging, and optoelectronic devices.

Thesis Supervisor: Nicola Marzari
Title: Associate Professor in Computational Materials Science
Acknowledgements

I would like to express my deepest gratitude and respect to my thesis advisor Prof. Nicola Marzari. He reminds me of the meaning of serendipity whenever I become too hesitant to take an unclear path. I am afraid that I have not fully grasped it; nevertheless, I find myself relishing few serendipitous moments when looking back on the past five years.

I would like to thank my committee members, Prof. Francesco Stellacci, Prof. John Joannopoulos, and Prof. Gerbrand Ceder, for their valuable suggestions, and also our collaborators, Prof. Marco Buongiorno Nardelli and Dr. Yudong Wu, for sharing their expertise at the initial stage of my thesis project.

I am indebted to the current and former members of our research group who have been great colleagues and friends. They are, in alphabetical order, Francesca Baletto, Nicola Bonini, Matteo Cococcioni, Ismaila Dabo, Boris Kozinsky, Heather Kulik, Nicholas Miller, Arash Mostofi, Nicolas Mounet, Damian Scherlis, Patrick Sit, Michael Tambe, Paolo Umari, Brandon Wood, and our kindest administrative assistant Kathryn Simons.

Life in a foreign country would have been so lonely without people who always make me feel at home: Yun Sung Kim, Inhee Chung, Jaeyeon Jung, Jeeyoung Choi, Taeyi Choi, Kisuk Kang, Sung Hoon Kang, and Wonjoon Jung.

I am very grateful to my dear friend Arum Yu with whom I have been sharing so many sweet and bitter memories. I express my warmest wishes for her future.

My daily life has been tireless dialogue with my faithful servant ahnbo, the keeper of data who never doses or sleeps (named after my best friend in high school). She has always been beside me since the first day we met, never losing her integrity, while
others have been suffering various failures and replacements torturing their masters.

Lastly, I owe too much to my family: my parents, sister (Eunsoo), brother-in-law (Donald), and two sweetest cousins (Hyegeun and Hyogeun). My small achievement would have been impossible without their love and care.

*I once read a silly fairy tale, called The Three Princes of Serendip: as their highness travelled, they were always making discoveries, by accidents and sagacity, of things which they were not in quest of: for instance, one of them discovered that a mule blind of the right eye had travelled the same road lately, because the grass was eaten only on the left side, where it was worse than on the right – now do you understand serendipity? ...*  

- from a letter by Horace Walpole
Contents

Introduction 19

1 First-principles calculations 21
  1.1 The Hohenberg-Kohn theorems 23
  1.2 The Kohn-Sham mapping 24
  1.3 Exchange-correlation functionals 25
  1.4 Practical implementations 26
    1.4.1 Bloch theorem 27
    1.4.2 Planewaves 28
    1.4.3 Pseudopotentials 29

2 Wannier functions 31
  2.1 Maximally-localized Wannier functions 32
    2.1.1 General k-point formalism 32
    2.1.2 Γ-point formalism 35
  2.2 Electronic structure in a Wannier representation 39
    2.2.1 Hamiltonian matrix in a periodic cell 39
    2.2.2 Band structure interpolation 42

3 Quantum transport 47
  3.1 Phase-coherent transport 48
    3.1.1 Ballistic transport 48
    3.1.2 The Landauer formula 49
3.1.3 Localization and fluctuations ........................................ 51
3.2 Quantum conductance from first-principles .......................... 52
  3.2.1 Hamiltonian matrix in a localized basis .......................... 54
  3.2.2 Transmission probability in the Green's function formalism ... 56
  3.2.3 Surface Green's functions ....................................... 58
  3.2.4 Summary ..................................................... 60

4 Pristine carbon nanotubes ............................................. 63
  4.1 Crystalline and electronic structure ................................. 64
  4.2 Maximally-localized Wannier functions and
      band structure interpolation ..................................... 69
  4.3 Real-space Hamiltonian matrix .................................... 75
      4.3.1 Periodic cell ........................................... 75
      4.3.2 Fragment ............................................. 76
  4.4 Quantum conductance ............................................. 78
  4.5 Experimental work on quantum transport ............................ 81

5 Disordered carbon nanotubes ......................................... 85
  5.1 Theoretical work on disordered carbon nanotubes ................ 86
  5.2 Experimental work on covalent functionalizations ............... 89

6 Functionalized carbon nanotubes:.................................... 95
    Hydrogens and aryl groups ........................................... 96
      6.1 Energetics and structure .................................... 96
      6.2 Band structure ............................................ 101
      6.3 Quantum conductance ....................................... 105

7 Functionalized carbon nanotubes:.................................... 111
    Carbenes and nitrenes ............................................. 112
      7.1 Energetics and structure .................................... 112
      7.2 Band structure and quantum conductance .................... 120
      7.3 Molecular homologues ....................................... 126
7.4 Conductance modulation via orbital rehybridization ............. 132

Conclusions ......................................................... 141

Bibliography ......................................................... 145
List of Figures

2-1 Disentanglement of a seven-dimensional subspace from the entangled bands of copper. ....... 34
2-2 Schematic description of a real-space integration method. ....... 40
2-3 Comparison between our real-space integration scheme and the Slater-Koster interpolation scheme. ............. 44
3-1 Schematic description of a single channel ballistic conductor. ....... 48
3-2 Schematic diagram of a reservoir-lead-conductor-lead-reservoir system. 49
3-3 Scattering of a planewave by a potential well. ............. 53
3-4 Illustration of the set-up for the calculations in a lead-conductor-lead system. ............. 55
3-5 Summary of the method. .................................. 61
4-1 Chiral and translational vectors of carbon nanotubes, and chiral and achiral carbon nanotubes. ............. 65
4-2 Brillouin zone of graphene and an illustration of the zone-folding method. 66
4-3 Structural parameters of a carbon nanotube in the tetragonal primitive unit cell used. ............. 67
4-4 Energy per carbon atom for a (5,5) and an (8,0) CNT as a function of the number of k-points in the z direction. ............. 68
4-5 Maximally-localized Wannier functions of the occupied states of graphene. 70
4-6 Maximally-localized Wannier functions of the occupied states of a (5,5) and an (8,0) CNT. ........................................ 70
4-7 Band structure of a (5,5) and an (8,0) CNT obtained from MLWFs interpolation scheme, using the MLWFs in the occupied subspace. . . 71
4-8 Disentangled maximally-localized Wannier functions of graphene. . . 73
4-9 Disentangled, maximally-localized Wannier functions of a (5,5) and an (8,0) CNT. .................................................. 74
4-10 Band structures of a (5,0) and an (8,0) CNT obtained from the interpolation scheme, using the MLWFs of the disentangled subspace. . . 74
4-11 Variation of Hamiltonian matrix element, $\langle \omega_0 | \hat{H} | \omega_n \rangle$, of a (5,5) CNT as a function of $|\langle r \rangle_0 - \langle r \rangle_n|$. ............................ 75
4-12 Band structure calculated from the Hamiltonian matrices to which different cutoff distances are applied. ................................. 76
4-13 Charge density and electrostatic potential of a fragment of a (5,5) CNT, $C_{110}H_{20}$. ..................................................... 77
4-14 $\langle \omega_0 | \hat{H} | \omega_n \rangle$ for the $C_{110}H_{20}$ fragment and band structure calculated from the Hamiltonian matrix of a central region in the fragment. . 78
4-15 Band structure, quantum conductance, and density of states of a (5,5) CNT. ......................................................... 80

5-1 Scaling of the mean-free-path as a function of the dopant concentration and the diameter of boron-doped (n,n) tubes. ...................... 88
5-2 Density of states and diffusivity of a boron-doped (10,10) CNT at three different energy levels. ............................................. 89
5-3 Absorption spectra of pristine and functionalized CNTs in dimethylformamide, illustrating the loss of characteristic structure following functionalization. .......................................... 91
5-4 Raman spectra of pristine and functionalized CNTs. .................... 91
5-5 Current as a function of gate voltage ($V_G$), for metallic CNTs after exposure to various concentrations of diazonium solutions. ........ 92
7-2 Three different configurations for a functional group CH$_2$ on a (5,5) CNT. .................................................. 113
7-3 $\Delta E$ for CH$_2$ and NH cycloaddition on an (n,n) CNT and formation energy of a pristine (n,n) CNT with respect to graphene. ............ 114
7-4 Sidewall equilibrium bond distance $d_{16}$ for (n,n) CNTs in the configurations S and O. ........................................... 115
7-5 Difference in $\Delta E$ between the S and O configuration for relaxed and unrelaxed (n,n) CNTs functionalized with CH$_2$. ........................ 116
7-6 Schematic description of a bent graphene sheet. ......................... 117
7-7 Relaxed structure and $d_{16}$ of a (6,6) CNT and a bent graphene sheet. 118
7-8 $\Delta E$ and $d_{16}$ for CH$_2$ functionalized (n,0) CNTs in the configurations S and P. .................................................. 119
7-9 $\Delta E$ and $d_{16}$ for (2n,n) chiral tubes functionalized with CH$_2$. .... 120
7-10 MLWFs of a (5,5) CNT functionalized with dichlorocarbene. .......... 121
7-11 Band structure of a DCC or an MCN functionalized (5,5) CNT in the configurations S and O. ........................................... 122
7-12 Quantum conductance for an infinite (5,5) CNT with one DCC or MCN group attached in the O open or closed configurations. .......... 123
7-13 Density of states projected onto $p_z$-MLWFs or $\sigma$-bonding MLWFs of C$_1$ and C$_6$ for three different CNTs: dichlorocarbene functionalized, hydrogen functionalized, and pristine (5,5) CNT. .................. 124
7-14 Quantum conductance for an infinite (5,5) CNT functionalized with one DCC group: comparison between a 5- and a 14-atomic-layer-long disordered central region. ........................................... 124
7-15 Quantum conductance for an infinite (5,5) CNT with thirty randomly distributed DCC groups in the open or closed (unstable) O configurations. ........................................... 125
7-16 Structure of two molecular homologues for functionalized carbon nanotubes. .................................................. 127
7-17 Potential energy surface as a function of $d_{16}$ for select cases of (a) 1 and (b) 2. ........................................ 129
7-18 Dominant interaction between the cyclopropane Walsh orbitals and the acetylene $p$-orbitals. ........................................ 130
7-19 Molecular structure of 1 of $X=C(NO_2)_2$. ........................................ 130
7-20 Spread of the MLWFs on the C$_1$-C$_{11}$ (or C$_6$-C$_{11}$) bond as a function of $d_{16}$. ........................................ 131
7-21 Maximally-localized Wannier Functions on the C$_1$-C$_{11}$ (or C$_6$-C$_{11}$) bond of 1 in Fig. 7-19. ........................................ 131
7-22 Potential energy surface as a function of $d_{16}$ for an $(n,n)$ CNT functionalized with C(CN)$_2$ in the $O$ configuration. ........................................ 133
7-23 Potential energy surface as a function of $d_{16}$ for an $(n,n)$ CNT functionalized with CH$_2$ in the $O$ configuration. ........................................ 133
7-24 Sidewall equilibrium bond distance $d_{16}$ for $(n,n)$ CNTs and for bent graphene sheets functionalized with C(CN)$_2$ in the $O$ configuration. ........................................ 134
7-25 Quantum conductance for an infinite (10,10) CNT with one or thirty randomly distributed C(CN)$_2$ groups in the $O$ open and closed configurations. ........................................ 135
7-26 Local density of states for the central 32-$nm$ segment of an infinite (10,10) CNT functionalized with thirty randomly distributed C(CN)$_2$ groups in the $O$ open and closed configurations. ........................................ 136
7-27 Walsh diagram showing the orbital energy evolution as a function of $d_{16}$ in the case of molecule 1 of $X=C(CN)_2$. ........................................ 137
7-28 Potential energy surface as a function of $d_{16}$ for the ground and for singlet excited states of molecule 1 (a) and 2 (b) of $X=C(CN)_2$. ........................................ 138
7-29 Isomerization pathway of 2o of $X=CH_2$ by photochemical excitations. ........................................ 140
7-30 Photochemical and thermal interconversion between several isomeric forms of a carbene functionalized fullerene. ........................................ 140
List of Tables

4.1 Structural parameters for \((n,n)\) armchair and \((n,0)\) zigzag nanotubes. 64

4.2 Parameters used for disentanglement of \(\pi\)- and \(\pi^*\)-bands in the case of
a \((5,5)\) CNT, an \((8,0)\) CNT, and graphene. ..................... 73

4.3 Spread of the maximally-localized Wannier functions of a \((5,5)\) CNT,
an \((8,0)\) CNT, and graphene. ................................. 73

6.1 Comparison of the reaction energies \(\Delta E_n\) of hydrogenation in a periodic
  cell or in a \(C_{110}H_{20}\) CNT fragment. ........................... 100

7.1 Experimental and theoretical \(d_{16}\) of 1. ............................. 128

7.2 Energy minimum conformations for 1 or 2. .......................... 128

7.3 Character table of point group \(C_{2v}\). ............................... 137
Introduction

It is not an overstatement to say that the pursuit of tininess has given a huge impetus to materials research in the last decade. This ever-growing field, represented by the beloved word “nano”, has created great excitement among diverse scientific communities, and unveiled phenomena that are unique to the nanometer scale. Out of the new physics emerging from the nano-world, understanding and controlling electron transport is of critical importance, considering its technological implications in engineering future electronic devices [1, 2].

Quantum effects prevail at the nanometer scale; classical approximations that treat electrons as particles are not valid any more as the wave character of electrons becomes more pronounced. At this length scale, it is more appropriate to consider electron transport as a transmission probability of electron waves [3], requiring that the Schrödinger equation that governs a quantum system be solved. While the exact solution of the Schrödinger equation is not available for many-electron systems, an alternative formulation (density-functional theory, established in 1960’s [4, 5]), has enjoyed a great practical success in recent decades in predicting the electronic structure of materials. This purely theoretical approach of determining physical properties by solving the quantum problem is often named first-principles or ab-initio. Current computer capabilities allow to simulate systems composed of hundreds of electrons on a common workstation. The increased capabilities on the theoretical side and the reduced sizes on the experimental side converge at the nanometer scale; indeed, the evolution of the field of nanoscience has been a textbook example of how theory and experiment can support and guide each other.

The main theme of this thesis is electron transport in nanoscale materials:
The first part of the work is focused on developing a theoretical tool for calculating quantum conductance in the phase-coherent transport regime. The method presented here combines well-established approaches with an original first-principles approach. Even if large scale first-principles simulations become more and more practical [6, 7], systems containing thousands of electrons or more are over the limit of the current capability. Our approach introduces a seamless bridge between accurate but relatively small first-principles calculations and large-scale model Hamiltonians, adopting maximally-localized Wannier functions as explicit tight-binding orbitals [8, 9].

The second part of the work is devoted to the application of this method to the study of electron transport in functionalized carbon nanotubes. Fifteen years after the discovery of nanotubes [10], their extraordinary electron transport properties still fascinate and engage scientists and engineers. Compelling evidences have been accumulated that metallic carbon nanotubes behave like an ideal one-dimensional quantum wire, exhibiting conductance close to the theoretical limit [11–13]. Such ideal behavior is usually plagued by various scattering sources, such as phonons, topological defects, impurities, etc.. On the other hand, electron transport could also be controlled by intentionally introducing scattering sources. Chemical functionalizations, which we deem to be a most promising way of achieving this goal, have proven to be extremely versatile in various applications, but their potential to control electron transport has just started to draw attention [14, 15]. To elucidate the scattering effects of different functional groups and to screen for optimal applications, electron transport in few paradigmatic, chemically-functionalized carbon nanotubes has been studied in detail; the present work provides a general picture of how chemistry will affect electron transport and suggests a novel approach toward conductance modulation in carbon-nanotube devices.
Chapter 1

First-principles calculations

Introduction

The state $\Psi$ of a quantum system is governed by the Schrödinger equation:

$$\hat{H}\Psi = E\Psi.$$  \hfill (1.1)

Under the Born-Oppenheimer approximation,$^1$ the Hamiltonian operator $\hat{H}$ of the time-independent Schrödinger equation for a system of $N$ electrons is given by \[16\]

$$\hat{H} = \hat{T}_e + \hat{V}_{ne} + \hat{V}_{ee} = \sum_{i}^{N} \left( -\frac{1}{2} \nabla_i^2 \right) + \sum_{i}^{N} v(r_i) + \sum_{i<j}^{N} \frac{1}{|r_i - r_j|}, \hfill (1.2)$$

$$v(r_i) = - \sum_{I} \frac{Z_I}{|r_i - R_I|}, \hfill (1.3)$$

where $E$ is the electronic energy, $r_i$ the coordinate of electron $i$, $R_I$ the coordinate of nucleus $I$, and $Z_I$ the charge of nucleus $I.$\(^2\)

The electronic energy $E$ and the $N$-electron many-body wavefunction $\Psi(r_1, r_2, \ldots, r_N)$ are obtained by solving this Schrödinger equation. The total energy $W$, including the

---

\(^1\)The Born-Oppenheimer approximation decouples the electronic and the nuclear wavefunctions: the positions of the nuclei enter the Schrödinger equation as a parameter and the electronic wavefunction is solved for any choice of these parameters.

\(^2\)The equation is written in atomic units: $\hbar = m_e = 4\pi\varepsilon_0 = 1.$
nucleus-nucleus interactions $V_{nn}$ is then given by

$$W = E + V_{nn} = E + \sum_{i<j} \frac{Z_i Z_j}{|\mathbf{r}_i - \mathbf{r}_j|}. \quad (1.4)$$

All the physical properties of a system (in the Born-Oppenheimer approximation) can be calculated in principle from the ground state wavefunction. This is the fundamental idea behind first-principles or ab-initio approaches: material properties are obtained from the solution of the Schrödinger equation without any experimental input. Though the idea itself is extremely appealing, the task of solving the many-body Schrödinger equation becomes rapidly intractable as the number of electrons increases.

A breakthrough was made by Hohenberg and Kohn [4] in 1964. They proved that electron density is a basic variable that uniquely determines the ground state $\Psi_0$ of an electronic system; that is the first formal proof of the density-functional theory. In principle, solving for the electron density greatly simplifies the problem. In practice though, we do not know the explicit form of the density-functional whose minimum corresponds to the solution of the Schrödinger equation. Kohn and Sham proposed a practical solution for this complex problem. They reformulated the problem for the interacting many-body system into that of an auxiliary non-interacting system, under the assumption that the two systems share the same ground state electron density [5]. This viewpoint proved itself very useful and the current implementations of the density-functional theory are based upon their formulation.

The first part of this chapter is dedicated to a brief introduction of these two seminal works that have made accurate first-principles calculations possible. A thorough discussion can be found in Ref. [16]. The second part concerns some practicalities in implementing the method, especially for extended solid-state systems.
1.1 The Hohenberg-Kohn theorems

Hohenberg and Kohn first provided a rigorous justification that the ground-state charge density \( n(\mathbf{r}) \) can be used as the basic variable for the many-body problem [4]. Their first theorem proves that, for a system of interacting particles under an external potential \( v(\mathbf{r}) \), \( v(\mathbf{r}) \) is a unique functional of the ground-state electron density \( n(\mathbf{r}) \). The external potential \( v(\mathbf{r}) \) then fixes the Hamiltonian and thus the ground state wavefunction \( \Psi_0 \); the ground state of a many-body system is therefore a unique functional of \( n(\mathbf{r}) \). Their second theorem introduces a formal variational principle on the charge density itself. Since \( \Psi_0 \) is a functional of \( n(\mathbf{r}) \), one can define

\[
F[n] \equiv \langle \Psi_0 | \hat{T}_e + \hat{V}_{ee} | \Psi_0 \rangle .
\]

(1.5)

\( F[n] \) is clearly a universal functional independent of \( v(\mathbf{r}) \) since the kinetic energy \( T_e \) and the electron-electron interaction energy \( V_{ee} \) are functionals only of \( n(\mathbf{r}) \). Then the ground state energy functional is defined as

\[
E[n] \equiv F[n] + \int v(\mathbf{r}) n(\mathbf{r}) \, d\mathbf{r} .
\]

(1.6)

For a given \( v(\mathbf{r}) \), ground state density \( n(\mathbf{r}) \) gives the ground state energy, and for any trial electron density \( \tilde{n}(\mathbf{r}) \),

\[
E_0 = E[n] \leq E[\tilde{n}] .
\]

(1.7)

\( E[n] \) assumes its minimum when the electron density is the exact ground state density \( n(\mathbf{r}) \) (see Ref. [16] for a clear introduction to this field).

The Hohenberg-Kohn approach represents a great simplification over the many-body Schrödinger equation. If the explicit form of the universal functional \( F[n] \) were known, the problem of determining the ground state would reduce to the problem of the minimization of a functional of the 3-dimensional electron density, i.e., a function of 3 coordinates, instead of \( 3N \).
1.2 The Kohn-Sham mapping

The celebrated density-functional $F[n]$, although well-defined, is not known in practice. In order to make progress, Kohn and Sham introduced an auxiliary non-interacting electron system which replaces the many-body electron system. Their approach is based on the assumption that the ground state charge density of a many-body electron system can be represented by that of an auxiliary non-interacting system.

In this non-interacting electron model, the exact wavefunction $\Psi$ is a Slater determinant

$$\Psi = \frac{1}{\sqrt{N!}} \det[\psi_1, ..., \psi_N], \quad (1.8)$$

and the charge density is given by

$$n(r) = \sum_{i}^{N} |\psi_i(r)|^2, \quad (1.9)$$

where $\{\psi_i\}$ are the $N$ lowest eigenstates of the Kohn-Sham equation:

$$\hat{H}_{KS}\psi_i = \left[ -\frac{1}{2} \nabla^2 + v_{KS} \right] \psi_i = \epsilon_i \psi_i, \quad (1.10)$$

where $v_{KS}$ is the effective one-electron Kohn-Sham potential defined as the potential for which the ground-state charge density for the non-interacting electrons is identical to that for the interacting electrons. In the Kohn-Sham approach, the ground-state energy functional is decomposed into

$$E[\{\psi_i\}] = T[\{\psi_i\}] + E_H[n] + E_{xc}[n] + \int v(r)n(r)dr, \quad (1.11)$$

where $T$ is the non-interacting kinetic energy given by

$$T[\{\psi_i\}] = -\frac{1}{2} \sum_{i} \langle \psi_i | \nabla^2 | \psi_i \rangle, \quad (1.12)$$
and $E_H[n]$ is the Hartree energy that represents the classical Coulomb interaction

$$E_H[n] = \frac{1}{2} \int \frac{n(r) n(r')}{|r - r'|} \, dr \, dr', \quad (1.13)$$

and $E_{xc}[n]$, that captures all remaining contributions, is the exchange-correlation energy that includes all complex many-body effects of exchange and correlation. Comparison between Eq. 1.6 and Eq. 1.11 shows this:

$$E_{xc}[n] = F[n] - (T[\{\psi_i\}] + E_H[n]) \quad (1.14)$$

$$= \langle \Psi | \hat{T}^e | \Psi \rangle - T[\{\psi_i\}] + \langle \Psi | \hat{V}_{ec} | \Psi \rangle - E_H[n].$$

$E_{xc}[n]$ is composed of the difference between the kinetic energy of the many-body and the non-interacting system, and non-classical part of the many-body electron-electron interaction term.

The effective one-electron Kohn-Sham potential can then be written as

$$v_{KS} \equiv v(r) + v_H(r) + v_{xc}(r) \equiv v(r) + \int \frac{n(r')}{|r - r'|} \, dr' + \frac{\delta E_{xc}[n(r)]}{\delta n(r)}. \quad (1.15)$$

The Kohn-Sham potential itself is dependent on the charge density; Eq.1.10 therefore needs to be solved in a self-consistent manner.

Though the Kohn-Sham approach involves solving $N$ one-electron wavefunction instead of the charge density, the explicit treatment of wavefunctions enables to approximate the exact kinetic energy term much more accurately. This, combined with reasonable approximations for the exchange-correlation functional has led to remarkable predictive accuracy and has made first-principles calculations a very successful and practical approach.

### 1.3 Exchange-correlation functionals

The exact form of the exchange-correlation functional is still unknown. However, having removed the kinetic energy and classical Hartree energy from $F[n]$ allows to
approximate $E_{xc}[n]$ with reasonable accuracy. The simple suggestion by Kohn and Sham is the local density approximation (LDA, then extended also to spin-polarized cases, LSDA). LDA assumes that the exchange-correlation functional is purely local: the exchange-correlational energy density at a position $r$ (and density $n(r)$) is defined as that of homogeneous electron gas with the same density $n(r)$,

$$
\varepsilon_{xc}(n(r)) = \varepsilon_{xc}^{hom}(n(r)).
$$

(1.16)

The total exchange-correlation energy is thus

$$
E_{xc}^{LDA}[n(r)] = \int n(r)\varepsilon_{xc}^{hom}(n(r))dr ,
$$

(1.17)

and

$$
v_{xc}^{LDA}(r) = \frac{\delta E_{xc}^{LDA}[n(r)]}{\delta n(r)} .
$$

(1.18)

The next level of approximation is the generalized gradient approximation (GGA) which takes the inhomogeneities into account, as well as important sum rules [17]:

$$
E_{xc}^{GGA}[n(r)] = \int f(n(r), \nabla n(r))dr .
$$

(1.19)

GGA gives more accurate results than LDA in many cases, but it is nonetheless not a systematic improvement. That can be improved by further powers of the gradient expansions.

1.4 Practical implementations

First-principles calculations in the present work are carried out in the planewave pseudopotential framework. We introduce here some of the important elements of the method; an exhaustive presentation can be found in Ref. [18].
1.4.1 Bloch theorem

Bloch theorem provides a powerful approach to determining the eigenstates in a crystalline system [19]. When the Hamiltonian operator $\hat{H}$ displays translational symmetry,

$$\hat{H}\psi_i(r) = \left[-\frac{1}{2}\nabla^2 + v(r)\right] \psi_i(r) = \varepsilon_i\psi_i(r) \quad (1.20)$$

where

$$v(r) = v(r + R) \quad \text{(for all } R \text{ in a Bravais lattice)}, \quad (1.21)$$

the eigenfunction $\psi_i$ can be chosen in the following form:

$$\psi_{nk}(r + R) = e^{i\mathbf{k}\cdot R}\psi_{nk}(R) \quad (1.22)$$

or equivalently,

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

$$u_{nk}(r + R) = u_{nk}(r) \quad (1.23)$$

A wave vector $\mathbf{k}$ is associated with each $\psi$, and the Schrödinger equation becomes separable:

$$\left[-\frac{1}{2} \left(\nabla + i\mathbf{k}\right)^2 + v(r)\right] u_{nk}(r) = \varepsilon_{nk} u_{nk}(r) \quad , \quad (1.24)$$

where $\{\psi_{nk}\}$ are orthonormal wavefunctions, i.e. $\langle \psi_{nk}|\psi_{n'k'}\rangle = \delta_{nn'}\delta_{kk'}$.

Eq. 1.24 can be solved within the unit cell (the unit of translational periodicity) of the crystal; however, this still needs to be solved for all $\mathbf{k}$ inside the Brillouin zone (BZ) in order to calculate e.g. the charge density

$$n(r) = \sum_{\mathbf{k}} \sum_n |u_{nk}(r)|^2 \quad (1.25)$$

or the total energy ($N$ is the number of the occupied bands at $\mathbf{k}$). While, in principle, an infinite number of $\mathbf{k}$ vectors must be sampled to obtain the exact result (equivalent to simulating a crystal of infinite dimensions), in practice, the number of $\mathbf{k}$ vectors can be systematically increased until the physical quantities of interest converge within
desired accuracy. A regular mesh of k-points $N_{kp,1} \times N_{kp,2} \times N_{kp,3}$ inside the BZ is commonly used, where $N_{kp,i}$ is the number of k-points along the primitive reciprocal lattice vector $b_i$.

### 1.4.2 Planewaves

In practice, Eq. 1.24 can be solved on a finite grid, or equivalently the electron wavefunctions can be expanded in a finite basis. The choice of basis functions determines efficiency and accuracy of the simulations.

A planewave basis set can be consistently used within periodic boundary conditions and thus is best suited for crystalline systems. An electron wavefunction in a planewave basis is expressed as

$$u_{nk}(r) = \sum_{G}^{N_{pw}} c_{nk}(G) e^{iG \cdot r}, \quad (1.26)$$

or

$$\psi_{nk}(r) = \sum_{G}^{N_{pw}} c_{nk}(G) e^{i(k + G) \cdot r}, \quad (1.27)$$

where $G$ is a reciprocal lattice vector. The sum is customarily taken over the set of $G$ vectors satisfying $|k + G|^2 < E_{cut}$, where $E_{cut}$ is the planewave cutoff energy. Then the total number of planewaves, $N_{pw}$, is proportional to $(E_{cut})^{3/2}$. Some of the advantages of the planewave basis are the following:

- The accuracy of simulations can be improved systematically by increasing $E_{cut}$. The cutoff energy $E_{cut}$ is the single variable that controls the error coming from the incompleteness of basis functions, irrespective of the atomic species involved.

- It does not depend on the position of atoms, so different atomic configurations can be compared on an equal footing.

- It allows to take advantage of the fast Fourier transformation between the real and the reciprocal space, greatly reducing the number of operations required.
A disadvantage of the planewave basis is that the number of basis functions is usually large compared to a localized basis, especially for the case of isolated systems.

1.4.3 Pseudopotentials

The idea of a pseudopotential is that of replacing the potential of the nucleus and the core electrons with an effective potential acting only on the valence electrons. Core electrons are tightly bound to the nucleus and remain unperturbed under different chemical environments; their role is limited to screening the potential of the nucleus. Incorporating core electrons into the pseudopotential is doubly advantageous from the computational point of view: the total number of electrons is reduced to the number of valence electrons, and a much smaller number of basis functions is required when only the smoother valence electron wavefunctions are involved.

A typical norm-conserving pseudopotential, where the valence pseudo-wavefunctions satisfy the orthonormality condition \( \langle \psi_i^{PS} | \psi_j^{PS} \rangle = \delta_{ij} \), is constructed following the prescriptions below [20]:

- All-electron and pseudo eigenvalues agree for a chosen reference atomic configuration.

- All-electron and pseudo atomic wavefunctions agree beyond a chosen core radius \( r_c \).

- The integrals from 0 to \( r \) of the all-electron and the pseudo charge densities agree for \( r > r_c \) for each valence state (norm conservation).

- The logarithmic derivatives of the all-electron and the pseudo wavefunctions and their first energy derivatives agree for \( r > r_c \).

Ultrasoft pseudopotential

Norm-conserving pseudopotentials require a high \( E_{cut} \) for first row and 3d transition metal elements since screening by core electrons for 2p or 3d orbitals is very
weak (there are no 1p or 2d orbitals to be orthogonal to). Since the computational cost grows as the planewave cutoff increases, a smoother pseudo wavefunction would be desirable. The ultrasoft pseudopotential formulation of Vanderbilt generates optimally smooth pseudo wavefunctions, by relaxing the norm-conserving condition [21]. The pseudo wavefunctions satisfy a generalized orthonormality condition $\langle \psi_i^{PS} | \hat{S} | \psi_j^{PS} \rangle = \delta_{ij}$, and the deficit of charge around the core is compensated by "augmentation" charge.
Chapter 2

Wannier functions

Introduction

Wannier functions were originally introduced by Wannier in 1937 [8]. They are orthonormal atomic-like wavefunctions that represent the same Hilbert space of the Bloch functions. Fourier transform of the extended Bloch functions is used to generate Wannier functions that are localized in real-space:

$$|w_{n\mathbf{R}}\rangle = \frac{V}{(2\pi)^3} \int_{BZ} |\psi_{nk}\rangle e^{-i\mathbf{k} \cdot \mathbf{R}} d\mathbf{k}, \quad (2.1)$$

where $V$ is the volume of real-space primitive cell and $\mathbf{R}$ is a Bravais lattice vector in real-space.

Wannier functions have the following properties:

- Span the same Hilbert space of Bloch functions.

- Constitute a complete orthonormal basis set in real-space. Indices $n$, $\mathbf{R}$ are used instead of $n$, $\mathbf{k}$, such that

$$\langle w_{n\mathbf{R}}|w_{n'\mathbf{R}'}\rangle = \delta_{nn'}\delta_{\mathbf{R}\mathbf{R}'}.$$  

- Have the same translational symmetry of the Bravais lattice:

$$w_{n\mathbf{R}}(\mathbf{r} - \mathbf{R}) = w_{n\mathbf{R}'}(\mathbf{r} - \mathbf{R}').$$
Though the concept of Wannier functions is straightforward, Wannier functions are not uniquely defined. The non-uniqueness comes from the phase indeterminacy in Bloch functions: any arbitrary phase factor added to an eigenfunction,

$$|\psi_{nk}\rangle \rightarrow e^{i\phi_{nk}}|\psi_{nk}\rangle$$

leaves all physical properties unchanged, but different choices of $e^{i\phi_{nk}}$ will certainly produce different Wannier functions. This inherent non-uniqueness together with the lack of a robust algorithm to construct Wannier functions has limited their utility, while several criteria have been used (for example, one can choose Wannier functions that have maximal overlap with guiding orbitals located at atom centers or bond centers [22]). We adopt here the approach recently proposed by Marzari and Vanderbilt [9]. The formalism and its extension by Souza et al. [23] are explained in Section 2.1. In Section 2.2, description of the electronic structure in a Wannier representation is discussed.

2.1 Maximally-localized Wannier functions

2.1.1 General k-point formalism

In the most general form, eigenstates belong to an isolated group of $N$ bands can be mixed together when constructing Wannier functions:

$$|w_{nR}\rangle = \frac{V}{(2\pi)^3} \int_{BZ} \sum_{m}^{N} U_{mn}^{(k)}|\psi_{mk}\rangle e^{-i\mathbf{R}\cdot\mathbf{k}} d\mathbf{k} .$$ (2.2)

The criterion introduced by Marzari and Vanderbilt to determine $\{U_{mn}^{(k)}\}$ is to minimize the sum of the mean square spread of the Wannier functions, defined as [9]

$$\Omega = \sum_{n}^{N} \left[ \langle r^2 \rangle_n - \langle r \rangle_n^2 \right] = \sum_{n}^{N} \left[ \langle w_{n0}|r^2|w_{n0}\rangle - \langle w_{n0}|r|w_{n0}\rangle^2 \right] .$$ (2.3)

The Wannier functions that minimize the spread functional $\Omega$ are named maximally-localized Wannier functions (MLWFs). The method has been widely used in many
applications for the following reasons:

- The minimization process is robust, and MLWFs provide a clear picture of chemical bonding [24].

- The centers of MLWFs, and their displacement under polarizing field have a close, formal connection to the macroscopic and microscopic polarization of an insulating system [9].

- The condition of “maximal localization” can be exploited by diverse methods that rely on real-space localized basis sets, such as $O(N)$ methods.

The expectation values $\langle r \rangle_n$ and $\langle r^2 \rangle_n$ are needed to calculate $\Omega$ in terms of the Bloch functions. They are given by [25]

$$\langle r \rangle_n = \frac{i}{2\pi^3} \int \frac{d\mathbf{k}}{(2\pi)^3} \langle u_{nk} | \nabla_k | u_{nk} \rangle ,$$

$$\langle r^2 \rangle_n = \frac{V}{(2\pi)^2} \int \frac{d\mathbf{k}}{(2\pi)^2} | | \nabla_k u_{nk} |)^2 .$$

A finite-difference expression for the above formulas on a regular mesh of k-points can be obtained as a function of the overlap matrices $M^{(k,b)}$, defined as

$$M^{(k,b)}_{mn} = \langle u_{mk} | u_{n,k+b} \rangle :$$

$$\langle r \rangle_n = -\frac{1}{N_{kp}} \sum_k N_k \sum_b W_b b \Im \ln M^{(k,b)}_{nn} ,$$

$$\langle r^2 \rangle_n = \frac{1}{N_{kp}} \sum_k N_k \sum_b W_b \left\{ \left[ 1 - |M^{(k,b)}_{nn}|^2 \right] + \left[ \Im \ln M^{(k,b)}_{nn} \right]^2 \right\} .$$

$b$ is a set of vectors pointing to the neighbors of a point $k$, $W_b$ is a weighting factor for the corresponding vector $b$, $N_b$ is the number of vectors $b$, and $N_{kp}$ is the total number of k-points. The spread functional $\Omega$ and its derivative with respect to $U$, $d\Omega/dU_{mn}^{(k)}$, can all be expressed in terms of $M^{(k,b)}_{mn}$. 

33
Disentanglement

A difficulty arises in the case of metallic systems since all the bands are connected together; there is no finite gap that allows to isolate a group of bands. In this case, a simple unitary transformation between occupied bands fails to localize the wavefunctions in real-space [26]. The band structure of copper in Fig.2-1 illustrates this challenge in metallic systems. Two sets of bands originate from atomic orbitals of different character: five narrow d-bands and one dispersive s-band are mixed together. The lowest five bands consist of either five d-bands or four d-bands plus one s-band depending on the position in the BZ. MLWFs from these bands are not localized due to the non-analytic disconnectivities in the Fourier transforms as two bands (one occupied and one unoccupied) cross. Including the six lowest bands would not solve the problem either, since the s-band is now mixed with higher-energy bands of different character.

Souza et al. [23] proposed a method that can deal with this complex “entangled” manifold. Their strategy is to select a \textit{maximally-connected} subspace out of the entangled space. That is achieved by minimizing the “gauge-variant” part of $\Omega$. $\Omega$ is decomposed into two terms [9]:

$$\Omega = \Omega_t + \tilde{\Omega}, \quad (2.9)$$
\[ \Omega_I = \frac{1}{N_{kp}} \sum_{k,b} W_b \left( N - \sum_{m,n} |M_{mn}^{(k,b)}|^2 \right) \]
\[ = \frac{1}{N_{kp}} \sum_{k,b} W_b \text{Tr}[\hat{P}_k \hat{Q}_{k+b}] , \]

where \( \hat{P}_k = \sum_n |u_{nk} \rangle \langle u_{nk}| \) and \( \hat{Q}_k = 1 - \hat{P}_k \). The gauge-invariant part \( \Omega_I \) is a measure of the change of character across the Brillouin zone; this is the quantity to be minimized in order to select a subspace composed of orbitals of similar character. For an isolated group of bands, the number of MLWFs \( (N) \) are the same as the number of bands \( (N_k) \); \( \Omega_I \) remains unchanged upon unitary transformation and minimizing \( \Omega \) is equivalent to minimizing \( \Omega_I \). When \( N < N_k \), a target \( N \)-dimensional subspace that minimizes \( \Omega_I \) is selected out of the \( N_k \)-dimensional original space. Fig. 2-1 shows the case where a seven-dimensional subspace is selected out of the entangled bands inside the outer window. An inner window sets the condition that all the bands inside it should be fully included in the disentangled subspace, which ensures that electronic structure inside the inner window is accurately reproduced in the Wannier representation. The number of MLWFs, and the inner and the outer window are the three variables that determine the subspace that is obtained.

### 2.1.2 Γ-point formalism

The size of the BZ is inversely proportional to the size of the unit cell in real-space. In the case of a large supercell containing hundreds of atoms, a single Γ-point sampling in the BZ is sufficient to converge charge density and energies. Maximally-localized Wannier functions at Γ, \( |w_n\rangle \), are then obtained from a unitary transformation of the eigenfunctions (for simplicity, \( \psi_n \) represents here \( \psi_{n\Gamma} \)):

\[ |w_n\rangle = \sum_m^N U_{mn} |\psi_m\rangle . \]  
(2.11)

The minimization procedure becomes simpler since it concerns a single unitary matrix \( U \) instead of \( N_{kp} \) unitary matrices. In addition, wave functions at Γ are real, due to time reversal symmetry \( (\psi_{n-k}(r) = \psi_{n,k}^*(r)) \), which further restricts \( U \) to be a real unitary matrix (i.e. orthogonal) and simplifies the algebra. In this case, the \( b \) vectors
connecting a point \( k \) to its neighbors are related to the primitive lattice vectors \( \mathbf{G}_l \) of reciprocal space, and the spread functional \( \Omega \) is defined as

\[
\Omega = \sum_{n} \sum_{l} W_l \left[ \langle w_n | (\mathbf{G}_l \cdot \mathbf{r})^2 | w_n \rangle - \langle w_n | \mathbf{G}_l \cdot \mathbf{r} | w_n \rangle^2 \right]
= \sum_{n} \sum_{l} W_l \left( 1 - |\langle w_n | e^{-i\mathbf{G}_l \cdot \mathbf{r}} | w_n \rangle|^2 \right) + \mathcal{O}(L^{-2})
= \Omega' + \mathcal{O}(L^{-2}) ,
\]

where \( L \) is the supercell dimension, and \( W_l \) and \( N_G \) are similarly defined as \( W_b \) and \( N_b \) in Eq. 2.7. In order for Eq. 2.12 to be valid, the weight factor \( W_l \) and the \( \mathbf{G}_l \) vectors must satisfy the following condition:

\[
\sum_{l} W_l (\mathbf{G}_l)_{\alpha}(\mathbf{G}_l)_{\beta} = \delta_{\alpha\beta} ,
\]

where \( \alpha \) and \( \beta \) denote \( x, y, z \) directions in Cartesian coordinates. \( \mathbf{G}_l \) and \( W_l \) for fourteen Bravais lattice are summarized in Ref. [27] and [28] (in the case of the simple cubic Bravais lattice, \( \mathbf{G}_l \) vectors are the three primitive reciprocal lattice vectors). We define the overlap matrices as in Eq. 2.6:

\[
M_{mn}^l = \langle u_m | u_n | \mathbf{G}_l \rangle = \langle u_m \rho | e^{-i\mathbf{G}_l \cdot \mathbf{r}} | \psi_n \rangle = \langle \psi_m \rho | e^{-i\mathbf{G}_l \cdot \mathbf{r}} | \psi_n \rangle .
\]

Minimizing \( \Omega \) is equivalent to minimizing \( \Omega' \) [24], within an \( \mathcal{O}(L^{-2}) \) error. \( \Omega' \) can be minimized with steepest-descents and conjugate-gradients [28] or Jacobi rotations [29]. When convergence is achieved, the center \( \langle \mathbf{r} \rangle_n \) and the spread \( \Omega'_n \) of an MLWF are obtained from

\[
\langle \mathbf{r} \rangle_n = - \sum_{l} W_l \mathbf{G}_l \text{Im} \ln M_{nn}^l ,
\]

\[
\Omega'_n = \sum_{l} W_l (1 - |M_{nn}^l|^2) .
\]
Disentanglement

A disentanglement procedure can also be adapted to the \(\Gamma\)-point case. A \(\Gamma\)-point version of the decomposition of \(\Omega'\) into \(\Omega_1\) and \(\tilde{\Omega}\) is as following:

\[
\Omega' = \sum_n \sum_l W_i (1 - |M_{nn}^l|^2) = \sum_l W_i \{ N - \sum_{mn} |M_{mn}^l|^2 \} + \sum_l W_i \sum_{mn, m \neq n} |M_{mn}^l|^2
\]

(2.17)

\[
\Omega_1 = \sum_l W_i \{ N - \sum_{mn} |M_{mn}^l|^2 \} = \sum_l W_i \text{Tr}[\hat{P}_l \hat{Q}_{G_l}].
\]

(2.18)

As in the case of multiple k-point sampling, minimizing \(\Omega_1\) provides an \(N\)-dimensional \textit{optimally-smooth} subspace out of the \(N_o\)-dimensional space chosen inside the outer window.

Practical implementation closely follows the formulas by Souza et al. [23]. The initial guess for the \(N\)-dimensional subspace \(\hat{P}^{(0)}\) consists of a spectrum of eigenstates that includes \(N_i\) states inside the inner window (an alternative would be to start from the projection of \(N\) “guiding” orbitals onto the original space). The condition of minimizing \(\Omega_1\) reduces to an iterative eigenvalue problem. At each \(i_{th}\) iteration step, the disentangled subspace \(\hat{P}^{(i)}\) is updated. An \(N_o \times N\) rectangular matrix \(A\) (such as \(A^\dagger A = I_N\)) is the quantity to be determined:

\[
\hat{P}^{(i)} = \sum_{n=1}^{N} |\phi_n\rangle \langle \phi_n| \quad \text{where} \quad |\phi_n\rangle = \sum_{m=1}^{N_o} A^{(i)}_{mn} |u_m\rangle.
\]

(2.19)

An \(N_o \times N_o\) complex Hermitian matrix \(Z^{(i)}\) which represents the projection of a nearby subspace at \(G_i\) onto the \(N_o\)-dimensional space at \(\Gamma\) is constructed:

\[
Z^{(i)}_{mn} = \langle u_m | \sum_l W_i \hat{P}^{(i)}_{G_l} | u_n \rangle
\]

\[
= \sum_j \sum_{l} W_i \left( \sum_p A^{(i)}_{pj} M_{mp}^l \right) \left( \sum_q A^{(i)}_{qj} M_{mq}^l \right)^*.
\]

(2.20)
The matrix $A$ at $(i+1)_{th}$ step is composed of $N$ eigenvectors of $Z(i)$ having the $N$ largest eigenvalues: each eigenvector constitutes a column of the $A(i)$ matrix. The procedure is iterated until $\hat{P}^{(i+1)} = \hat{P}^{(i)}$ within a convergence threshold. In practice, the algorithm is optimized such that the $N_i$-dimensional space inside the inner window is always included in $\hat{P}$, which means that an $(N-N_i)$-dimensional subspace is extracted from the $(N_o-N_i)$-dimensional subspace and the size of $Z$ is $(N_o-N_i) \times (N_o-N_i)$.

In general, the rectangular matrix $A$ is complex since $Z$ is a complex Hermitian matrix. However, we would like to have a real matrix $A$ that would generate real Wannier functions. A complex part could derive from a trivial phase factor. In that case, each component of $A$ could be written as $A_{mj} = e^{i\phi_j} A_{mj}^R$, with $A_{mj}^R$ real. Projection to the $N_o$-dimensional space should remove this phase factor. The projection matrix $P$ is

$$P_{mn} = \sum_j A_{mj}(A^\dagger)_{jn} = \sum_j A_{mj} A_{nj}^* = \sum_j e^{i\phi_j} A_{mj}^R e^{-i\phi_j} A_{nj}^R = \sum_j A_{mj}^R A_{nj}^R.$$  

If these assumptions were true, $P$ should have all real components, but we found small complex components. These complexity could be intrinsic, or derive from the finite-difference representation of the operators. We have not investigated this matter further.

In order to exploit the simplicity of a formalism based upon real wavefunctions, rather than complex, we add the constraint that the disentangled subspace should be real. Then the task is to find a real subspace $\hat{R}$ that maximizes $\text{Tr}[\hat{P}\hat{R}]$. The complex Hermitian matrix $P$ is decomposed into the real symmetric matrix $P_S$ and the real anti-symmetric matrix $P_A$:

$$P = P_S + iP_A.$$  

It can be easily shown that $\text{Tr}(P_A R) = 0$ when $R$ is a real symmetric matrix, and then $\text{Tr}[\hat{P}\hat{R}]$ becomes

$$\text{Tr}(PR) = \text{Tr}((P_S + iP_A)R) = \text{Tr}(P_SR) + i\text{Tr}(P_AR) = \text{Tr}(P_SR).$$  

(2.23)
The problem thus reduces to finding the real symmetric matrix $R$ that has the maximum overlap with $P_S$ under the restriction that $\hat{R} = \sum_{n=1}^{N} |r_n\rangle\langle r_n|$ and $\langle r_m | r_n \rangle = \delta_{mn}$. The mathematical formalism is the same as the minimization of $\Omega_i$: $\hat{R}$ is obtained from the eigenvectors $|r_n\rangle$ corresponding to the $N$ largest eigenvalues of the $N_o \times N_o$ matrix $P_S$. The eigenvalues of $P_S$ are a measure of the size of the complex part of $P$: when $P$ itself is real, all eigenvalues will be one. In practice, we found that if the constraint is applied from the beginning, the iterative minimization procedure becomes unstable, which implies $\|P_A\|$ is large at first few iteration steps. We perform the minimization until a converged complex matrix $A$ is obtained. From that point onwards, the additional step for $\hat{R}$ is inserted, and $\hat{R}^{(i)}$ replaces $\hat{P}^{(i)}$ when constructing $Z^{(i)}$. The procedure is iterated until $\hat{R}^{(i+1)} = \hat{R}^{(i)}$.

### 2.2 Electronic structure in a Wannier representation

#### 2.2.1 Hamiltonian matrix in a periodic cell

Our transformation into real-space localized Wannier functions allows for the simplicity of tight-binding models with the accuracy of first-principles calculations. This is because the Hamiltonian matrix in the Wannier-functions basis will be diagonally dominant, since MLWFs are localized in real-space. However, one should keep in mind that Wannier functions as constructed are also periodic, whose periodicity is determined by the number of k-points sampled, and that should be considered very carefully when constructing the real-space Hamiltonian matrix. The number of k-points $N_{kp,i}$ sampled along the $b_i$ primitive reciprocal lattice vector is linked to the Born-von Karman real-space periodicity along the corresponding lattice vector $a_i$:

$$w_n(r) = w_n(r + N_{kp,i} \times a_i). \quad (2.24)$$
We elaborate this issue for a one-dimensional system extended in the $z$ direction. In the case of Γ-point sampling, MLWFs are $|w_n\rangle = \sum_m U_{mn}|\psi_m\rangle$ and the Hamiltonian matrix in $|w_n\rangle$ basis is simply

$$H^{(\text{rot})} = U^\dagger H U.$$  

(2.25)

However, $H^{(\text{rot})}$ is not truly the real-space Hamiltonian we want. According to Eq. 2.24, MLWFs have a periodicity of $L$ when a single Γ-point is sampled:

$$w_n(z) = w_n(z + L).$$  

(2.26)

The situation is illustrated in Fig. 2-2. The MLWFs $w_1$ and $w_2$ are localized, but periodical repeat in one-dimensional supercell of length $L$. If we focus on $0th$ supercell, the matrix element $\langle w_1 | \hat{H} | w_2 \rangle$ must be close to zero, but in reality a periodic image of $w_1$ or $w_2$ (colored in white-gray) would make $\langle w_1 | \hat{H} | w_2 \rangle$ different from zero. We assign an index to each periodic image: $|\omega_n^i\rangle$ represents the well-defined (thanks to localization) portion of $n_{th}$ MLWF centered at $i_{th}$ supercell, in the effective range of $[z_n^i - L/2, z_n^i + L/2)$, where $z_n^i$ is the center of $|\omega_n^i\rangle$ defined as $z_n^i = z_n^0 + iL \ (0 \leq z_n^0 < L)$. Then $|w_n\rangle$ is

$$|w_n\rangle = \sum_{i=-\infty}^{\infty} |\omega_n^i\rangle.$$  

(2.27)
\( H_{mn}^{(rot)} = (U^\dagger H U)_{mn} \) can be decomposed into the sum of the contributions from each periodic image:

\[
\langle w_m | \hat{H} | w_n \rangle = \langle \omega_m^0 | \hat{H} | \omega_n^0 \rangle = \sum_{i=-\infty}^{\infty} \langle \omega_m^i | \hat{H} | \omega_n^i \rangle = H_{mn}^{00} + H_{mn}^{01} + H_{mn}^{01} . \tag{2.28}
\]

The element \( H_{mn}^{ij} \) is strictly zero whenever \( |z_m^i - z_n^j| \geq L \).

The fact that the Hamiltonian matrices are symmetric (i.e. real Hermitian) implies:

\[
\begin{align*}
H_{mn}^{00} &= H_{nm}^{00} \\
H_{mn}^{01} &= H_{mn}^{10} = H_{mn}^{01} . \tag{2.29}
\end{align*}
\]

When the index \( n \) is sorted according to \( z_n^0 \), the matrix \( H_{mn}^{01} \) \( (H_{mn}^{01}) \) becomes strictly lower (upper) triangular:

\[
\begin{align*}
H_{mn}^{01} &= 0 \text{ if } m < n \\
H_{mn}^{01} &= 0 \text{ if } m > n . \tag{2.30}
\end{align*}
\]

Applying Eq. 2.29 to Eq. 2.28 results in

\[
\langle w_m | \hat{H} | w_n \rangle = H_{mn}^{00} + H_{mn}^{01} + H_{mn}^{01} . \tag{2.31}
\]

Either \( H_{mn}^{01} \) or \( H_{mn}^{01} \) will be zero according to Eq. 2.30. To identify the contributions of each term, we calculate matrix elements \( H_{mn}^{ij} \) via integration on a real space grid.\(^1\)

For example, the integral in \( \langle \omega_1^0 | \hat{H} | \omega_2^0 \rangle \) is done only over the real space grid where the effective regions of \( |\omega_1^0\rangle \) and \( |\omega_2^0\rangle \) overlap each other (hatched area in the middle of 0th cell); \( \langle \omega_1^0 | \hat{H} | \omega_2^0 \rangle \) obtained in this way will certainly be negligible. Each periodic image should be treated as a different localized orbital. Hence \( H_{mn}^{ij} \) \( (\text{not } H_{mn}^{(rot)}) \) should be taken as the elements of the real-space Hamiltonian matrix.

Our truncation of periodic images is somewhat arbitrary; however, it becomes exact whenever the MLWF’s are entirely contained in the supercell. The concept of

---

\(^1\)The real space grid used is the same as the potential grid. The number of grid points is \( \sqrt{3}E_{\text{cut}} \times L/\pi \), where \( E_{\text{cut}} \) is in \( \text{Ry} \) and \( L \) in Bohr.
separating periodic images can be easily extended to two- or three-dimensional cases. The boundary is defined by the Wigner-Seitz cell at the center of each periodic image; this is how the boundary is defined in the one-dimensional example. The problem persists even in the case of multiple k-point sampling. The Hamiltonian matrix at each \( \mathbf{k} \) is calculated by applying the same transformation as for the wavefunctions (Eq. 2.2):

\[
H^{(\text{rot})}(\mathbf{k}) = (U^{(\mathbf{k})})^\dagger H(\mathbf{k}) U^{(\mathbf{k})},
\]

and then it is Fourier-transformed into the lattice \( \mathbf{R} \) within the Wigner-Seitz supercell centered around \( \mathbf{R} = 0 \):

\[
H^{(\text{rot})}_{mn}(\mathbf{R}) = \frac{1}{N_{kp}} \sum_{\mathbf{k}} e^{-i\mathbf{k} \cdot \mathbf{R}} H_{mn}^{(\text{rot})}(\mathbf{k}) = \langle w_{m0}|\hat{\mathcal{H}}|w_{n\mathbf{R}} \rangle.
\]

\( |w_{n\mathbf{R}} \rangle \) has a periodicity of the Wigner-Seitz supercell, which would cause the same difficulty. Real-space integration is not so convenient since the Wigner-Seitz supercell is \( N_{kp} \) times larger than the unit cell, whereas they are of the same size in the case of \( \Gamma \) sampling. Another way to truncate is to sample enough k-points and neglect the Hamiltonian matrix elements beyond a certain cutoff distance. For example, in the case of one-dimensional system of lattice vector \( \mathbf{a}_3 \) and \( N_{kp,3} \) k-points, one can apply a criterion like

\[
\langle \omega_{m0}^i | \hat{\mathcal{H}} | \omega_{n\mathbf{R}}^j \rangle = \begin{cases} 
\langle w_{m0}|\hat{\mathcal{H}}|w_{n\mathbf{R}} \rangle & \text{if } |z_{m0}^i - z_{n\mathbf{R}}^j| < (N_{kp,3} \times |\mathbf{a}_3|)/2 \\
0 & \text{if } |z_{m0}^i - z_{n\mathbf{R}}^j| \geq (N_{kp,3} \times |\mathbf{a}_3|)/2
\end{cases}.
\]

2.2.2 Band structure interpolation

From a Wannier representation, one can construct an extended \( \{\phi_{nk}\} \) basis at any \( \mathbf{k} \) vector using Bloch sums of Wannier functions:

\[
\phi_{nk}(\mathbf{r}) = \frac{1}{\sqrt{N_{\mathbf{R}}}} \sum_{\mathbf{R}} e^{i\mathbf{k} \cdot \mathbf{R}} \omega_{n\mathbf{R}}(\mathbf{r}).
\]
In the one-dimensional case shown in the last section, the Hamiltonian matrix at \( k \) is

\[
\langle \phi_{mk} | \hat{H} | \phi_{nk} \rangle = e^{-ikL}H_{nn}^{01} + H_{mn}^{00} + e^{ikL}H_{mn}^{01} = H_{mn}^{00} + e^{ikL}H_{mn}^{01} + e^{-ikL}H_{mn}^{01}.
\] (2.36)

Diagonalizing this Hamiltonian matrix gives us eigenvalues \( \{ \epsilon_{nk} \} \) and eigenvectors \( \{ \psi_{nk} \} \) at any \( k \):

\[
\psi_{nk}(r) = \sum_{m} b_{nk}(m) \phi_{mk}(r).
\] (2.37)

Comparison with the standard non-self-consistent band structure calculation explains why the Wannier representation is advantageous. In the standard approach, eigenfunctions are expanded in the original \( N_{pw} \) planewave basis set:

\[
\psi_{nk}(r) = \sum_{G} c_{nk}(G)e^{i(k+G)\cdot r}.
\] (2.38)

The number of basis function is much larger than the number of MLWFs. In addition, the procedure of constructing the Hamiltonian matrix at each \( k \) is more complicated and much more expensive compared to the Wannier representation where parameterized Hamiltonian matrices are used. Two small \( N \times N \) matrices, \( H^{00} \) and \( H^{01} \), are all we need to generate the band structure over the whole BZ.

We note that the band structure shown in Fig. 2-1 is generated in a different fashion. The band structure is calculated by the Slater-Koster interpolation (or Fourier interpolation):

\[
H_{mn}^{(rot)}(k) = \sum_{R} e^{ikR}H_{mn}^{(rot)}(R).
\] (2.39)

In order to explain the difference, we illustrate in Fig. 2-3 the corresponding real-space Hamiltonians for the two cases. There is only one Wannier function (filled circles in the top panel) per unit cell; the WFs linked together by the real-space periodicity are drawn in the same color. Five-times longer supercell is used for \( \Gamma \) sampling, and five \( k \)-points are used for the simulation with one unit cell, so that the Wigner-Seitz supercell is the same as the supercell for \( \Gamma \) sampling; the electronic
Figure 2-3: Comparison between our real-space integration scheme and the Slater-Koster interpolation scheme. Top: Wannier functions are drawn as circles. The unit cell contains one WF. Both the supercell for Γ-sampling and the Wigner-Seitz supercell are 5-times longer than the unit cell. The numbers correspond to the Hamiltonian matrix element between the WF at the center (black) and the WF where the number is written. The WFs with the same color are related by the periodicity of the (Wigner-Seitz) supercell. Bottom: band structures obtained from the two cases. While they overlap at the \( \mathbf{k} \) vectors sampled (marked as blue lines), slight discrepancies can be seen, which comes from the truncation error.

structures obtained from the two calculations must be identical. The numbers written at the WFs correspond to the Hamiltonian matrix element with the WF at the center (colored in black): top row for our scheme and bottom row for the Slater-Koster scheme (these numbers are arbitrarily chosen and are not physical values). Now it is clearly seen that our real-space integration scheme divides \( H_{mm}^{(rot)} \) between the two nearest periodic images, whereas the Slater-Koster method assigns the whole element to the first-nearest periodic image sitting inside the Wigner-Seitz supercell. The actual cutoff distance for our case is the same as the supercell length, while it is half of it for the Slater-Koster interpolation (similar to the criterion in Eq. 2.34). The band structures obtained from the two methods are slightly different (the
The advantage of both methods is that they guarantee that exact eigenvalues are reproduced at the original $k$ vectors sampled: the eigenvalues from the two calculations exactly overlap at these points (marked as blue lines) in the band structures presented. In some cases, one can observe discrepancies even at the $k$ vectors sampled owing to truncation [30]. In any case, the truncation error or the discrepancies between different schemes can be easily controlled by sampling enough number of k-points or by increasing supercell size, and this is only a technical issue.

\footnote{The elements of the real-space Hamiltonians obtained from our method may converge to zero more smoothly; however, it requires running the first-principles code again, while the Slater-Koster interpolation (or other simple truncation schemes like the one suggested in Eq. 2.34) can be done in the post-processing stage.}
Chapter 3

Quantum transport

Introduction

Classical theories of transport treat electrons as particles. Scattering by phonons, impurities, electrons, etc. govern the motion of electrons and thus their mobility. As the size of devices shrinks to the characteristic scattering length, electrons can pass through the device region experiencing only a few number of scattering incidences. Under these circumstances, an electron wavefunction can be spread over the whole device and classical approximations start to fail.

The most important characteristic length which sets the boundary between quantum and classical transport is the phase-relaxation length, $L_\varphi$. The phase relaxation length is the length over which an electron retains its phase coherency as a wave. Whenever the length of a device $L$ becomes comparable to $L_\varphi$, electrons must be treated as waves instead of particles. Scattering by phonons, electrons, magnetic impurities are main sources that decrease $L_\varphi$.

Another important characteristic length is the momentum-relaxation length, $L_m$. The momentum-relaxation length is the average length over which an electron travels before the motion becomes uncorrelated with its initial momentum. Scattering by phonons and impurity atoms decrease $L_m$.

Relative sizes of $L_m$, $L_\varphi$ and $L$ determine the transport regime. Classical particle-

---

1Scattering sources need to have internal degree of freedom to randomize phase. Elastic scattering by impurity atoms do not affect $L_\varphi$ [31].
like behavior is observed when $L_\varphi < L_m \ll L$. The focus here is the transport regime where phase coherency is preserved, i.e. $L \ll L_\varphi$. A brief introduction to phase-coherent transport is given in Section 3.1, and a practical method of calculating conductance is presented in Section 3.2.

### 3.1 Phase-coherent transport

#### 3.1.1 Ballistic transport

We start from the ideal case where an electron is transmitted through a conductor without being scattered. This is realized when $L \ll L_\varphi, L_m$ and called \textit{ballistic} transport. Fig. 3-1 schematically describes a ballistic conductor. Here we assume a single channel case, i.e., one allowed $k$ vector for a given energy. An electron injected from the left reservoir passes through the conductor with a transmission probability of one. Then the current in $(+)$ direction is

$$ I^+ = \frac{2e}{L} \sum_k v_k f(E_k - \mu_1) = \frac{2e}{L} \sum_k \frac{1}{\hbar} \frac{\partial E_k}{\partial k} f(E_k - \mu_1) = \frac{2e}{\hbar} \int_{-\infty}^{+\infty} f(E - \mu_1) dE , \quad (3.1) $$

where $v_k$ is the group velocity, $\mu_1$ the chemical potential of the left reservoir, $f(E - \mu_1)$ the Fermi-Dirac distribution function, and $\frac{1}{\hbar} \frac{\partial E_k}{\partial k} f(E_k - \mu_1)$ the carrier density at $k$ (a
factor of 2 appears due to spin degeneracy). Similarly, the current in \((-\) direction is

\[ I^- = \frac{2e}{\hbar} \int_{-\infty}^{+\infty} f(E - \mu_2) \, dE, \]  

(3.2)

where \(\mu_2\) is the electrochemical potential of the right reservoir. The net current under finite voltage \(V = (\mu_1 - \mu_2)/e\) between the two electrodes is (Fig. 3-1)

\[ I = I^+ - I^- = \frac{2e}{\hbar} \int_{-\infty}^{+\infty} [f(E - \mu_1) - f(E - \mu_2)] \, dE = \frac{2e^2}{\hbar} \frac{(\mu_1 - \mu_2)}{e}. \]  

(3.3)

Then, the conductance of a single-channel ballistic conductor is

\[ G_0 = \frac{dI}{dV} = \frac{2e^2}{\hbar} = 77.5 \mu S. \]  

(3.4)

\(G_0\) is called the conductance quantum.

### 3.1.2 The Landauer formula

We consider the situation where a disordered conductor is placed in the middle of ideal ballistic leads. The disordered conductor transmits an electron wavefunction with a transmission probability \(T < 1\). Fig. 3-2 schematically shows the overall geometry of this reservoir-lead-conductor-lead-reservoir system.

Landauer derived a formula for the conductance of such system [32]. A disordered
conductor of length \( L \) consists of \( n \) random scatterers with a reflection probability \( r \), giving an overall reflection probability of \( R = 1 - T \). The carrier density on the left side of the conductor is \( 1 + R \) and that on the right side is \( 1 - R \), giving a density gradient

\[
\nabla n = -2R/L . 
\]

(3.5)

The net particle flux \( j \) is

\[

j = -D \nabla n = v (1 - R) ,
\]

(3.6)

where \( v \) is velocity and \( D \) is a diffusion coefficient.

The diffusion coefficient is then given by

\[

D = \frac{vL (1 - R)}{2R} .
\]

(3.7)

From the Einstein relation, which links resistivity to diffusion, we have that the total resistance \( R \) is

\[

R = \frac{2R}{1 - R} \left( \frac{\partial \mu}{\partial n} \right)_T \frac{1}{e^2 v} ,
\]

(3.8)

where \( \mu \) is the chemical potential and \( T \) is the temperature. Simplifying the above equation [33] and including spin degeneracy leads to the Landauer formula

\[

G = \frac{2e^2 \mathcal{T}}{\hbar R} .
\]

(3.9)

One could note that the derivation by Landauer appears inconsistent with Eq. 3.4 when \( \mathcal{T} = 1 \): Eq. 3.4 predicts a finite resistance whereas the Landauer expression indicates zero resistance. The discrepancy is due to the fact that the resistance is of different origin. The finite resistance in Eq. 3.4 is the contact resistance between the lead and the reservoir while the resistance in Eq. 3.9 is that of the conductor itself. The contact resistance is attributed to the limited number of transverse modes in a narrow lead. The reservoir is a macroscopic entity having continuous electronic states, but the narrow lead can only accommodate a single transverse mode: therefore, it cannot transmit all the electron wavefunctions in the reservoir, which accounts for
the finite contact resistance. The overall resistance is the sum of the two:

\[ R = R_{\text{contact}} + R_{\text{conductor}} \]
\[ = \frac{h}{2e^2} + \frac{h}{2e^2} \mathcal{R} \]
\[ = \frac{h}{2e^2} \frac{1}{\mathcal{T}}. \]  

(3.10)

The overall conductance of the system in Fig. 3-2 is then

\[ G = \frac{2e^2}{h} \mathcal{T}. \]  

(3.11)

Extension to a multichannel \((N_{ch})\) conductor is given by [33,34]

\[ G = \frac{2e^2}{h} \text{Tr}(t^\dagger t), \]  

(3.12)

where \(t\) is \(N_{ch} \times N_{ch}\) transmission matrix with elements \(t_{ab}\) corresponding to the transmission coefficient from an ingoing channel \(a\) to an outgoing channel \(b\).

### 3.1.3 localization and fluctuations

We now discuss several features of phase-coherent transport that are distinct from classical transport. First, when an electron suffers frequent elastic scattering while phase coherency is retained \((L_m < L < L_p)\), resistance grows exponentially as \(L\) increases. This is different from the classical limit, where \(\mathcal{R}/(1 - \mathcal{R})\) in Eq. 3.8 is simply additive as a function of the reflection probability of each scatterer \(r\):

\[ \frac{\mathcal{R}}{1 - \mathcal{R}} = n \frac{r}{1 - r}. \]  

(3.13)

Eq. 3.13 is the familiar Ohm’s law, since the number of scatterers \(n\) is proportional to \(L\). In the phase-coherent regime, on the other hand, phase shifts of wavefunctions between adjacent scatterers must be taken into account. When a uniform distribution of the phase shifts between 0 and \(2\pi\) is assumed, an ensemble average of \(\mathcal{R}/(1 - \mathcal{R})\)
yields \[ \left\langle \frac{R}{1-R} \right\rangle = \frac{1}{2} \left( \frac{1+r}{1-r} \right)^n - \frac{1}{2} \right. \] (3.14)

In the limit of large \( n \) (or \( L \)), resistance grows exponentially, which is characteristic of electron transport in the localized regime. The typical localization length \( \xi \) is given by the number of channels available multiplied by the momentum-relaxation length

\[ \xi = N_{ch} \times L_m \] (3.15)

and resistance \( R \) in this localized regime \( ( \xi \ll L) \) is

\[ R \propto e^{L/\xi} \] (3.16)

Even before the system enters into the localized regime, if electrons show diffusive behavior \( (L_m \ll L < \xi) \), the ensemble average of the quantum conductance \( \langle G_Q \rangle \) will be smaller than the classical conductance \( G_{CL} \), due to the enhanced backscattering

\[ \langle G_Q \rangle \sim G_{CL} - \frac{2e^2}{h} \] (3.17)

Another interesting feature of this regime is that the magnitude of conductance fluctuations among different random configurations is constant, regardless of the absolute value of the conductance. The variance of the normalized conductance, \( g \equiv G/(2e^2/h) \), is [35]

\[ \langle \delta g^2 \rangle \approx 1 \] (3.18)

Detailed discussion on the Landauer formula and phase-coherent transport can be found in Refs. [31] and [36].

### 3.2 Quantum conductance from first-principles

In principle, the transmission probability \( T \) can be obtained by solving the Schrödinger equation, as shown for the familiar potential well model in Fig. 3-3. An analytic ex-
Figure 3-3: Scattering of a planewave by a potential well. The transmission probability can be obtained by solving the Schrödinger equation at three regions, and matching the wavefunctions at the boundaries. Note that the transmission probability is a function of $E$.

The expression of $T(E)$ is available in this simple case. However, when the potential is replaced by a real atomic potential, the solution is not straightforward, and several methods for calculating conductance from first-principles have been developed. The system at hand consists of semi-infinite left and right leads, plus a disordered region, as shown in the top panel of Fig. 3-4, which is the center part of Fig. 3-2.\(^2\)

First-principles calculation can obtain the electron wavefunctions in the left and right leads, treated as infinite and perfectly periodic (e.g. complex band structure [37]), or semi-infinite (e.g. Green’s function approaches [38]), or approximate (e.g. a jellium model [39]). The difficulty arises when evaluating transmission probability through the conductor, since that breaks translational symmetry. One reciprocal-space approach has been developed, based on the complex band structure, that explicitly calculates the transmission matrix by matching wavefunctions in the two semi-infinite regions and in the conductor region [37,40]. More widely-used approaches adopt a real-space description of the system, using a localized basis. An added benefit is that the Hamiltonian matrix is diagonally dominant in a localized basis, which allows to treat large disordered system in $O(N)$ fashion. We summarize here the method of Nardelli [38], that is applicable to any type of localized basis. This will be used as our main tool to characterize the electronic transport in disordered systems.

\(^2\)It is not necessary to include the left and right reservoirs since the contact resistance is trivially obtained from the number of eigenchannels in the left and right leads.
3.2.1 Hamiltonian matrix in a localized basis

Our solution to deal with infinite but non-periodic systems is to describe the Hamiltonian in a real-space localized basis. As shown in Fig. 3-4, the overall Hamiltonian matrix is divided into five submatrices, which are of the left \( H_L \) and the right \( H_R \) leads, the conductor \( H_C \), and the coupling regions between the leads and the conductor \( H_{LC}; H_{CR} \):\(^3\)

\[
H \rightarrow H_L + H_{LC} + H_C + H_{CR} + H_R. \tag{3.19}
\]

These submatrices have an infinite size except \( H_C \). However, in a localized basis \( \omega_i \) the matrix element \( \langle \omega_i | \hat{H} | \omega_j \rangle \) will approach zero as \( | \langle r \rangle_i - \langle r \rangle_j | \) increases \( (| \langle r \rangle_i = \int r | \omega_i(r) |^2 dr) \). Therefore, one can define a principal layer as a region large enough to have interactions only with the nearest neighboring principal layer, i.e. \( \langle \omega_i^n | \hat{H} | \omega_j^n \rangle = 0 \) (\( \omega_i^n \) is \( i_{th} \) basis function belonging to the \( n_{th} \) principal layer) [41]. As shown in Fig. 3-4, each lead is composed of a semi-infinite array of principal layers: the semi-infinite \( H_{\{L,R\}} \) is then constructed by arranging \( H_{L}^{00} \) and \( H_{L}^{01} \) repeatedly in a block tridiagonal form \( (H_{\{L,R\}}^{00} \text{ stands for the matrix elements between the orbitals in the same principal layer and } H_{\{L,R\}}^{01} \text{ for the orbitals in consecutive principal layers, as described in the bottom left panel of Fig. 3-4). Then the Hamiltonian matrix of the whole system has the following form:

\[
\begin{pmatrix}
H_L & H_{LC} & 0 \\
H_{LC}^\dagger & H_C & H_{CR} \\
0 & H_{CR}^\dagger & H_R
\end{pmatrix}
= \begin{pmatrix}
\ddots & \ddots & \vdots \\
\cdots & H_{L}^{00} & H_{L}^{10} & 0 & 0 & 0 & \cdots \\
\cdots & H_{L}^{01} & H_{L}^{00} & h_{LC} & 0 & 0 & \cdots \\
\cdots & 0 & h_{LC}^\dagger & H_C & h_{CR} & 0 & \cdots \\
\cdots & 0 & 0 & h_{CR}^\dagger & H_{R}^{00} & H_{R}^{01} & \cdots \\
\cdots & 0 & 0 & 0 & H_{R}^{01} & H_{R}^{00} & \cdots \\
\ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix}. \tag{3.20}
\]

\(^3\)The interaction between the left and right lead, \( H_{LR} \), can be suppressed by choosing the conductor region long enough.
Figure 3-4: Illustration of the set-up for the calculations in a lead-conductor-lead system. The whole system is described as an array of principal layers. Left: lead is studied in a periodic supercell containing one principal layer of the lead. Right: conductor and leads are studied in a periodic supercell containing the conductor plus at least one principal layer of each lead.

\( H_{(L,R)}^{00} \) and \( H_{(L,R)}^{01} \) are obtained from first-principles simulations of a lead supercell comprising at least one principal layer of the lead under periodic boundary conditions (Section 2.2.1). To obtain \( H_{LC} \) (or \( H_{CR} \)) and \( H_{C} \), a conductor supercell including the conductor region and one principal layer of the left (or right) lead is simulated (the bottom right panel of Fig. 3-4). The supercell must be large enough such that the potential profile inside the lead part in a conductor supercell converges to that of an unperturbed lead. This ensures a seamless connection from the conductor to the semi-infinite leads. When the left and the right lead are of the same kind, calculations will involve two separate simulations to determine five submatrices: one for \( H_{L}^{00} \) and \( H_{L}^{01} \), and one for \( H_{LC}, H_{CR}, \) and \( H_{C} \). When they are of different kind, three separate simulations and two more matrices \( H_{R}^{00} \) and \( H_{R}^{01} \) are obviously required.
3.2.2 Transmission probability in the Green's function formalism

Evaluating the transmission probability from these matrices is still not a trivial task. We use here the multichannel extension of the Landauer formula (Eq.3.9), expressed in terms of Green's functions [42]:

$$ G(E) = \frac{2e^2}{\hbar} \text{Tr}(tt') = \frac{2e^2}{\hbar} \text{Tr}(\Gamma_L G^r_C \Gamma_R G^a_C) ,$$

(3.21)

where $G^{(r,a)}_C$ are the retarded and advanced Green’s functions of the conductor, and $\Gamma_{\{L,R\}}$ are the functions that describe the coupling of the conductor to the leads. The retarded Green’s function for the whole system satisfies

$$ G^r = [E + i\eta - H]^{-1} \quad (\eta \to 0^+) ;$$

(3.22)

similarly for the advanced one, we have

$$ G^a = [E - i\eta - H]^{-1} \quad (\eta \to 0^+) .$$

(3.23)

The advanced Green’s function $G^a$ is the Hermitian conjugate of the retarded one, i.e. $G^a = (G^r)^\dagger$. We solve for the retarded case, unless specified otherwise. Calculating Green’s functions from the previous equations involves inversion of an infinite matrix for the whole system. We exploit the fact that the Green’s function can be partitioned into submatrices corresponding to the conductor and the leads:

$$ \begin{pmatrix} G_L & G_{LC} & G_{LR} \\ G_{CL} & G_C & G_{CR} \\ G_{RL} & G_{RC} & G_R \end{pmatrix} = \begin{pmatrix} \epsilon - H_L & -H_{LC} & 0 \\ -H_{LC}^\dagger & \epsilon - H_C & -H_{CR} \\ 0 & -H_{CR}^\dagger & \epsilon - H_R \end{pmatrix}^{-1} ,$$

(3.24)
where $\epsilon = E + i\eta$.

After some algebra, a simplified expression for $G_C$ is obtained:

$$G_C = (\epsilon - H_C - \Sigma_L - \Sigma_R)^{-1},$$

(3.25)

where $\Sigma_L (= H_{LC}^\dagger G_L H_{LC})$ and $\Sigma_R (= H_{CR} G_R H_{CR}^\dagger)$ are self-energy terms arising from the interaction with the semi-infinite leads. $H_C + \Sigma_L + \Sigma_R$ can thus be viewed as an effective Hamiltonian for the conductor interacting with the semi-infinite leads. We note in passing that the imaginary part of the Green’s function gives the local density of states for the conductor part [31]:

$$\mathcal{N}_C(E) = -\frac{1}{\pi} \text{Im}(\text{Tr}[G_C(E)]).$$

(3.26)

The evaluation of $\Sigma_L$ (or $\Sigma_R$) appears to require multiplication of infinite-sized $H_{LC}$ and $G_L$. However, we only need to consider the non-zero part of of $H_{LC}$, labelled $h_{LC}$ in Eq. 3.20, which is a small $N_L \times N_C$ submatrix ($N_L$ and $N_C$ are the number of basis functions in the principal layer of the lead and in the conductor, respectively). Then the required portion of $G_L$ is an $N_L \times N_L$ submatrix that is in contact with the conductor. This matrix is called the surface Green’s function, and we label it as $G_{L0}$, since it is the Green’s function of the 0th principal layer located at the surface of the semi-infinite left lead. $\Sigma_L$ is thus given by

$$\Sigma_L = H_{LC}^\dagger G_L H_{LC} = h_{LC}^\dagger G_{L0} h_{LC}. $$

(3.27)

Once $G_{L0}^{(0)}$ and then $\Sigma_{\{L,R\}}$ are known, the coupling functions $\Gamma_{\{L,R\}}$ in Eq. 3.21 can be readily obtained [31]:

$$\Gamma_{\{L,R\}} = i[\Sigma_{\{L,R\}} - \Sigma_{\{L,R\}}^a],$$

(3.28)

where the advanced self-energy $\Sigma_{\{L,R\}}^a$ is the Hermitian conjugate of the retarded self-energy $\Sigma_{\{L,R\}}^r$. The Green’s function of the conductor $G_C$ can be calculated by
plugging $\Sigma_{(L,R)}$ in Eq. 3.25, and then we have all the matrices required to evaluate 
The transmission probability using Eq. 3.21. The problem narrows down to obtaining 
the surface Green’s function.

### 3.2.3 Surface Green’s functions

We follow an efficient procedure by López Sancho et al. [43] to obtain the surface 
Green’s functions of the semi-infinite left and right leads. Their formalism allows to 
include $2^i$ principal layers at every $i_{th}$ iteration and thus assures fast convergence. 

Going back to the definition of Green’s function, in a principal-layer decomposition, 
we have

$$
\begin{pmatrix}
\epsilon - H^{00} & -H^{01} & 0 & \cdots \\
-H^{01\dagger} & \epsilon - H^{00} & -H^{01} & \cdots \\
0 & -H^{01\dagger} & \epsilon - H^{00} & \cdots \\
\vdots & \vdots & \vdots & \ddots
\end{pmatrix}
\begin{pmatrix}
G^{00} & G^{01} & G^{02} & \cdots \\
G^{10} & G^{11} & G^{12} & \cdots \\
G^{20} & G^{21} & G^{22} & \cdots \\
\vdots & \vdots & \vdots & \ddots
\end{pmatrix} = I . 
\quad (3.29)
$$

Comparing the $(m+1,1)$ submatrix of the left and the right side gives the following 
set of equations:

$$m = 0 , \quad (\epsilon - H^{00})G^{00} = I + H^{01}G^{10}$$

$$m = 1 , \quad (\epsilon - H^{00})G^{10} = H^{01\dagger}G^{00} + H^{01}G^{20}$$

$$\cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cd...
It can be seen that after Eq.3.31 is applied \(i\) times, we have

\[
G^{n0} = t_i G^{n-2^i,0} + \tilde{t}_i G^{n+2^i,0}, \tag{3.33}
\]

where

\[
t_i = (I - t_{i-1} \tilde{t}_{i-1} - \tilde{t}_{i-1} t_{i-1})^{-1} t_i^2, \\
\tilde{t}_i = (I - t_{i-1} \tilde{t}_{i-1} - \tilde{t}_{i-1} t_{i-1})^{-1} \tilde{t}_i^2. \tag{3.34}
\]

The first line of Eq. 3.30 tells that \(G^{00}\) can be obtained by solving for \(G^{10}\). This can be solved recursively, using Eq. 3.33 (\(G^{2^i0} = t_i G^{00} + \tilde{t}_i G^{2^{i+1}0}\) when \(n = 2^i\)):

\[
G^{10} = t_0 G^{00} + \tilde{t}_0 G^{20} = (t_0 + \tilde{t}_0 t_1) G^{00} + \tilde{t}_1 G^{40} = (t_0 + \tilde{t}_0 t_1 + \cdots + \tilde{t}_0 \cdots \tilde{t}_{i-1} t_i) G^{00} + \tilde{t}_i G^{2^{i+1}0}. \tag{3.35}
\]

The calculation is repeated until \(||t_i||, ||\tilde{t}_i|| < \varepsilon\) with an arbitrarily small threshold \(\varepsilon\). We note that, as implied from the subscripts in Eq. 3.35, \(i_{th}\) iteration effectively includes \(2^i\) layers. \(G^{10}\) is then given by

\[
G^{10} = T G^{00}, \tag{3.36}
\]

with

\[
T = t_0 + \tilde{t}_0 t_1 + \cdots + \tilde{t}_0 \cdots \tilde{t}_{i-1} t_i. \tag{3.37}
\]

Finally, from Eq. 3.30 and Eq. 3.36, the surface Green's function is

\[
G^{00}(E) = (\varepsilon - H^{00} - H^{01} T)^{-1}. \tag{3.38}
\]

\(t_i\) and \(\tilde{t}_i\) will converge to zero since they are at least order \(2^i\) in \(H^{01}\). As discussed in Section 2.2.1, \(H^{01}\) is a strictly lower triangular matrix. The characteristic polynomial of \(H^{01}\) calculated from \(\det |H^{01} - \lambda I| = 0\) will be simply \(\lambda^N = 0\) when \(H^{01}\) is an \(N \times N\) strictly lower or upper triangular matrix, which proves \((H^{01})^N = 0\).
3.2.4 Summary

As a summary of these method chapters, we outline in Fig. 3-5 the full prescription for calculating band structure and quantum conductance from first-principles simulations. The procedure requires an electronic-structure code [44], a Wannier localization code [45], and a tight-binding transport code [46].
First-Principles Simulations
Electronic Ground State
(Section 1.2)
Bloch Functions $\psi$ at $\Gamma$
(Section 1.4)

Disentanglement and Localization
$\Omega = \Omega_I + \tilde{\Omega}$
Maximally-Localized Wannier Functions
$w = U\psi$
(Section 2.1.2)

Real-space Hamiltonian matrix
Lead: $H^{00}, H^{01}$
Conductor: $H_C, h_{LC}, h_{CR}$
(Sections 2.2.1 and 3.2.1)

Surface Green's Function
$H^{00}, H^{01} \rightarrow G^{00}$
(Section 3.2.3)

Green's Function Formalism
$H_C, h_{LC}, h_{CR}, G^{00}$
(Section 3.2.2)

Quantum Conductance
Density of States
(Section 3.2.2)

Figure 3-5: Summary of the method.
Chapter 4

Pristine carbon nanotubes

Introduction

Since their discovery in 1991 [10], carbon nanotubes (CNTs) have attracted much attention for their exceptional electronic, mechanical, and thermal properties and significant progress has been made in understanding, characterizing, and controlling their physical properties [47, 48]. Among their characteristics, electronic transport properties stand out:

- Metallic carbon nanotubes are ideal quasi-one-dimensional quantum wires that can sustain extremely high current densities.
- The high surface-to-volume ratio makes them an ideal candidate for sensors.
- The high aspect-ratio and mechanical stability qualifies them for electron emitters.

These are only some of the wide varieties of applications foreseen for carbon nanotubes.

In this chapter, we discuss first the electronic structure of carbon nanotubes. The method described in Chapters 2 and 3 is expanded and validated here, specifically for carbon nanotubes. In the last section, we summarize experimental findings that highlight phase-coherent transport in CNTs.
Table 4.1: Structural parameters for \((n,n)\) armchair and \((n,0)\) zigzag nanotubes. \(r\) is the radius, \(N_A\) the number of carbon atoms in a unit cell, \(C_h\) the chiral vector, and \(T\) the translational vector.

|                | \(|C_h|\) | \(|T|\) | \(r\) | \(N_A\) |
|----------------|----------|--------|-------|--------|
| armchair       | \(3n\) \(d_{CC}\) | \(\sqrt{3} \cdot d_{CC}\) | \(3n\) \(d_{CC}/(2\pi)\) | \(4n\) |
| zigzag         | \(\sqrt{3}n\) \(d_{CC}\) | \(3d_{CC}\) | \(\sqrt{3}n\) \(d_{CC}/(2\pi)\) | \(4n\) |

### 4.1 Crystalline and electronic structure

The structure of a carbon nanotube can be represented as a rolled-up graphene sheet (graphene is one layer of graphite). The unit cell of a carbon nanotube is defined by the chiral vector \(C_h\) and translational vector \(T\). \(C_h\), which is the roll-up direction and defines the circumference of the tube, is expressed as a linear combination of the two lattice vectors of graphene, \(a_1\) and \(a_2\), as in Fig. 4-1:

\[
C_h = na_1 + ma_2 \equiv (n, m) \quad (n, m \text{ are integers, } 0 \leq |m| \leq n).
\] (4.1)

The index \((n,m)\) of a CNT is thus derived from that of the chiral vector.

The translational vector \(T\) defines the translational period along the tube axis and is expressed as

\[
T = \frac{2m + n}{d_R}a_1 - \frac{2n + m}{d_R}a_2,
\] (4.2)

where \(d_R\) is the greatest common divisor of \((2m + n)\) and \((2n + m)\) [47]. The chiral vector \(C_h = 4a_1 + 2a_2\) and the translational vector \(T = 4a_1 - 5a_2\) of a \((4,2)\) CNT are shown in Fig. 4-1. Carbon nanotubes are categorized into chiral and achiral ones according to their symmetry. Achiral CNTs have mirror symmetry planes containing the tube axis. These are the \((n,n)\) armchair and \((n,0)\) zigzag tubes. All other tubes are chiral. A chiral tube and two types of achiral tubes are shown in the right panels of Fig. 4-1. The unrelaxed values of the structural parameters can be expressed in terms of the indices \(n\) and \(m\) and the carbon-carbon bond distance \(d_{CC}\) (1.42\(\text{Å}\) in graphene). These structural parameters for \((n,n)\) armchair and \((n,0)\) zigzag tubes.
Like their crystalline structure, the electronic structure of a CNT can be derived from that of two-dimensional graphene; Fig. 4-2 illustrates the relationship between the two. Graphene is a zero-gap semiconductor: conduction and valence band touch each other at the corners of the Brillouin zone (marked as $K$). In the case of carbon nanotubes, confinement of the electron wavefunction along the circumference only allows discretized wave vectors in the direction of $C_h$; the red lines indicate the available $k$ vectors for a $(5,5)$ and for an $(8,0)$ CNT. When these lines pass through $K$, a CNT will be metallic; otherwise it is semiconducting. In terms of the $(n,m)$ index, this zone folding can be summarized as following:

$$\text{mod}(n - m, 3) = \begin{cases} 
0 & : \text{CNT is metallic} \\
1, 2 & : \text{CNT is semiconducting} 
\end{cases}.$$  \tag{4.3}

Thus a $(5,5)$ CNT is metallic while an $(8,0)$ CNT is semiconducting. Since zone
folding ignores curvature and the accompanying orbital rehybridizations, exceptions exist in small diameter tubes [50]. However, armchair CNTs are always metallic and have two degenerate eigenstates at the Fermi energy, regardless of diameter.

A more detailed discussion can be found in Refs. [47] and [48], and references therein.

Simulation details

Our First-principles calculations have been performed using density-functional theory, within the planewaves and ultrasoft pseudopotentials framework. The Perdew-Burke-Ernzerhof (PBE) generalized gradient approximation is used throughout for the exchange-correlation functional [17], together with a planewave cutoff of 30 Ry
Figure 4-3: Structural parameters of a carbon nanotube in the tetragonal primitive unit cell used. (a) $a$ and $b$ are cell dimensions in the $x$ and $y$ directions. $r$ is a radius of a CNT. A vacuum region of length $V_{ac}$ minimizes the interaction between periodic images. Cell dimension $c$ in the $z$ direction for (a) an armchair tube and for (b) a zigzag tube.

and a charge density cutoff of 240 $Ry$. This level of theory and parameters used have been shown to accurately predict the crystalline structures and the physical properties of diamond and graphite [51]. We use our Quantum-ESPRESSO package as a main tool for the simulations [44]. For $\Gamma$-point electronic structure calculations and structural optimizations, the Car-Parrinello (damped) molecular dynamics algorithm is used [52,53].

Fig. 4-3 shows the structural parameters of the tetragonal primitive cell inside which a carbon nanotube is placed. Periodic boundary conditions generate mirror images in each direction. That naturally creates an infinite nanotube along the $z$ direction, but in the $x, y$ directions the cell dimension should be large enough to
Figure 4-4: Energy per carbon atom for a (5,5) and an (8,0) CNT as a function of the number of k-points in the z direction. Finite smearing parameter improves convergence. Top panels: (5,5) CNT. Bottom panels: (8,0) CNT.

To avoid interactions between mirror images: a vacuum of length $V_{ac}$ is inserted for this purpose. The cell dimensions $a, b, c$ in the $x, y, z$ directions are

$$a = b = 2r + V_{ac}, \quad c = |T|.$$ 

As discussed in Section 1.4.1, k-point convergence must be checked. In the $x, y$ directions, only a single k-point is required since the system is isolated in the $xy$ plane. In Fig. 4-4 we plot energy per atom for a (5,5) and (8,0) CNT as a function of the number of k-points along the tube axis. The strong fluctuation and slow convergence in the case of (5,5) CNT (panel (a)) is a characteristic of metallic systems. A cold smearing of 0.03 $Ry$ slightly improves convergence [54], and an accuracy of 1 $meV/atom$ is achieved when $1 \times 1 \times 9$ ((5,5) CNT) or $1 \times 1 \times 6$ ((8,0) CNT) regular mesh is used. This information will be used to determine sampling in Chapters 6 and 7.

The method introduced in Chapter 2 is aimed at large systems where single $\Gamma$-
point sampling is sufficient. The primitive unit cell of a (5,5) and an (8,0) CNT is too small to apply Γ-point sampling. In order to make our Γ-point calculations reasonably accurate, the cell dimension in the $z$ direction is increased. A five-times or three-times longer supercell is used, with a cell dimension $c$ of

$$(5,5) \text{ CNT}: c = 5\sqrt{3}d_{CC}, \quad (8,0) \text{ CNT}: c = 9d_{CC}.$$  

A vacuum $V_{ac} = 12\text{Å}$ between sidewalls is used in the $x, y$ directions.

4.2 Maximally-localized Wannier functions and band structure interpolation

Maximally-localized Wannier functions are generated following the procedure described in Section 2.1.2. This is done as a post-processing step after the ground state Kohn-Sham orbitals have been obtained. The calculation of overlap matrices needed in Eq. 2.14 has already been implemented by Wu and Sharma [55] in Quantum-ESPRESSO; disentanglement and localization has been part of the present work.

Let us examine the MLWFs of graphene first. A $4 \times 4$ supercell (32 atoms) is used. MLWFs for the occupied bands comprise $\sigma$-bonding MLWFs at every bond center and $\pi$-bonding MLWFs at every other carbon atom, as shown in Fig. 4-5 (there are two equivalent localization minima, corresponding to the two choices of sublattices on which the $\pi$-bonding MLWFs sit). The MLWFs of a (5,5) and an (8,0) tube in Fig. 4-6 are only slightly different from those of graphene: curvature breaks symmetry, and so there are two inequivalent $\sigma$-bonding MLWFs instead of only one. The center of the $\sigma$-MLWFs is located slightly outside the tube, whereas that of $\pi$-bonding MLWFs is located slightly inside the tube. The fact that we find localized MLWFs in graphene and a (5,5) CNT is somewhat surprising. The reason might be due to the fact that these are zero-gap semiconductors. In truly metallic systems, we should not expect

\footnote{The isosurface value is chosen such that it contains 70% of the total charge of an MLWF.}
to obtain localized Wannier functions from the occupied manifold alone. In the case of a (5,0) CNT, an MLWF with a large spread is found at the center of the tube.

The band structure of a (5,5) and an (8,0) CNT is then calculated following the prescriptions of Section 2.2.2. This is shown in Fig. 4-7, where dots represent a standard (non-self-consistent) diagonalization everywhere in the BZ and the solid lines come from the Wannier interpolation. To make the comparison easier, the BZ of the supercell is unfolded. Blue lines indicate the $k$ vectors which correspond to $\Gamma$ for the supercell; eigenvalues are accurately reproduced at these points. The quality of the interpolated bands obtained from the occupied MLWFs is quite poor:
Figure 4-7: Band structure of a (5,5) and an (8,0) CNT obtained from MLWFs interpolation scheme, using the MLWFs in the occupied subspace (solid lines). Red dots are from full diagonalization using Quantum-ESPRESSO. The agreement between the two is poor when the MLWFs of the occupied space (Fig. 4-6) are used as a basis set. The deviation in the case of the metallic (5,5) CNT is more pronounced than for the semiconducting (8,0) CNT.

the eigenvalues (solid lines) deviate significantly from the correct values (red dots). The semiconducting (8,0) CNT behaves better than the metallic (5,5) CNT, but the agreement is not satisfactory either.

In more explicit terms, our MLWFs interpolation implies that the periodic part of the Bloch function at an arbitrary $k$ can be expressed as a linear combination of the periodic functions at $\Gamma$:

$$u_{nk} = \sum_{m} b_{nk}(m) u_{m\Gamma} .$$

Deviation from the exact result originates in restricting the summation only up to the $N$ occupied bands. In a (5,5) CNT, $N = 200$ while the virtually complete planewave basis set has $N_{pw} = 5200$, which makes a huge difference. Of course using up to $N_{pw}$ MLWFs would achieve the same accuracy of the standard method, but our goal is to find a minimal but virtually complete basis. What we are trying to do is similar in spirit with the $k \cdot p$ method [19, 56]. In $k \cdot p$, eigenvalues at $k$ are calculated

---

2This is not entirely true for the Wannier interpolation scheme though.
from a perturbative approach without explicitly diagonalizing the Hamiltonian, giving around band extrema [57]

\[ E_{nk} = E_{n\Gamma} + \frac{|k|^2}{2} + \sum_m \frac{|k \cdot (u_{m\Gamma}|p|u_{m\Gamma})|^2}{E_{n\Gamma} - E_{m\Gamma}}. \]  

Contribution from high-energy eigenstates becomes smaller and smaller as \( |E_{n\Gamma} - E_{m\Gamma}| \) increases. Therefore, including enough unoccupied states will systematically improve the quality of the interpolation; indeed, we see that even in our case the interpolation error is much reduced at energies few eV below the Fermi level. In our approach though, an optimal subset of unoccupied states that contributes the most to the \( k \)-space dispersion will be selected by the disentanglement procedure.

The disentanglement procedure is applied here to both graphene and carbon nanotubes. For graphene, the relevant bands near the Fermi level are \( \pi \)- and \( \pi^* \)-bands derived from a linear combination of \( p_z \)-orbitals. There is one \( \sigma \)-orbital per each bond and one \( p_z \)-orbital per each carbon atom, which add up to 2.5 orbitals per carbon atom. They are successfully disentangled from the outer window that includes the full set of \( \pi \)- and \( \pi^* \)-bands. Subsequent localization generates \( \sigma \)- and \( p_z \)-MLWFs as shown in Fig. 4-8. The \( \sigma \)-bonding MLWF is the same as the one in Fig. 4-5 since the subspace of \( p_z \)-MLWFs does not mix. The same strategy is applied to the (5,5) and (8,0) CNT. The resulting disentangled MLWFs are shown in Fig. 4-9. A slight change in the \( \sigma \)-bonding MLWFs is observed: \( \sigma \)- and \( \pi \)-band are not exactly separated in this case. Parameters for the disentanglement procedure are summarized in Table 4.2. The interpolated band structures of Fig. 4-10 obtained from diagonalization in the disentangled subspace show great improvement over the ones obtained from the occupied subspace alone (Fig. 4-7). The number of basis functions has been increased by only 0.5 orbitals per atom, i.e. still orders of magnitude smaller than the complete \( N_{pw} \) basis. The result underscores the importance of the disentangling

\(^{3}\)The specific algorithm we use to obtain the ground state wavefunction gives randomly rotated Kohn-Sham eigenstates since the electron density and the ground state energy are invariant upon unitary transformation [53]. Eigenstates need to be recovered before the disentanglement procedure, in order to select the inner and outer energy window.
Figure 4-8: Disentangled maximally-localized Wannier functions of graphene: (a) $\sigma$-bonding MLWF; (b) $p_z$-MLWF. $\sigma$- and $\pi$-manifolds are separated by symmetry; the $\sigma$-bonding MLWF is therefore identical to that of the occupied subspace shown in Fig. 4-5(a).

Table 4.2: Parameters used for disentanglement of $\pi$- and $\pi^*$-bands in the case of a (5,5) CNT, an (8,0) CNT, and graphene. $E_i$ denotes the upper bound of the inner window, and $E_o$ denotes that of the outer window. The reference is set to the Fermi level for the (5,5) CNT and graphene, and the highest occupied level for the (8,0) CNT.

<table>
<thead>
<tr>
<th></th>
<th>$N$</th>
<th>$N_i$</th>
<th>$N_o$</th>
<th>$E_i$</th>
<th>$E_o$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(5,5)</td>
<td>250</td>
<td>223</td>
<td>300</td>
<td>2.81</td>
<td>6.82</td>
</tr>
<tr>
<td>(8,0)</td>
<td>240</td>
<td>219</td>
<td>320</td>
<td>3.61</td>
<td>8.05</td>
</tr>
<tr>
<td>graphene</td>
<td>80</td>
<td>64</td>
<td>167</td>
<td>0</td>
<td>13.87</td>
</tr>
</tbody>
</table>

procedure, that allows us to extract a maximally-connected subspace. Some more advantages also follow: The disentangled subspace covers also exactly bands up to few $eV$ above the Fermi level (these are needed to describe electronic transport under finite bias or varying gate voltage). Also the spreads of the MLWFs obtained from the disentangled space are smaller (as summarized in Table 4.3).

Table 4.3: Spread of the maximally-localized Wannier functions of a (5,5) CNT, an (8,0) CNT, and graphene.

<table>
<thead>
<tr>
<th>spread ($\text{Å}^2$)</th>
<th>occupied</th>
<th>disentangled</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\sigma$</td>
<td>$\pi$</td>
</tr>
<tr>
<td>(5,5)</td>
<td>0.652 / 0.719</td>
<td>2.344</td>
</tr>
<tr>
<td>(8,0)</td>
<td>0.696 / 0.657</td>
<td>2.540</td>
</tr>
<tr>
<td>graphene</td>
<td>0.551</td>
<td>1.885</td>
</tr>
</tbody>
</table>
Disentangled, maximally-localized Wannier functions of a (5,5) and an (8,0) CNT. Two \( \sigma \)-bonding MLWFs (red and blue) and a \( p_z \)-MLWF (violet) are shown. \( \sigma \)- and \( \pi \)-manifold are mixed due to curvature; the \( \sigma \)-bonding MLWF is therefore different from that of the occupied subspace shown in Fig. 4-6(a).

Band structures of a (5,0) and an (8,0) CNT obtained from the interpolation scheme using the MLWFs of the disentangled subspace (Fig. 4-9 as a basis set). Solid lines are from our method and red dots are from Quantum-ESPRESSO code. The agreement between the two is excellent.
4.3 Real-space Hamiltonian matrix

4.3.1 Periodic cell

The accuracy of transferable real-space Hamiltonian matrices and of the band structure interpolation relies on the localized character of MLWFs. In the case of a (5,5) and an (8,0) CNT, the spread of the \( p_z \) orbitals in the \( z \) direction is 0.36 Å\(^2\) and 0.29 Å\(^2\) respectively, which is small compared to the supercell length of \( \sim 12 \) Å, and our truncation error will be negligible. Fig. 4-11 shows a plot of \( \langle \omega_0 | \hat{H} | \omega_n \rangle \) for a (5,5) CNT, where \( \omega_0 \) is a \( p_z \)-MLWF located at \( z = 0 \) and \( \omega_n \) sweeps through all the MLWFs moving along the tube axis: in the bottom panel, yellow dots mark the center of these MLWFs. Two results are compared: one is from a 100-atom supercell and the other is from a 240-atom supercell. The overall agreement of two results is good. As the

---

Figure 4-11: Variation of Hamiltonian matrix element, \( \langle \omega_0 | \hat{H} | \omega_n \rangle \), of a (5,5) CNT as a function of \( |\langle r \rangle_0 - \langle r \rangle_n| \). \( \omega_0 \) is a \( p_z \)-MLWF located at \( z = 0 \), and \( \omega_n \) sweeps through MLWFs along the tube axis as illustrated at the bottom.

---

\(^4\)The degree of localization and its decaying behavior has been discussed. Kohn proved that in one-dimensional case Wannier function decays exponentially [58], and He and Vanderbilt showed that there is power-law prefactor, \( w(x) \approx x^{-\alpha}e^{-bx} \) [59].
distance between two orbitals is increased, \( \langle \psi_0 | \hat{H} | \psi_n \rangle \) decays almost exponentially. It becomes smaller than \( 10^{-2} \text{eV} \) when the distance is larger than \( 3\sqrt{3}d_{CC} \). Parameterization up to the third-nearest neighbor in graphene, which is separated by \( 2d_{CC} \), already gives a good description of band structures of CNTs [60]. We tested a few different cutoff distances, as indexed in Fig. 4-11: i.e. we neglect \( \langle \psi_0 | \hat{H} | \psi_n \rangle \) when \( |z_0 - z_n| \) is above the cutoff distance. Band structures obtained applying these four cutoff distances are plotted in Fig. 4-12. The contribution of MLWFs separated by more than \( 2\sqrt{3}d_{CC} \) is not appreciable, and the principal layer can be reduced to two or three times of the conventional unit cell as a very good approximation. However, this result does not imply that the supercell size itself can be reduced. One should keep in mind that a supercell should be big enough to make \( \Gamma \)-point sampling reasonably accurate and to avoid truncation errors due to the periodic images of MLWFs.

4.3.2 Fragment

It would be interesting to compare MLWFs obtained for a nanotube fragment with those in the infinite tube. If the fragment is long enough, the charge density in the center region will converge to the infinite limit; the question is how fast MLWFs will
Figure 4-13: Charge density and electrostatic potential of a fragment of a (5,5) CNT, $C_{110}H_{20}$. Ending carbon atoms are saturated with hydrogen atoms. Charge density in the center of the fragment converges to that of the infinite tube.

also converge. A fragment of a (5,5) CNT, $C_{110}H_{20}$, is used as a test case. Atomic coordinates of 90 carbon atoms in the center are fixed to those of the infinite (5,5) tube; 20 atoms at each end are relaxed. Fig. 4-13 shows the charge density and the electrostatic potential profile. The plotted values are obtained as following: a planar average in the $xy$ plane is first taken and then it is averaged over one atomic layer ($\sqrt{3}d_{CC}$) along the $z$ direction, resulting in a constant value for the infinite tube. Charge density of the center region surrounded by a red box converges to that of an infinite (5,5) tube. Disentanglement and localization$^5$ of 295 orbitals, corresponding to 2.5 orbitals per carbon and 1 orbital per hydrogen, generates a regular array of MLWFs very similar to that in Fig. 4-9(a). $\langle \omega_0 | \hat{H} | \omega_n \rangle$, plotted in the left panel of Fig. 4-14, proves that MLWFs in the center region of the fragment are indeed converged to those of an infinite tube. In the right panel of Fig. 4-14, we show that the band structure calculated from the tight-binding parameters obtained from the Hamiltonian matrix of the center part (solid lines) closely matches that of an infinite tube (red dots).

$^5$The concept of “disentanglement” does not fit to the isolated molecular systems: eigenstates are discretized and “band” does not exist. In this case, it is better to regard disentanglement and localization as a two step procedure of minimizing the overall spread.
Figure 4-14: $\langle \omega_0 | \hat{H} | \omega_n \rangle$ for the C$_{110}$H$_{20}$ fragment and band structure calculated from the Hamiltonian matrix of a central region in the fragment. (a) Variation of $\langle \omega_0 | \hat{H} | \omega_n \rangle$ as compared with that of an infinite (5,5) tube (Fig. 4-11). The two results agree very well. (b) Band structure derived from the Hamiltonian matrix of the central region surrounded by a red box in Fig. 4-13 (solid lines). The overlapping with the band structure of an infinite tube (red dots) is also remarkable.

The fact that bulk MLWFs are recovered in the center of the fragment is important. In the lead-conductor-lead geometry shown in Fig. 3-4, a conductor supercell consists of the disordered conductor and one principal layer of the lead. It is not guaranteed a priori that the MLWFs in the lead part inside the supercell would be the same as the MLWFs in the bulk lead. The result shown here is encouraging since it suggests that MLWFs could be determined by local charge density, exhibiting weak dependence on an environment.

### 4.4 Quantum conductance

As the final step in the validation of our method, the quantum conductance is calculated following the procedure outlined in Section 3.2. We use the DOSQC code of Nardelli [38, 46]: it calculates quantum conductance and local density of states for
given Hamiltonian matrices. The formalism in Section 3.2 can be applied to any kind of localized basis set [6, 61-63]. Compared to other localized basis, MLWFs have the following advantages:

- Our starting point is a first-principles calculation using planewave basis set; this is a virtually complete set and more accurate than most methods using localized basis sets.

- The information is seamlessly transferred to real-space without losing the accuracy of the original calculations.

- MLWFs are a minimal basis set, which implies minimum computational cost in calculating quantum conductance.

A similar approach has also been introduced by Calzolari et al. [64,65] and Thygesen et al. [66]; the focus here is to extend the method to large scale systems.

One technical point is worth mentioning before discussing the result. Special care must be taken to ensure that the sign of MLWFs is consistent. $w_n(r)$ and $-w_n(r)$ are obviously degenerate and the localization process can not differentiate between the two. Sign ambiguities become a critical issue when dealing with lead-conductor-lead geometries. The sign of MLWFs in the lead principal layer included in the conductor supercell must be the same as that of MLWFs in the lead supercell, otherwise an incorrect conductance will be obtained. In the case of $s$-like MLWFs, the phase can be reset by imposing for the sign of $\int w_n(r)dr$ (which is the coefficient of $e^{i\mathbf{k}\cdot\mathbf{r}}$ for $\mathbf{g} = 0$) to be positive. More generally, inspection on few more planewave coefficients (or projection to atomic orbitals) is sufficient to check the sign of MLWFs.

Our results are shown in Fig. 4-15, where the band structure, quantum conductance and density of states of a (5,5) CNT are plotted. Most notably, it is seen that starting from the eigenstates at $\Gamma$, where the system has $\sim 2\, eV$ pseudogap, the full metallic band structure with two conduction eigenchannels at the Fermi level is recovered. The conductance also corresponds exactly to the one predicted from the band structure: the number of ballistic channels ($N_{ch}$) is the same as the number of available eigenstates at a given energy, which gives the quantum conductance
Figure 4-15: Band structure, quantum conductance, and density of states of a (5,5) CNT. The folded Brillouin zone clearly shows a sizable pseudo-gap at \( \Gamma \); solid lines are calculated from our method and red dots are from Quantum-ESPRESSO code. The agreement between the two results is excellent. When only the occupied subspace is used (dotted lines), totally incorrect results are obtained.

\[ G(E) = \frac{2e^2}{h} N_{ch}(E). \]

The results prove that the real-space surface Green's function accurately incorporates the response coming from the semi-infinite lead, that can be also confirmed by the density of states, where the van Hove singularities appear at the band extrema and a non-zero density of states is found around the Fermi level (the quantum conductance and density of states obtained from the occupied states (shown in dotted lines) are completely wrong, as is the case for the band structure in Fig. 4-7(a)). It is now clearly shown that the formalism adopted here can efficiently generate the same quality of information that would otherwise require sampling of huge number of k-points. The method is now validated and it will be extensively used to characterize disordered nanotubes in the following chapters.
4.5 Experimental work on quantum transport

As can be inferred from the strong carbon-carbon bond of graphite and diamond, a carbon nanotube is very resilient; in addition it is an ideal quasi-one-dimensional conductor with typical diameters of $1\sim 2 \text{ nm}$ and high aspect ratios of $10^4 \sim 10^5$. Here we review some of the experimental evidence demonstrating that phase coherency can be preserved in sub-micron carbon nanotubes.

A common experimental setup consists of a carbon nanotube lying on a substrate with two metal electrodes deposited on top or at the bottom. Electron transport mechanisms are derived from the current-voltage characteristics as a function of gate voltage, source-drain voltage, temperature, and length of the tubes. Experimental observation of the conductance close to the ideal value, $G = \frac{4e^2}{h} (= 155 \text{ S})$, is very rare [12,13,67] since a transparent contact with electrodes and defectless CNTs need to be realized. However, observation of phase-coherent transport is more common.

Tans et al. [68] measured the current-voltage characteristics of single-walled CNT as a function of temperature. Though the conductance was only a few percent of the ideal value, the temperature dependence of conductance suggests resonant tunneling through a single molecular level, implying that the electron wavefunction is extended over the entire region ($\sim 140 \text{ nm}$) between the two contacts.

To our knowledge, Frank et al. [11] first presented evidence of near-ballistic transport. The measured conductance found was half of the theoretical value. The conductance of multi-walled CNT increases step-wise as the tube is pushed into the liquid metal electrode; the step-wise increase is attributed to the contribution from an increasing number of shells in multi-walled CNT. Exceptionally high current densities ($> 10^7 \text{ A/cm}^2$) were observed, enough to burn out CNTs if heat dissipation would have taken place along the CNT, indicating non-dissipative electronic transport. The result was quite surprising, considering that the experiment was performed at room temperature with 4-$\mu$m-long CNT and hinted that the mean-free-path of inelastic scattering could be quite long.

Liang et al. [67] found oscillations in the differential conductance as a function of
gate and source-drain voltages in CNT devices hundreds of nanometer long. They were able to relate the peak positions of the two-dimensional differential conductance plots to the theoretically-predicted interference pattern originating from the back scattering at the interface of the carbon nanotube and the electrode, suggesting phase coherency over the entire device.

Kong et al. [69] measured conductance as a function of temperature and gate voltage. They obtained low-contact resistance using Pd electrodes. Conductance decreases as temperature is increased due to the enhanced inelastic scattering. Sharp oscillations of conductance as a function of gate voltage found at low temperature were attributed to resonance scattering by defects.

Javey et al. [13] measured the length dependence of conductance in several metallic tubes. They estimated the mean-free-path for acoustic phonon scattering to be about 300 nm. They found that the conductance of 60-nm-long metallic CNT approaches 90% of the theoretical limit, regardless of temperature. At high bias, optical phonons become an important inelastic scattering source [70], and differential conductance decreases. The estimated mean-free-path of optical phonon scattering was about 15 nm. Park et al. [12] also found similar current-voltage characteristics. Length-independent resistance is observed for short length CNTs (< 200 nm) at low bias, while classical linear dependence is found at high bias, which again highlights the loss of phase coherency. Recent theoretical study by Lazzeri et al. [71] on the optical phonon scattering corroborates these experimental findings: the experimental current-voltage characteristics were nicely reproduced by their hot-phonon model, which assumes a non-equilibrium distribution of optical phonons (an effective temperature of several thousands K) and the inelastic scattering by them at high bias.

Though the quality of the nanotubes and the assumptions made to analyze the experimental data will affect the derived characteristic lengths, numerous experimental findings support that a metallic single-walled carbon nanotube is a long quantum wire that retains phase coherency even up to few μm. For short nanotubes below 100 nm, scattering by acoustic phonons is not significant. In that case, the most important scattering mechanism in low bias limit will be the elastic scattering by defects [69,72]
(followed by the inelastic scattering by optical phonons as the bias is increased). In the next chapter, we will discuss how disorder can affect the conductance of CNTs.
Chapter 5

Disordered carbon nanotubes

Introduction

As discussed in Chapter 4, a perfect pristine metallic carbon nanotube will have a conductance of $4e^2/h$. However, carbon nanotubes are bound to have various kinds of disorder, e.g. topological defects, vacancies, substitutional atoms, and sidewall chemical groups that are introduced during the growth or purification process. Differently from two- or three-dimensional systems, disorder in one-dimensional conductor is known to be a strong perturbation since paths are so narrow that an electron is always subject to the influence of disorder. To our knowledge, experimental techniques that can quantitatively measure the number of defects and can directly identify atomic configurations around a defect are not well-established yet. There is no general consensus on the defect concentration in a CNT, that must depend also on the details of processing. In some cases, chemical vapor deposition technique can produce high-quality CNTs with defect concentrations as low as one per $\mu m$, which has been confirmed by a selective electrochemical deposition [73] or has been estimated from the mean-free-path of defect scattering [74]. Analysis of an STM image combined with theoretical calculations could single out the most likely defect structure among all possible candidates [75], but it is not practical to characterize every defect in this way.

Disorder, in general, degrades the excellent physical properties of CNTs, but at
the same time it opens up the opportunity to modulate their physical properties. Therefore, in order to design CNT-based devices, a thorough understanding of the role of defects must be accomplished. In Section 5.1 we briefly summarize theoretical studies on the electronic structure and transport properties of CNTs in the presence of disorder. One of the most significant type of defects will be provided by covalent sidewall functionalizations. Compared to the weak interaction of physisorbed molecules or a substrate, a covalent bonding is robust and stable: heat treatment at temperature as high as 500°C [76] is needed to remove covalent groups. If these chemical groups could be attached in a controlled way, they could serve the purpose of modulating conductance [14]. We summarize currently-available covalent functionalization methods and the physical properties of functionalized tubes in Section 5.2.

5.1 Theoretical work on disordered carbon nanotubes

Experimentally, it is still a difficult task to introduce a desired amount of defects to a carbon nanotube of a specified chiral vector, and that has hindered quantitative analysis of the effect of disorder. Theoretical studies have proven to be a useful tool to compensate for the lack of experimental control. Quantum-conductance calculations in the Landauer approximation have become a popular tool to estimate scattering effect in disordered CNTs. We summarize first several important theoretical studies on phase-coherent transport in disordered CNTs.

Chico et al. [77] have shown that scattering by a single vacancy can reduce the conductance at the Fermi level by 50%. They modeled a vacancy by removing an atom from a tight-binding Hamiltonian. Apart from the issue that whether such simplified tight-binding model can describe the vacancy correctly (they neglect structure relaxation and charge redistribution around the vacancy), their results clearly demonstrate that even a single defect can strongly alter the transport properties in
this quasi-one-dimensional material.

White and Todorov [78] have highlighted an important aspect which makes metallic nanotubes long ballistic conductors [11]. Their two-band model suggests that a CNT feels the effect of disorder averaged over the circumference of the tube, and as a result, the mean-free-path of defect scattering will be proportional to the diameter. For \((n,n)\) armchair tubes, the elastic mean-free-path \(L_m\) is derived as

\[
L_m = \frac{6V_0^2}{(2\sigma_e^2 + 9\sigma_o^2)} n ,
\]

where \(V_0\) is an average off-diagonal element, and \(\sigma_e^2\) and \(\sigma_o^2\) are the variance of diagonal (on-site) and off-diagonal (overlap) elements respectively. For example, a nitrogen doped \((10,10)\) CNT with \(\sigma_e = 0.06\,eV\) \((\Delta\epsilon = 2.5\,eV, \ 0.06\%\ \text{doping})\) will have \(L_m\) about 7.5 \(\mu\)m. Another interesting point is that in the case of metallic CNTs, the number of eigenchannel \((N_{ch})\) is fixed to two, regardless of a diameter. Since localization length is given by \(\xi = N_{ch} \times L_m\), quasi-ballistic motion can be maintained even at the onset of localization.

Anantram and Govindan [79] have compared scattering by strong and weak scattering centers. Their numerical study on weak disorder and localization length is in quite good agreement with the analytic study of White and Todorov. At a fixed disorder strength, the transmission coefficient increases as the diameter is increased. Quantum conductance at the Fermi level is not significantly affected by weak disorder, whereas strong disorder tends to open a transmission gap which is widened as the defect density is increased (the vacancy model studied by Chico et al. certainly belongs to this category).

Choi et al. [80] analyzed scattering by boron or nitrogen atoms using first-principles calculations. While the conductance at the Fermi level is largely unaffected by these substitutional atoms, the quasi-bound states and the conductance dips around the lower (for boron) or the upper (for nitrogen) subband edge reflect the acceptor or the donor character of these atoms. Their study elaborates how two eigenchannels, originating from \(\pi\) - and \(\pi^*\)-bands, are scattered by the defects; a key factor is the
existence of mirror symmetry planes along the tube axis over which these two bands display well-defined parity. Boron, nitrogen or mono-vacancies break the symmetry and result in mixing of the two eigenchannels; both $\pi$- and $\pi^*$-channels are scattered, with a transmission probability of 50%. On the other hand, Stone-Wales defect and divacancies preserve the mirror symmetry and scatter either a $\pi$- or $\pi^*$-channel alone, without mixing the two channels.

Latil et al. [81] employed a different method in evaluating the conductance. Instead of calculating the transmission coefficient explicitly, they obtain the diffusion coefficient as a function of time by evolving wave packets and then derive the conductance from the Kubo relation. Their results are in accord with previous studies [78,79], in the sense that the elastic mean-free-path is inversely proportional to the dopant concentration and proportional to the diameter (Fig. 5-1). The time dependence of the diffusion coefficients reveals the difference between different transport regimes (Fig. 5-2): depending on the energy level chosen, the transport regime can vary from ballistic to diffusive to localized.

Gómez-Navarro et al. [15] measured the conductance of metallic nanotubes irradiated with Ar$^+$ ions. They suggested that divacancies generated by the ion irradiation are responsible for the decreased conductance. The presence of exponentially-increasing resistance as a function of nanotube length strongly supports the assump-
Figure 5-2: Density of states and diffusivity of a boron-doped (10,10) CNT at three different energy levels (figure from Ref. [81]). The asymptotic behavior of diffusivity as a function of time determines the relevant transport regime: quasi-ballistic regime at $E_1$, diffusive regime at $E_F$, and localized regime at $E_2$.

All of the theoretical studies summarized here assume a non-interacting electrons picture. Other interesting phenomena, such as Coulomb blockade, Kondo effect, and Luttinger liquid behavior are not covered in the present work.

5.2 Experimental work on covalent functionalizations

Carbon nanotubes are usually produced in bundles and adhesion between the tubes limits solubility and further characterization and processing. Chemical functionalizations are one of the promising way of overcoming these hurdles. Furthermore, they open up a new avenue to exploit the unique electronic properties of CNTs. Several routes for covalent functionalizations have been developed for diverse applications [82–84]: First, functionalized CNTs dissolve in various solvents [85], at variance from the insoluble, bundled pristine tubes. Second, functionalizations can be used to separate semiconducting tubes from metallic ones [86]. For device applications, it is
crucial to sort carbon nanotubes according to their electronic structure; there is no experimental method at this stage that can selectively grow either metallic or semiconducting tubes. One of the current strategies is to short-circuit and burn metallic carbon nanotubes by flowing large amount of current [87]. Strano et al. proposed a different way to select CNTs by chemical functionalizations, based on their finding that metallic tubes are more reactive to diazonium salts than semiconducting ones [86]: indeed, differences in electronic structure can give rise to different reactivities [88]. The method has been confirmed and utilized by others [89–91]. Third, the exceptionally high surface-to-volume ratio makes CNTs promising candidates for sensors. Chemical groups can add sensing capability and selectivity, e.g. by attaching an acceptor that interacts strongly with a specific target molecule [92].

With regard to the electronic structure, the most notable consequence of covalent functionalizations is that they introduce $sp^3$-type bonding in the otherwise fully delocalized $\pi$-bonding network. This could change the electronic structure to a great extent [93]. As an example, we show the effect on CNTs of 4-bromophenyl functionalizations [76]. The degree of functionalization corresponds to a 4% coverage of sidewall carbon atoms (the ratio is estimated from weight loss in thermogravimetric analysis). Fig. 5-3 shows the absorption spectra of pristine and functionalized tubes. Characteristic absorption peaks that originate in the van Hove singularities are blurred after functionalization. Raman spectra in Fig. 5-4 show an increased peak height at disorder mode around 1290 $cm^{-1}$ after functionalization. All these findings point to the disruption of the perfect $\pi$-bonding manifold and the symmetry-breaking induced by sidewall chemical groups. Similar results have been found in other types of functionalizations, and absorption and Raman spectra have become a standard way to characterize functionalized tubes [86, 94]. Increased electrical resistance is also a signature of functionalization [14, 95, 96]: Fig. 5-5 shows the variation in the resistance of metallic carbon nanotubes as a function of the concentration of functionalizing agents used. Conductance of the tubes drops by two orders of magnitudes for the highest concentrations. This result demonstrates that chemical groups attached to the sidewall degrade the excellent transport properties of CNTs.
Figure 5-3: Absorption spectra of pristine and functionalized CNTs in dimethylformamide, illustrating the loss of characteristic structure following functionalization (figure from Ref. [76]): (a) pristine CNTs; (b) 4-bromophenyl functionalized CNTs. Peaks in the absorption spectra arising from van Hove singularities in the density of state disappear after functionalization.

Figure 5-4: Raman spectra of pristine and functionalized CNTs (figure from Ref. [76]): (a) pristine CNTs; (b) 4-bromophenyl functionalized CNTs. The peak at 1290 cm$^{-1}$ is attributed to disorder. The intensity of the peak is greatly increased after functionalization.

Fig. 5-6 summarizes the currently available functionalization methods. According to their local bonding structure, we categorized them into two groups (Fig. 5-7). The first group includes hydrogens, aryls, fluorines: all groups that form a single
Figure 5-5: Current as a function of gate voltage ($V_G$), for metallic CNTs after exposure to various concentrations of diazonium solutions (figure from Ref. [14]). The current is independent of gate voltage, confirming metallicity of the tubes. The degrading effects of functionalization are evidenced by the monotonic decrease of conductance with increasing concentration of the functionalizing agent.

Figure 5-6: Summary of currently available functionalization methods (figure from Ref. [96]).

bond with a sidewall carbon atom. The second group includes carbenes, nitrenes, azomethine ylides. These attack the sidewalls via a cycloaddition reaction and form a three- or five-membered ring that always involves two carbon atoms on the side-
Figure 5-7: Classification of the functionalization methods of Fig. 5-6, according to the local bonding structure between the functional group and the sidewall carbon atoms: (a) single bond; (b) cycloadditions.

wall. Considering the wide application of covalent functionalizations, it is important to enhance our knowledge on how covalent functionalization modifies the electronic properties of CNTs both qualitatively and quantitatively. The energetics of functionalizations, and the electronic structure and the conductance of functionalized CNTs will thus be discussed in Chapter 6 and 7 using the theoretical tools introduced in Chapters 2 and 3.
Chapter 6

Functionalized carbon nanotubes:
Hydrogens and aryl groups

Introduction

Our first class of functionalized carbon nanotubes is characterized. Hydrogens [97] and aryl groups [76] are chosen as representatives of addends that form a single bond with sidewall carbon atoms (Fig. 5-7(a)).

In the phase-coherent regime, conductance fluctuates strongly whenever scattering centers (here, functional groups) change their relative position, owing to the interference of the electron wavefunctions. For this reason, finding out the energetically or thermodynamically stable distribution of functional groups is a first step toward full characterization of covalent functionalizations. Previous studies on fluorination highlighted the importance of ordering on the sidewall. Inspired by the experimental findings that a fluorinated tube has C\textsubscript{2}F stoichiometry [98] and that fluorine atoms preferentially form a ring along the circumference of the tubes [99], stable conformations of fluorinated CNTs have been studied. First-principles calculations have shown that the size of the band gap ranges from 0 to a few eV depending on the patterns of fluorination [100,101]. Here, we will cover only a limited number of configurations: an exhaustive search for the global minimum would require statistical mechanics techniques [102-105] outside the scope of this work.
The most prominent consequence of covalent functionalizations is saturation of the conjugated π-bonding network of the sidewall carbon atoms. Hydrogen is the simplest group that brings this characteristic feature. It is of special interest to investigate how metallicity is affected by functionalization since in the extreme case of a fully hydrogenated tube, a sizable gap is opened [106,107]. The electronic structure and the transport properties of functionalized armchair metallic carbon nanotubes is discussed in Sections 6.2 and 6.3. Aryl groups are more versatile since charge transfer can be controlled by replacing the hydrogens in a phenyl ring with different electropositive or electronegative chemical groups. The results from various aryl groups will be compared to assess the relevance of charge-transfer effects.

Simulation details

A typical simulation cell for a (5,5) CNT will contain 100 atoms plus the functional groups. Γ-point sampling is used for structural optimization. Atomic coordinates are relaxed until the force in each crystallographic direction is smaller than 0.02 eV/Å. A 1x1x5 k-point mesh and a cold smearing of 0.02 Ry [54] is used for single-point energy calculations after optimizing the atomic coordinates.

A functionalized tube may require a larger amount of vacuum to avoid interactions with periodic images. Fig. 6-1 shows energy as a function of a cell dimension a for a pristine and a functionalized tube, and the difference in the last column. The difference converges faster than the absolute energies, and it converges within ~5 meV for a vacuum length of $V_{ac} = 12\,\text{Å}$. This vacuum length will be consistently used in Chapters 6 and 7.

6.1 Energetics and structure

In the case of hydrogenation, chemisorption of a single hydrogen or a hydrogen pair have been studied most extensively. Li et al., [108] studied the energetics of a hydrogen pair on fragments of armchair tubes (from (3,3) to (6,6)), employing a semi-empirical quantum chemical method. They found that a hydrogen dimer fa-
Figure 6-1: Energy of a pristine and a functionalized CNT as a function of the cell dimension $a$ in the $x$ and $y$ directions. (a) 100 atoms of a pristine (5,5) CNT. (b) 100 atoms of a pristine (5,5) CNT plus two phenyl groups. (c) Energy difference between (a) and (b): it converges within $\sim 5$ meV when the vacuum length is 12 Å (i.e. the cell dimension $a$ is 19 Å).

...vors nearest-neighbor sites and generally it is more stable when an odd number of C-C bond connects the two hydrogen atoms (as in the ortho and para position of a benzene ring). A pair of hydrogen atoms is also found to be more stable than an atomic hydrogen [109, 110]. The tendency for pairing has also been observed in graphite and fullerene, close relatives of carbon nanotubes. A recent STM study shown in Fig. 6-2 has clearly demonstrated that the presence of hydrogen dimers on graphite. Even number of addends are found covalently attached to fullerenes, and usually two groups are located nearby [111, 112]. Since the bonding structure of a carbon nanotube can be described as a resonant structure of alternating single and double bond, attaching a functional group to a sidewall carbon atom activates nearby carbon atoms by introducing a dangling bond, which can expedite subsequent reactions between neighboring carbon atoms and functional groups, to minimize the disruption of the conjugated $\pi$-bonding network.

Reaction energies upon addition of hydrogen (H), nitrophenyl ($C_6H_4NO_2$), amino-phenyl ($C_6H_4NH_2$), and phenyl ($C_6H_5$) are calculated. A stable molecular form, where a hydrogen atom saturates the dangling bond of a ligand, is taken as a reference state: for example, a benzene is the reference state for a phenyl. Reaction energies are thus
defined as

\[
\begin{align*}
\text{CNT} + (\text{FG}_1)\text{H} &\rightarrow \text{CNT}(\text{FG}_1) + \frac{1}{2}\text{H}_2 \quad \Delta E_1, \\
\text{CNT}(\text{FG}_1) + (\text{FG}_2)\text{H} &\rightarrow \text{CNT}(\text{FG}_1)(\text{FG}_2) + \frac{1}{2}\text{H}_2 \quad \Delta E_2,
\end{align*}
\]

where FG\(_n\) stands for \(n\)th functional group attached. Our results are summarized in Fig. 6-3. Attaching the first functional group is endothermic in all cases. The absolute value of \(\Delta E_n\) is not very meaningful since it would be exothermic if atomic hydrogen were considered as a reference state [115,116] and in any case the chemical potential (or partial pressure) should also be considered. What is more relevant is the difference between \(\Delta E_1\) and \(\Delta E_2\). \(\Delta E_2\) on a few neighboring sites, (a), (β), (γ), and (δ) in Fig. 6-3, is significantly lower than \(\Delta E_1\), which indicates a strong tendency for pairing.\(^1\) When the second hydrogen is located far from the first one, \(\Delta E_2\) is found to be much larger than -0.34 eV (the lowest value in Fig. 6-3(a)). Fig. 6-4 shows the unfolded supercell; the second hydrogen added is marked as a red dot (two different positions have been explored). For further validation, we tested two configurations for the addition of a third hydrogen. \(\Delta E_3\) for the third hydrogen addition (defined

\(^1\)Instead of adding one more group, changing the charge state of the nanotube could stabilize an isolated functional group; we have not considered this option in detail.
Figure 6-3: $\Delta E_2$ in eV as a function of the position of the second group FG$_2$ on a (5,5) CNT (figure from Ref. [114]). FG$_1$ is attached in the position marked by X ($\Delta E_1$ is indicated in parentheses) FG$_1$ and FG$_2$ are, respectively, (a) H and H, (b) nitrophenyl and H, (c) aminophenyl and H, (d) phenyl and phenyl (the nitro and amino residues are in the para position). The arrow lies parallel to the nanotube axis, and we denote with $\alpha$, $\beta$, $\gamma$, and $\delta$ the arrangements for four of the most stable pairings.

Figure 6-4: $\Delta E_2$ in eV when the second hydrogen atom is added on two of the most distant sites from the first hydrogen on a (5,5) CNT. The unfolded supercell of a (5,5) CNT is shown for clarity. X indicates the position of the first hydrogen, and a red dot indicates the second one.

in the same manner as for $\Delta E_2$ in Eq. 6.1) turns out to be higher than $\Delta E_2$ (Table 6.1).

It should be mentioned that the size of the simulation cell can significantly affect the result. Arellano et al. [117] and Lu et al. [116] suggested that hydrogen pairs attached to a narrow diameter CNT can induce unzipping of the tube. They used a periodic cell that actually makes a continuous row of hydrogen pairs, which brings
Figure 6-5: Fragment of a (5,5) CNT saturated with hydrogens: C_{110}H_{20}. This fragment is used to compare the reaction energies of hydrogenation with those obtained in a periodic cell. Carbon atoms are labelled to mark the positions of hydrogen atoms.

Table 6.1: Comparison of the reaction energies $\Delta E_n$ of hydrogenation in a periodic cell or in a C_{110}H_{20} CNT fragment. The indices in the second column indicate the positions of hydrogens (labelled as in Fig. 6-5). $\Delta E_n$ is the reaction energy of adding the $n_{th}$ hydrogen when there are $(n-1)$ hydrogens at the positions listed in parentheses. Note that the finite length of the fragment results in different reaction energies while they should be the same on an infinite tube (e.g. 1-6 and 2-3' pairs).

<table>
<thead>
<tr>
<th>$\Delta E_n$ (eV)</th>
<th>periodic</th>
<th>fragment</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta E_1$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.63</td>
<td>0.65</td>
</tr>
<tr>
<td>2</td>
<td>0.77</td>
<td></td>
</tr>
<tr>
<td>$\Delta E_2$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1), 2</td>
<td>-0.34</td>
<td>-0.43</td>
</tr>
<tr>
<td>(1), 6</td>
<td>-0.09</td>
<td>-0.19</td>
</tr>
<tr>
<td>(2), 3'</td>
<td></td>
<td>-0.03</td>
</tr>
<tr>
<td>$\Delta E_3$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1, 2), 6</td>
<td>0.42</td>
<td>0.46</td>
</tr>
<tr>
<td>(1, 2), 4</td>
<td>0.76</td>
<td>0.71</td>
</tr>
</tbody>
</table>

huge distortions and consequent carbon-carbon bond breaking. When a hydrogen pair is attached to a twice-larger cell, the carbon nanotube remains stable. To see whether the simulation cell size is large enough to suppress the interaction between periodic images, we also calculated reaction energies on a fragment of a (5,5) CNT. The geometry of the fragment (C_{110}H_{20}) is shown in Fig. 6-5. The main result found is that the periodic cell simulations are well reproduced in the fragment calculations (Table 6.1): paring is favored and (a) is more stable than (b). The good agreement
between the two confirms that the effect of periodic images is not significant in our case.

We thus draw the conclusion that a pair of functional groups should rather be regarded as a unit of functionalization. The most favored pair position is (α) when a pair of hydrogens is added; bulkier groups favor a different position, such as (γ) (Fig. 6-3(d)).

6.2 Band structure

The band structure of functionalized (5,5) carbon nanotubes was generated by the interpolation method explained in Section 2.2.2 using Hamiltonian matrices in the MLWFs basis constructed at the Γ-point. Fig. 6-6 shows the band structure of a (5,5) CNT functionalized with two hydrogens, a nitrophenyl and a hydrogen, or an aminophenyl and a hydrogen, sitting in the (α) configuration of Fig. 6-3. For comparison, we also plot the band structure of a pristine (5,5) CNT. In Fig. 6-6(a), the band structure as obtained from standard methods with a full planewave basis is shown as solid dots; agreement between the two results is again excellent as it was the case for a (5,5) pristine tube. Surprisingly, these band structures are largely insensitive to the chemical nature of the attached groups. The similarity hints that a common factor determines the topology of the modified band structures. We compare in Fig. 6-7 MLWFs before and after functionalizations. MLWFs naturally reflect chemical bonding, and provide insights into how the electronic structure is affected upon modifications of chemical bonds [65, 66]. The formation of a covalent bond with a functional group transforms the original half-filled $p_z$-MLWF into a fully occupied $sp^3$-bonding MLWF whose on-site energy is lowered by $\sim 7\, eV$. The huge difference in the on-site energy effectively removes the $sp^3$-bonding MLWF from the $\pi$-manifold which is responsible for the electronic transport in CNTs.

\footnote{We note that whenever unoccupied states are included, some arbitrariness ensues from choosing the inner and outer window (Fig. 2-1). One must apply a consistent rule to obtain comparable results to the pristine tube: simply, MLWFs of the pristine tube should be recovered away from a functional group.}
Figure 6-6: Band structure of a (5,5) CNT decorated with a periodic array of functionalization pairs in position $\alpha$ (figure from Ref. [114]): (a) H pair (dots are from a quantum-ESPRESSO calculation, for comparison); (b) nitrophenyl and H; (c) aminophenyl and H; (d) model calculation, where the two $p_z$-MLWFs on the functionalized sidewall carbons have been removed from the calculation; (e) pristine nanotube.

Figure 6-7: Comparison of MLWFs before and after functionalizations. (a) The three inequivalent MLWFs obtained for a pristine (5,5) CNT, clearly corresponding to a chemical picture of $p_z$ and $sp^2$-bonding orbitals. (b) The same MLWFs for the case when an aminophenyl group (1% coverage) has been covalently attached. The $p_z$-MLWF is replaced by an $sp^3$ bonding MLWF.

To confirm this assumption, we constructed a model Hamiltonian matrix whose $p_z$-MLWFs corresponding to the functionalization sites are removed. This can be done in practice by setting all the matrix elements involving the relevant $\omega_p$ ($p_z$-
Figure 6-8: Comparison of charge transfer from four different aryl groups, $R = \text{NH}_2$, H, COOH, and NO$_2$. (a) Averaged charge density of the functionalized tubes. X indicates the position of the aryl group. (b) $\Delta \rho$ of the two extreme cases, aminophenyl($-\text{C}_6\text{H}_4\text{NH}_2$) and nitrophenyl($-\text{C}_6\text{H}_4\text{NO}_2$)$\times 10^{-5}\text{e/au}^3 \approx 0.3 \text{ |e|/supercell}$.

MLWF) to zero: $\langle \omega_i | \hat{H} | \omega_p \rangle = \langle \omega_p | \hat{H} | \omega_i \rangle = 0$, or equivalently by removing the corresponding columns and rows from the Hamiltonian matrix of a pristine (5,5) CNT. This approximation has been used in tight-binding models to simulate a vacancy or covalently-attached groups, without a formal justification [77, 118]. The band structure obtained from this model Hamiltonian is shown in Fig.6-6(d); it is closely similar to those obtained in the functionalized cases, confirming that the foremost effect of functionalization is a rehybridization from $sp^2$ to $sp^3$.

Now we turn our attention to the effects of charge transfer. Charge transfer by four phenyl moieties, 4-R-phenyl (‘4’ denotes that the residue $R$ is in para position) with $R = \text{NH}_2$, H, COOH, and NO$_2$ are compared in Fig. 6-8(a). The propensity of charge transfer to the nanotube follows the usual sequence, $\text{NH}_2 > \text{H} > \text{COOH} > \text{NO}_2$. When a nitro group is replaced by an amino group, roughly $0.1 \text{ |e|/supercell}$ is transferred to the nanotube backbone: the amount of charge transfer is estimated from
Figure 6-9: Band structure of a (5,5) CNT with two \( p_z \)-MLWFs removed. These are located at the four different paired positions of Fig. 6-3: (a) \( \alpha \); (b) \( \beta \); (c) \( \gamma \); (d) \( \delta \); (e) pristine CNT as a reference.

the flat region in Fig. 6-8(b) where the difference in charge densities is plotted. The actual value is expected to be higher since the difference must be greater around the functional group marked as X. As can be seen from the graph, charge density is not fully converged to that of a pristine tube, which implies that the charge distribution can vary with simulation-cell size and the number given here should be taken as an estimate.\(^3\)

Since removing a \( p_z \)-MLWF captures the effect of adding a covalent bond, we calculate band structures for other pairings directly from the Hamiltonians for pristine CNTs. The results are shown in Fig. 6-9. Different patterns have a stronger effect than chemical identity, which implies that the distinction among different functional species would mainly come from the preference to a certain topology for the arrangements on the sidewall of the tube.

\(^3\)It has been discussed that extra charges in metallic carbon nanotubes and graphite could be fairly delocalized [119,120].
6.3 Quantum conductance

A band structures hints at the electronic-structure modifications that follow functionalization. However, one should not be tempted to say that a (5,5) CNT having a functional group pair in (δ) position will be perfectly metallic, judging from the band structure in Fig. 6-9(d). This would be true only if periodic structure were assumed, and that makes the size of the band gap a function of the simulation cell size or defect density [121]. Perfect periodicity in the functionalizations will almost never be the case in real disordered materials. A more realistic approach would then be to evaluate the scattering by a single functional group on an infinite tube and to see how scattering is enhanced when more and more groups are added.

To this purpose, we calculated the quantum conductance of functionalized tubes. Fig. 6-10 shows the quantum conductance profile of an infinite (5,5) CNT with a pair of hydrogens in the configurations (α) and (β). It turns out that scattering by a single pair is not strong enough to open a gap at the Fermi energy (this could not be evinced from the band structures in Fig. 6-9). The energy levels most affected are 1 eV above and below the Fermi level, where conductance drops by 50%. These conductance dips are related to the flat bands (i.e. localized states) found at these energy levels: perturbation caused by functionalization is most prominent, and as a consequence electrons cannot propagate well and instead are localized around the defect sites, which also can be seen from the enhanced projected density of state shown in Fig. 7-13. On the contrary, a relatively broad band dispersion is found around the Fermi level even if a small gap appears: we expect that the size of this gap will decrease as the supercell size is increased. Scattering by a pair of phenyls at (α) follows a similar pattern to the hydrogen case. We again confirm that the approximation of removing p-MLWFs (solid blue) produces qualitatively the same result as the full first-principles simulations with hydrogens and phenyls, and this model Hamiltonian will thus be used as a representative Hamiltonian of covalently-functionalized CNTs.

The next question is how the conductance would change when more and more groups are added. The central conductor will now consist of thousands of carbon
Figure 6-10: Quantum conductance for an infinite (5,5) CNT with one functional group pair attached in the center. (a) α. (b) β. Quantum conductance calculated from the model Hamiltonian (solid blue) is in close agreement with the full first-principles calculations for the hydrogen and phenyl pair ((b) panel does not have the phenyl pair case; in both panels the dashed line shows the pristine CNT case).

Figure 6-11: Infinite (5,5) CNT functionalized in its central region by an array of phenyl pairs.

atoms with tens of functional groups, as shown in Fig. 6-11. We start from the case of 1000 carbon atoms with 10 randomly arranged pairs in position (α) in an otherwise perfect, infinite tube. The number of possible choices is on the order of $1000C_{10}$ (this number will become smaller when symmetry and overlap between pairs are taken into account). As discussed before, in the phase-coherent regime, conductance fluctuates strongly among different configurations; average over a large number of configurations is required to obtain a quantitatively accurate result. Averages over 5 and 10 sets of random configurations are compared in Fig. 6-12. The averaged
quantum conductance for two independent 5 sets (red and blue in the panel (a)) is qualitatively the same. Increasing the number of configurations to 10 suppresses fluctuations, but averaging over 5 sets seems enough to capture the basic results.\textsuperscript{4}

The averaging scheme is also an issue: when conductance is smaller than the conductance quantum, a geometric average gives a more accurate estimate than an arithmetic average since $\ln(g)$, and not $g$ itself, is normally distributed \cite{122}. In the panel (b) of Fig. 6-12, the arithmetic average $\langle g \rangle$ and the geometric average $\exp(\ln(g))$ are plotted. The agreement is perfect when conductance is greater than the conduc-

\textsuperscript{4}It is interesting to see the magnitude of standard deviation. The size of the error bars in Fig. 6-12 is not strongly dependent on the conductance and is around the conductance quantum ($\sim \frac{2e^2}{h}$), which is discussed in Section 3.1.3.

Figure 6-12: Quantum conductance for an infinite (5,5) CNT with ten pairs of $p_z$-MLWFs removed from the $\pi$-manifold. Averages are over 5 or 10 sets, and arithmetic and geometric averages are compared. The positions of $p_z$-MLWFs (in the $\alpha$ pairing) are randomly generated. (a) Comparison between averages over two independent five random configurations. (b) Comparison between the arithmetic and geometric averages over ten random configurations.
Figure 6-13: Quantum conductance for an infinite (5,5) tube with one, ten, and thirty randomly distributed single ligand or pairs of ligands in four different configurations: (a) α; (b) β; (c) γ; (d) δ; (e) single ligand.

The geometric average tends to be smaller than the arithmetic one when \( g < 1 \), which is expected in the localized regime [123], but the deviations do not introduce major differences. For simplicity, an arithmetic average over 5 random configurations is used for the following data.

We then calculate the quantum conductance of an infinite (5,5) CNT with one, ten, and thirty functional group pairs, or single functional groups. The results are summarized in Fig. 6-13: the pairs are arranged in positions (α), (β), (γ) or (δ), corresponding to each panel in sequence, and the bottom panel is for the isolated group case. The covalently-bonded groups act as strong scattering centers, reduc-
ing the conductance at the Fermi energy by 19%(\(\alpha\)), 8%(\(\beta\)), 21%(\(\gamma\)), 17%(\(\delta\)) and 42%(single) respectively. All four paired cases turn out to be less destructive than an isolated functional group, despite the fact that the number of groups is two times larger. The pairing of functional groups not only stabilizes the structure but also weakens the scattering around the Fermi level, which confirms the idea that one can reduce the disruption of the conjugated network by pairing.

The cumulative effect of tens of functional groups is very strong. Once a random array of functional groups covers the nanotube for tens of nanometers - comparable to the lengths of pristine tubes for which quasi-ballistic transport has been observed [13,124] - the conductance of the tube drops dramatically; such dramatic decrease in the conductance has also been recently confirmed experimentally [14,125]. Among the patterns studied, (\(\beta\)) turns out to be the weakest scattering center, since it is a paired defect that preserves the mirror plane containing the tube axis and is almost transparent to either of the two Bloch states with well-defined parity over the mirror plane (a similar result has been discussed for symmetric Stone-Wales defects on a (10,10) CNT [37]). It is worth mentioning that the position of conductance dips can be dependent on functional groups [126], though this effect blurs when tens of groups are attached.

Summarizing our results so far for hydrogen and aryl group functionalizations, we emphasize that the foremost effect of covalent functionalizations is an sp\(^2\) to sp\(^3\) rehybridization, which effectively removes \(p_z\)-orbitals from the \(\pi\)-manifold. These rehybridizations create strong scattering centers that significantly affect conductance. In addition, these groups have a strong tendency to cluster together.

As a final remark, we do not intend to convey the idea that chemistry is not relevant. Besides determining the topology and patterns of attachment, the charge transfer will modulate the work function and has central importance when a contact with other materials is considered. In the next chapter, we will see how chemistry can play a significant role in a very unusual way.
Chapter 7

Functionalized carbon nanotubes: Carbenes and nitrenes

Introduction

The second category of covalent functionalizations we consider is that of cycloadditions (Fig. 5-7(b)), where a functional group saturates a double bond between two sidewall carbon atoms. Two common types of the cycloaddition are the [2+1] cycloaddition of carbenes or nitrenes [94,97,127-131] and 1,3-dipolar addition of azomethine ylides [85,118,132].

Addition of carbenes (-CR₂) or nitrenes (-NR) to the sidewall of a CNT creates an interesting bonding structure. The carbon or nitrogen “bridgehead” atom in these groups forms a cyclopropane-like three-membered ring structure with two neighboring carbon atoms on the sidewall. At variance from the stable five-membered ring created by 1,3-dipolar addition, a three-membered ring is under huge strain energy (cyclopentane (ΔᵢH° = -76.4 kJ/mol) is in fact much more stable than cyclopropane (ΔᵢH° = 53.3 kJ/mol) [133]). It is interesting to examine these functionalizations in fullerenes, first.¹ Cycloaddition of CH₂ on C₆₀ yields two isomeric forms depending on whether a carbene saturates a [5,6] or a [6,6] ring fusion edge, as shown in Fig. 7-1.

¹Functionalizations of fullerenes have a long history. Most of the functionalization approaches applied to carbon nanotubes have already been realized in fullerenes.
Figure 7-1: Two isomeric forms of methylene (CH₂) functionalized fullerene (figure from Ref. [134]): (a) methanofullerene; (b) fulleroid.

The [6,6] ring edge remains intact whereas a carbene group on a [5,6] edge cleaves the bond between two carbon atoms; the optical properties of two isomers are very distinct from each other [134]. Theoretical studies have predicted similar bond cleaving on the sidewall of a CNT [121, 135–142]. However, the scope of previous studies has been limited to structure optimizations and energetics in most cases. The effect of bond cleaving on the electronic transport properties has not been appreciated enough, nor the differences among various carbene groups.

In the present work, we have performed extensive studies on this class of cycloadditions. In Section 7.1, energetics and equilibrium structures are discussed, giving special attention on the diameter dependence of these properties. Electronic structure and quantum conductance are discussed in Section 7.2. The crucial role of chemistry in controlling electron transport is highlighted in Section 7.3, leading to the discussion of a novel conductance-switching mechanism in Section 7.4.

### 7.1 Energetics and structure

The curved sidewall of a CNT provides few independent sites for a cycloaddition, differentiated by the relative angle of the C-C bond with respect to the tube axis: two sites for achiral tubes and three for chiral tubes. The focus here remains on metallic armchair CNTs. Fig. 7-2 (a) and (b) show the two options available on
Figure 7-2: Three different configurations for a functional group CH$_2$ on a (5,5) CNT (figure from Ref. [143]): (a) skewed S (b) orthogonal O, with an intact (“closed”) sidewall bond (c) orthogonal O, with a broken (“open”) sidewall bond.

an armchair carbon nanotube, which are equivalent to the (α) and (β) positions of a pair of hydrogen atoms in Fig. 6-3. They are labelled as “S” (skewed) and “O” (orthogonal), to emphasize the angle of the sidewall carbon bond with respect to the tube axis.

Our simulation cells include 12$n$ carbon atoms in the ($n,n$) CNT plus one functional group; this is three times longer than the conventional unit cell of an armchair CNT. A 1×1×4 regular mesh of k-points is used for all structural relaxations, together with 0.03 $Ry$ of cold smearing. The k-points mesh is doubled (1×1×8) for single-point ground-state energy calculations performed with the optimized atomic coordinates.

The ground state energies for the S and O configurations are calculated first for the simplest addends in the class: CH$_2$ and NH. Reaction energies are defined as

$$\Delta E_{\text{CNT}} = E_{\text{CNT-func}} - E_{\text{CNT}} - E_{\text{func}} \quad (\text{func=CH}_2 \text{ or NH}) .$$  

(7.1)

In Fig. 7-3(a), we plot these reaction energies as a function of 1/r (this curvature is calculated from the atomic coordinates of pristine tubes and the deformation ensuing from functionalization is not taken into account). The reaction energies are found to
Figure 7-3: $\Delta E$ for CH$_2$ and NH cycloaddition on an $(n,n)$ CNT and formation energy of a pristine $(n,n)$ CNT with respect to graphene. (a) Energy change $\Delta E$ upon functionalizations; the orthogonal O configuration is always more stable than the skewed S one, irrespective of the chemical nature of the addends. (b) Extra energy per atom with respect to a flat graphene sheet; the energy is proportional to $1/r^2$.

be linearly dependent on curvature, following the formula

$$\Delta E_{\text{CNT}} = \Delta E_{\text{graphene}} - \frac{C}{r},$$

where $C$ is a constant which is dependent on the functional group and the attachment site. The same linear dependence on curvature has been reported for H, F, OH, and COOH functionalizations [126, 144]. This behavior clearly demonstrates higher reactivity of small diameter tubes. The extra energy per unit length of an $(n,n)$ CNT with respect to flat graphene is proportional to $1/r$ (Fig. 7-3(b)), which accounts for the enhanced reactivity of small diameter tubes. For both CH$_2$ and NH, the O configuration is more stable than the S configuration for all the diameters considered. The energy difference is $0.5\sim1\,\text{eV}$ for the range of diameters experimentally synthesized.
Figure 7-4: Sidewall equilibrium bond distance $d_{16}$ for $(n,n)$ CNTs in the configurations S and O. The sidewall bond on the O site is broken in all the $(n,n)$ CNTs considered, while it remains intact on the S site.

(1~2 nm); at room temperature armchair CNTs will strongly favor the O site.

In Fig. 7-4, the C$_1$-C$_6$ bond distance (i.e. $d_{16}$ from Fig. 7-2(c)) is plotted. This distance, for the O configuration, is much larger than the usual C-C distance (1.54 Å in diamond and 1.42 Å in graphite). This clearly indicates that that the sidewall bond is broken, as shown in Fig. 7-2(c). The stabilization of O is partly due to bond breaking and the accompanying strain-energy release. To measure the contribution from bond breaking, reaction energies are calculated under a constraint in which the carbon nanotube backbone is not relaxed. The results are then compared with the fully relaxed case in Fig. 7-5: the difference between the relaxed and the unrelaxed case is attributed to the extra stabilization energy due to the bond opening.

Bond cleaving is strongly affected by the curvature and the angle with respect to the tube axis: e.g., the bond remains intact on a graphene sheet, and the S configuration never breaks the bond. This suggests that a transition from closed
Figure 7-5: Difference in $\Delta E$ between the S and O configuration for relaxed and unrelaxed $(n,n)$ CNTs functionalized with CH$_2$.

to open form should occur when the curvature is increased. As shown in Fig. 7-4, a transition point is not found up to a (12,12) CNT for both CH$_2$ and NH. The cost of the simulations, however, becomes prohibitively expensive upon increasing the diameter, and so we decided to estimate the critical diameter from a curved graphene sheet. The structural parameters of this curved graphene sheet are described in Fig. 7-6. The arc length $l$ is half of the circumference of a (6,6) CNT, and the dimension in the $z$ direction is the same as that of a CNT. The orthorhombic unit cell contains 72 atoms plus two functional groups; the atoms are symmetrically positioned so that the unit cell belongs to the $mmm$ point group. The coordinates are generated varying
Figure 7-6: Schematic description of a bent graphene sheet. The arc length $l$ is fixed to half of the circumference of a (6,6) CNT and $\theta$ is varied, which determines in turn the other structural parameters.

$\theta$, and all parameters are calculated according to the following formulas:

$$
\begin{align*}
  l &= 9d_{CC} , \\
  r &= \frac{l}{\theta} , \\
  s &= 2r \sin \frac{\theta}{2} , \\
  a &= -2r \cos \frac{\theta}{2} + \sqrt{(2r \cos \frac{\theta}{2})^2 + V_{ac}^2 + 4V_{ac}r} , \\
  b &= 2s , \\
  c &= 3\sqrt{3}d_{CC} ,
\end{align*}
$$

(7.3)

where $a$, $b$, $c$ are cell dimensions in the $x$, $y$, $z$ direction, respectively. The cell dimension $a$ is adjusted such that the vacuum length $V_{ac}$ remains constant when $\theta$ is varied. For geometry optimizations, $1 \times 2 \times 4$ k-points mesh is used.

The optimized structure of a curved graphene sheet with $\theta = 180^\circ$ and a (6,6) CNT are compared in Fig. 7-7(a). While one should not expect an exact one-to-one matching between the two since the electronic structures are not exactly identical, and the arc length is chosen somewhat arbitrarily (restricting the shape of bent graphene), we still observe that the results for the relaxed $d_{16}$ of the CH$_2$ functionalized CNTs are closely reproduced by our bent graphene model (Fig. 7-7(b)). The initial configuration is for ideal curved sheet with the bond closed. As the curvature is increased, a sharp transition from closed to open takes place for CH$_2$ around a diameter of 2.4 nm, or equivalently an (18,18) tube.
Figure 7-7: Relaxed structure and $d_{16}$ of a (6,6) CNT and a bent graphene sheet. (a) The cross-section of a CH$_2$ functionalized (6,6) CNT is compared with that of a bent graphene sheet with $\theta = 180^\circ$. (b) $d_{16}$ of CNTs and bent graphene sheets. Bent graphene sheet allows to sample regularly-spaced curvature starting from zero. A transition from closed to open is located; the critical curvature corresponds approximately to that of an (18,18) CNT.

Zigzag and chiral tubes

To explore the phenomenology of bond cleaving, simulations on zigzag and chiral tubes have been carried out. In the literature, bond cleavage is predicted for zigzag and chiral tubes [115, 136, 138, 140, 141]. Exactly like armchair tubes, $(n,0)$ zigzag tubes provide two independent attachment sites: one is parallel to the tube axis (labelled as P) and the other is skewed (labelled as S). Zigzag tubes ranging from $n=5$ to 14 have been tested. Simulation cells include $8n$ carbon atoms for an $(n,0)$ CNT, plus one functional group. The dimension in the $z$ direction is $c = 2 \times (3d_{CC})$, comparable to the $c = 3 \times (\sqrt{3}d_{CC})$ of an armchair tube. Identical k-point meshes, smearing, and force-convergence criteria are used.

The results are presented in Fig. 7-8: as found in armchair tubes, we see that the bond cleaving occurs when the two carbon atoms make a relatively large angle with the tube axis. Again, the bond cleaved S configuration is more stable than the closed
Figure 7-8: $\Delta E$ and $d_{16}$ for CH$_2$ functionalized ($n$,0) CNTs in the configurations S and P. (a) $\Delta E$ defined in Eq. 7.1. $\Delta E$ for the P configuration is not a monotonic function of curvature as in Eq. 7.2. (b) Sidewall equilibrium bond distance $d_{16}$.

P configurations. We note that the S configuration of a zigzag tube makes a larger angle with the tube axis than that of an armchair tube: the ideal values are 60° and 30°, respectively.

The reaction energies for the P configuration show a more complex behavior. Depending on their $q$, defined as $q \equiv \text{mod}(n - m, 3)$, zigzag tubes can be divided into three groups [145]. The results in Fig. 7-8(a) show that the metallic zigzag tubes with $q = 0$ are the most reactive among the three. The pattern in the reaction energies as a function of diameter is consistent with other works [126,140,146].

When bond breaking occurs, the reaction energies are largely governed by ionic relaxations (or equivalently by curvature). On the other hand, the subtle differences in electronic structure becomes more pronounced when the curvature effect is suppressed. This behavior is not expected in the case of armchair tubes since they all belong to the same group with $q = 0$.

\footnote{A smooth dependence on curvature was found by Gülseren et al. [147]. This is not consistent with the present work, and the different characteristics might be due to the smaller unit cell used [139,148].}
Finally, (2n,n) chiral tubes, n=1 to 3, have been tested. Simulation cells include 28n carbon atoms for a (2n,n) CNT, plus one functional group. The dimension in the z direction is \( c = 3\sqrt{7}d_{CC} \), and 1\times1\times4 k-points mesh is used. The three independent attachment sites are labelled as \( \theta_i \). The ideal values of \( \theta_i \) are 79.1°, 40.9°, 19.1° for i=1, 2, 3 respectively. Bond cleaving is observed only for the largest \( \theta_i \) (i=1). Interestingly, the reaction energies do not change monotonically as a function of \( \theta_i \), and the sequence of \( \theta_2 \) and \( \theta_3 \) is switched.

Simulations on zigzag and chiral tubes show some subtle effects that are not present in armchair tubes; here we limit ourselves to armchair tubes, leaving the rest as future work.

### 7.2 Band structure and quantum conductance

Our main focus is to investigate how bond cleaving affects electron transport in functionalized tubes. We then switch to realistic functional groups that have already been
Figure 7-10: MLWFs of a (5,5) CNT functionalized with dichlorocarbene. In striking contrast with the case of \( sp^3 \) rehybridization (Fig. 6-7), disentanglement and localization generates \( p_z \)-MLWFs on \( C_1 \) and \( C_6 \) when bond cleaving takes place: a clear signature that the original \( sp^2 \) hybridization is recovered.

synthesized experimentally, such as dichlorocarbene (labelled henceforth as DCC) [97] and methoxycarbonylnitrene (MCN) [127]. In order to compare with the results obtained for the hydrogen and aryl functionalizations, a unit cell of the same dimensions (100 carbon atoms plus one functional group) is used. A single DCC or MCN group is added to a (5,5) CNT on the S or O site. As is the case of \( \text{CH}_2 \) and NH, structural relaxations lead to a bond-opened configuration for attachments on the O site.

In the last section, strain release by bond breaking was highlighted. Here we would like to focus on the orbital rehybridization that follows bond breaking. We have seen that the major effect of a single-bond forming group is \( sp^3 \) rehybridization and we were able to reproduce the band structure of functionalized tubes by simply removing \( p_z \)-MLWFs from the \( \pi \)-manifold. Bond opening drastically changes this situation: \( C_1 \) and \( C_6 \), after cleavage, go back to an environment in which they have only three nearest neighbors and recover their \( sp^2 \)-bonding character. We show in Fig. 7-10 the relevant \( p_z \)-MLWFs on \( C_1 \) and \( C_6 \); these MLWFs would disappear if the bond remained intact. The difference in the on-site energy for these \( p_z \)-MLWFs with respect to the pristine case is only \( \simeq 0.5 \text{ eV} \); this degree of perturbation is even weaker than that of nitrogen or boron substitution [78,81] which also preserve the \( sp^2 \)-bonding character.
Fig. 7-11 shows the band structure of DCC and MCN functionalized tubes. Again, the band structure is more influenced by the functionalization pattern than the chemical details of the functional group itself. As one might expect, the band structures for S site functionalization closely resemble those of pattern (α) in Fig. 6-9. The foremost effect of the bond opening is apparent: the band structure of the O open configuration is strikingly similar to that of the pristine tube. The functionalized tube retains two eigenstates at the Fermi level, very distinct from the band structure of pattern (β) in Fig. 6-9.

To further analyze the effects of bond cleaving, the quantum conductance of an infinite (5,5) CNT with a single DCC or MCN group attached is calculated. The results shown in Fig. 7-12 confirms that bond cleaving indeed brings only a minor perturbation: the quantum conductance almost approaches its ideal value over a wide energy range around the Fermi level. Comparison with the O closed case allows to single out the effects of bond cleaving. The atomic coordinates of the nanotube backbone are not relaxed to obtain the O closed configuration (one should keep in
mind that this geometry is unstable for DCC or MCN on a (5,5) CNT. The results are plotted in red (DCC) and orange (MCN) lines. Two broad dips appear (as in Fig. 6-10(b)), and the enhanced scattering highlights the importance of bond cleaving.

It is clear that the $sp^2$-rehybridization taking place in the O open configuration is key to restoring the original properties of the tube, and to allow for a chemical ligand to act as an extremely weak perturbation to the original electronic structure. The $p_z$-MLWFs on C₁ and C₆ are slightly misaligned, but still take part in the $\pi$-manifold. To visualize this, we compare in Fig. 7-13 the projected density of states (PDOS) for three cases: two $p_z$-MLWFs in a DCC functionalized CNT, two $sp^3$-bonding MLWFs in a hydrogenated CNT, and two $p_z$-MLWFs of in a pristine CNT. The PDOS of the DCC case closely follows that of the pristine tube except for a very strong peak around $1\,eV$. The peak originates from a quasi-bound state [80] and coincides with the gap in Fig. 7-11 and the conductance dip in Fig. 7-12. The PDOS of the hydrogenated case is in general much smaller and does not have the characteristic van Hove singularities. The two small bumps seen at $1\,eV$ and $-1.2\,eV$ also originate from quasi-bound states.

To simulate realistic functionalized tubes tens of nanometers long, we embed the
Figure 7-13: Density of states projected onto $p_z$-MLWFs or $\sigma$-bonding MLWFs of C$_1$ and C$_6$ for three different CNTs: dichlorocarbene functionalized, hydrogen functionalized, and pristine (5,5) CNT.

Figure 7-14: Quantum conductance for an infinite (5,5) CNT functionalized with one DCC group: comparison between a 5- and a 14-atomic-layer-long disordered central region. The two results are in very good agreement.

Hamiltonian $H_C$ of a conductor functionalized with a single group into the Hamiltonian matrix of a pristine CNT, at randomly generated positions. The original conductor part includes 14 atomic layers, which are too long to be used to this purpose. A segment of 5 atomic layers around the functional group is taken and tested against the original calculation; the agreement is very good (Fig. 7-14) and thus a $H_C$ corresponding to five atomic layers is used from now on. There are ten available choices for O sites along the circumference of a (5,5) CNT. The 5-layer $H_C$ is rotated
Figure 7-15: Quantum conductance for an infinite (5,5) CNT with thirty randomly distributed DCC groups in the open or closed (unstable) O configurations. 70% of the original conductance is preserved in the open configuration. The conductance for the closed configuration is significantly lower.

to generate ten matrices. The position of the functional group along the axis of the tube and the choice among the ten sites are determined by a random number generator. Quantum conductance for a 43-nm-long central disordered region composed of 3500 carbon atoms and 30 randomly distributed functional groups is shown in Fig. 7-15. Even after adding thirty groups, the functionalized tubes still preserve ~70% of the original conductance. Scattering in the (unstable) closed configuration would be much stronger, which again emphasizes the critical role of the bond cleaving.

In conclusion, our early results that covalent functionalizations introduce $sp^3$-bonding and rapidly degrade the metallicity of CNTs does not apply to this class of functionalizations, precisely because bond-breaking restores an $sp^2$ environment. Our result strongly suggests that non-destructive covalent functionalization is possible, which opens a new way of adding chemical functionality while preserving the unique electronic transport properties of carbon nanotubes.
7.3 Molecular homologues

We have not considered up to now the role of chemistry, since orbital rehybridization has been key to rationalizing the transport properties of single-bond ligands, or of cycloadditions. Nevertheless, orbital rehybridization is more subtle in the case of cycloadditions. The critical diameter at which a closed to open transition takes place (Fig. 7-7(b)) will be determined by the chemistry of the functional species, and that transition will be accompanied by a switch from low to high conductance, as shown in Fig. 7-15. Thus, chemistry will affect the conductance by controlling the critical diameter below which the tubes are in the bond-broken, high-conductance configuration. Two directions can now be envisioned to accomplish conductance modulation on the same tube:

- If a functional group displays both an open and a closed stable form, conductance could be controlled by interconversion between the two valence tautomers.

- In-situ chemical changes of the functional group itself (e.g. by oxidation or protonation) may induce a transition between the two forms.

In order to streamline our search for an addend for which both the open and closed configurations could be stable, we first screen several functional groups on isolated molecules. The bridged 1,6-X-[10]annulene in Fig. 7-16(a) turns out to be an excellent molecular homologue of a functionalized CNT [138]. It has the same local bonding structure and symmetry as a functionalized carbon nanotube. The substitutional group X dictates the preference for the 1,6-methano[10]annulene derivative (henceforth labelled as 1o) or for its valence tautomer, a bisnorcaradiene derivative (1c), corresponding to the open and closed configurations of a functionalized nanotube [149–154]. Similar tautomerization takes place in a pyrene derivative (Fig. 7-16(b)), where 8,16-X-[2.2]metacyclophane-1,9-diene (2o, open) and 10b,10c-X-cis-10b,10c-dihydropyrene (2c, closed) have been reported [155]. The molecules 1 and 2 are thus used to screen for the optimal X groups. In the literature, hydrogen and halogens have been reported to stabilize the open annulene form (both experimentally
Figure 7-16: Structure of two molecular homologues for functionalized carbon nanotubes: (a) 1,6-methano[10]annulene; (b) 8,16-methano[2.2]metacyclophane-1,9-diene.

and theoretically [149,150,156,157]), while a cyano group is known to favor the closed bisnorcaradiene form (a possible coexistence with the open form is not excluded) [153]; in this latter case, two minima have been predicted theoretically [157].

A variety of groups corresponding to X= CH$_2$, NH, SiH$_2$, C(NO$_2$)$_2$, C(≡N)$_2$, C(≡CH)$_2$, C(CH$_3$)$_2$, C(COOH)$_2$, CCl$_2$, C(NH$_2$)$_2$, C$_6$H$_4$O (1,1-(4-oxocyclohexa-2,5-dieno)-) [158], and C$_{13}$H$_8$ (9,9-fluoreno-) [158] have been tested in order to identify the factors that determine stability in the open or closed form. The molecules are placed in a cubic cell of size (15.9 Å)$^3$, and Γ-sampling is used. We compare first our results with previous studies. Table 7.1 summarizes $d_{16}$ from experiments and from different first-principles calculations. Our theoretical predictions are consistent with the result from second-order Møller-Plesset perturbation theory (MP2) [159,160] and Becke three parameter Lee-Yang-Parr hybrid functional (B3LYP) [161,162].

We show our calculated potential energy profiles for select groups in Fig. 7-17(a) and (b) and summarize the stable forms in Table 7.2. Among all substituents considered, C(CN)$_2$, C(CCH)$_2$, C$_6$H$_4$O, C$_{13}$H$_8$ and SiH$_2$ have a well-defined minimum in the closed configuration. The driving force for C(CN)$_2$ to stabilize the closed form was attributed to π-electron withdrawing [166], but it has been argued that this simple

\textsuperscript{3}MP2 is a second-order improvement over Hartree-Fock [16]: the first-order energy is the Hartree-Fock energy and the second-order energy provides a sizeable fraction of the electron correlation energy. B3LYP is a hybrid functional, which includes LSDA exchange-correlation, Hartree-Fock exchange and GGA exchange-correlation: the portion of each term is determined by fitting to experimental data. These are two popular and accurate quantum chemistry methods.
Table 7.1: Experimental and theoretical $d_{16}$ of 1. Theoretical values calculated from MP2, B3LYP, and DFT-GGA (this work) show very good agreement. Note that the calculations consider an isolated molecule at 0 K limit while experimental values are obtained from X-ray spectroscopy of a crystal at finite temperature.

<table>
<thead>
<tr>
<th>X</th>
<th>Expt.</th>
<th>MP2</th>
<th>B3LYP</th>
<th>This work</th>
</tr>
</thead>
<tbody>
<tr>
<td>CH$_2$</td>
<td>2.235$^a$</td>
<td>2.251$^f$</td>
<td>2.279$^g$</td>
<td>2.278</td>
</tr>
<tr>
<td>CF$_2$</td>
<td>2.269$^b$</td>
<td>2.268$^f$</td>
<td>2.296$^f$</td>
<td>2.300</td>
</tr>
<tr>
<td>C(CN)$_2$</td>
<td>1.542$^c$</td>
<td>1.599, 2.237$^f$</td>
<td>1.558, 2.253$^h$</td>
<td>1.572, 2.245</td>
</tr>
<tr>
<td>C(CH$_3$)$_2$</td>
<td>1.836$^d$</td>
<td>2.156$^f$</td>
<td>2.168$^f$</td>
<td>2.151</td>
</tr>
<tr>
<td>NH (open)$^e$</td>
<td></td>
<td></td>
<td>2.237$^g$</td>
<td>2.239</td>
</tr>
<tr>
<td>SiH$_2$</td>
<td></td>
<td></td>
<td>1.622, 2.513$^g$</td>
<td>1.637, 2.507</td>
</tr>
</tbody>
</table>

$^a$Ref. [156], $^b$Ref. [163], $^c$Ref. [153], $^d$Ref. [151], $^e$Ref. [164], $^f$Ref. [157], $^g$Ref. [165], $^h$Ref. [154]

Table 7.2: Energy minimum conformations for 1 or 2. The character inside a parenthesis corresponds to a less stable conformation.

<table>
<thead>
<tr>
<th>X</th>
<th>CH$_2$</th>
<th>NH</th>
<th>SiH$_2$</th>
<th>C(NO$_2$)$_2$</th>
<th>C(CN)$_2$</th>
<th>C(CCH)$_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>o</td>
<td>o</td>
<td>o (c)</td>
<td>o</td>
<td>o (c)</td>
<td>o (c)</td>
</tr>
<tr>
<td>2</td>
<td>o</td>
<td>o</td>
<td>o (c)</td>
<td>o</td>
<td>c (o)</td>
<td>c (o)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>X</th>
<th>C(CH$_3$)$_2$</th>
<th>C(COOH)$_2$</th>
<th>CCl$_2$</th>
<th>C(NH$_2$)$_2$</th>
<th>C$_6$H$_4$O</th>
<th>C$_{13}$H$_8$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>c (o)</td>
<td>c</td>
<td>c</td>
</tr>
<tr>
<td>2</td>
<td>c (o)</td>
<td>o</td>
<td>o (c)</td>
<td>o</td>
<td>c</td>
<td>c</td>
</tr>
</tbody>
</table>

The $\pi$-acceptor rule is not comprehensive enough [157]; our results also do not conform to this rule, since another $\pi$-accepting group, C(NO$_2$)$_2$, strongly favors the open form. We have found that the critical factor that stabilizes the closed form is the alignment of the $\pi$-electron system of X with respect to the tangential $p_\sigma$ orbital [167] of the cyclopropane ring shown in Fig. 7-18. When partially occupied $p$-orbitals of X lie parallel to the $p_\sigma$ orbital and effectively withdraw electron density from the cyclopropane HOMO, the anti-bonding interaction between C$_1$ and C$_6$ is weakened, and therefore the bond becomes stronger [168, 169]. All carbene groups stabilizing
Figure 7-17: Potential energy surface as a function of $d_{16}$ for select cases of (a) 1 and (b) 2. C(CN)$_2$ (violet) and C(CCH)$_2$ (blue) show a double-well minimum in both 1 and 2. C(NO$_2$)$_2$-rot (dashed red) indicates the unstable conformation; this configuration strongly stabilizes the closed form, in striking contrast with the case where the two NO$_2$ groups are left at equilibrium (solid red).

The closed configuration satisfy this condition. The argument is strongly supported by examining the energy minimum conformation for $X=$C(NO$_2$)$_2$; the sidewall bond switches from open to closed upon rotation of the two NO$_2$ groups by 90°. Fig. 7-19
Figure 7-18: Dominant interaction between the cyclopropane Walsh orbitals and the acetylene p-orbitals (figure from Ref. [169], few annotations are added).

Figure 7-19: Molecular structure of 1 of X=C(NO₂)₂. (a) Equilibrium conformation: two oxygen atoms in the NO₂ group lie on a line parallel to the C₁-C₆ bond (b) Rotated conformation: two oxygen atoms in the NO₂ group lie on a line perpendicular to the C₁-C₆ bond.

shows clearly that the rotation of the NO₂ groups places the p orbitals of the NO₂ groups parallel to the p σ orbital on C₁₁. The potential profiles for the two cases, plotted in Fig. 7-17, are strikingly different. Electron withdrawal and delocalization effect can also be seen in the change of spread for the MLWFs on the C₁-C₁₁ and C₆-C₁₁ bonds (Fig. 7-20). MLWFs are more spread and the bond length of C₁-C₁₁ is longer in the case of the closed-form stabilizing groups (Fig. 7-21). The stabilizing interaction and the mixing between the π-electron system of X and the cyclopropane ring could also be responsible for the enhanced substituent effects found in functionalized fullerenes. Dicyanomethanofullerene has a strong electron withdraw-
Figure 7-20: Spread of the MLWFs on the C\textsubscript{1}-C\textsubscript{11} (or C\textsubscript{6}-C\textsubscript{11}) bond as a function of \( d_{16} \). The cases in which molecule 1 has X=CH\textsubscript{2}, C(NO\textsubscript{2})\textsubscript{2}, C(NO\textsubscript{2})\textsubscript{2}-rot, and C(CN)\textsubscript{2} are compared. The MLWFs of X=C(NO\textsubscript{2})\textsubscript{2}-rot and C(CN)\textsubscript{2} are more spread than those of X=CH\textsubscript{2} and C(NO\textsubscript{2})\textsubscript{2} which do not have a closed minimum.

Figure 7-21: Maximally-localized Wannier Functions on the C\textsubscript{1}-C\textsubscript{11} (or C\textsubscript{6}-C\textsubscript{11}) bond of 1 in Fig. 7-19: (a) equilibrium conformation; (b) rotated conformation. The structures shown are the optimized ones at \( d_{16} = 1.5 \text{Å} \). The isosurfaces include 95% of the total charge of the MLWFs.

The planarization effect \cite{170}. 9,9-fluorenofullerene influences the reduction potential more strongly than diphenylmethnaofullerene does \cite{158, 171}; the molecular structure of the two
fullerene derivatives are very similar except for the plane in which \( \pi \)-electron lies, as is the case for \( X = \text{C(NO}_2\text{)}_2 \) rotated and at equilibrium.

Now we turn our attention to the open-form stabilizing groups. The energy difference between the closed and the open form spans a wide range even among the species which have a single open minimum for both 1 and 2. In terms of \( E(d_{16} = 1.6) - E(\text{open minimum}) \) of 1, groups are ordered as

\[
\text{CH}_2 > \text{NH} > \text{C(NO}_2\text{)}_2 > \text{C(COOH)}_2 > \text{C(NH}_2\text{)}_2.
\]

A hand-waving rule that would determine this sequence is not apparent. Choi et al. [157] proposed that the sequence is determined by the energy difference between the singlet and triplet minima of a corresponding cyclopropane derivative (\( \text{CH}_2-X-\text{CH}_2 \)); this is supported by their calculation results. Since the above sequence shows that \( X = \text{CH}_2 \) is a strong stabilizer of the open form, most groups would undergo the closed-to-open transition at a diameter smaller than 2.4 \( nm \) where the closed-to-open transition for \( \text{CH}_2 \) was observed (Fig. 7-7(b)).

Functional groups displaying a stable double minimum are of central importance for our goals of conductance modulation. Among the substituents screened, we find that SiH\(_2\), C(CCH)\(_2\) and C(CN)\(_2\) show most clearly the presence of two minima in their potential energy surface. We choose here \( X = \text{C(CN)}_2 \) as the most promising candidate since both 1c [153] and a C\(_{60}\) derivative [170] have already been synthesized and thus we study in the next section the effects of this group on a CNT.

### 7.4 Conductance modulation via orbital rehybridization

Inspired by the finding of a double minimum for the isolated molecular homologues, we examined the potential energy surface for an armchair CNT functionalized with

---

\(^4\)It is likely that small diameter tubes will have the open form regardless of chemical groups since large curvature induces the bond opening.
Figure 7-22: Potential energy surface as a function of $d_{16}$ for an $(n,n)$ CNT functionalized with C(CN)$_2$ in the O configuration (figure from Ref. [143]). Both the (10,10) and (12,12) CNTs display a double-well minimum. Dashed lines are only a guidance to the eye.

Figure 7-23: Potential energy surface as a function of $d_{16}$ for an $(n,n)$ CNT functionalized with CH$_2$ in the O configuration.

$X=C(CN)_2$. The results are shown in Fig. 7-22. The potential energy surfaces reflect closely those found in the molecular homologues. The critical role of curvature in controlling the stability of the open or the closed form is also apparent. A unique minimum in the open form is found in small-diameter tubes, as is generally the case for cycloadditions. As the diameter is increased, the signature of a closed minimum
Figure 7-24: Sidewall equilibrium bond distance $d_{16}$ for $(n,n)$ CNTs and for bent graphene sheets functionalized with C(CN)$_2$ in the O configuration. A double minimum is seen in the bent graphene simulations in the range of $\theta/2 \in [47^\circ, 63^\circ]$. This range overlaps with the curvature of a (10,10) and a (12,12) CNT. To locate the double minimum region, two initial configurations are tried: open (hollow symbols) and closed (solid symbols).

starts to appear, first as an inflection point for the (5,5) CNT ($\phi = 0.69 \text{ nm}$), then as a local minimum for the (10,10) CNT ($\phi = 1.36 \text{ nm}$) as in 1c, and finally as a global minimum for the (12,12) CNT ($\phi = 1.63 \text{ nm}$) as in 2c. For comparison, we plot in Fig. 7-23 the potential energy surfaces of a (5,5) and an (8,8) CNT functionalized with CH$_2$. A steep rise in the energy upon decreasing $d_{16}$ is seen, in clear contrast with the C(CN)$_2$ case. Simulation result of bent graphene in Fig. 7-24 corroborates the existence of double minimum. A double minimum is located for bent graphene in the range of $\theta/2 \in [47^\circ, 63^\circ]$,$^5$ overlapping closely with the carbon nanotube case.

$^5$A narrow range of double minima in the region $\theta/2 \in [28^\circ, 31^\circ]$, roughly corresponding to $n \in [18, 19]$ for an $(n,n)$ CNT, is found for CH$_2$. 

134
Figure 7-25: Quantum conductance for an infinite (10,10) CNT with one or thirty randomly distributed C(CN)$_2$ groups in the open and closed configurations (figure from Ref. [143]): (a) single group; (b) 30 functional groups (averaged over 10 different random configurations). Quantum conductance for a pristine (10,10) CNT is shown as a reference (dashed line).

As shown in Fig. 7-15, quantum conductance is dramatically influenced by the bonding and hybridization of the sidewall carbons. We compare in Fig. 7-25 the quantum conductance for the two structural minima of a C(CN)$_2$ functionalized (10,10) CNT. The scattering by a single group turns out to be negligible, especially for the open form: the conductance around the Fermi energy is extremely close to its ideal value (panel (a)). As the number of functional groups is increased, the difference between the two structural minima (open and closed) becomes rapidly apparent (panel (b)). Even after attaching 30 groups, the quantum conductance of the open form approaches its ideal value, but that of the closed form drops noticeably. Likewise, the characteristic van Hove singularities are preserved in the density of states of the open form while they are blurred in the closed form (Fig. 7-26).

Such dramatic change in the conductance could find many applications in CNT
devices, and several strategies, including photochemical, electrochemical, and thermal mechanisms, can be envisioned to promote interconversion between the two tautomeric forms. Photochemical and electrochemical interconversions in general rely on the energy shift of the frontier orbitals along a reaction coordinate ($d_{16}$ in our case): population or depopulation of those frontier orbitals drives the system in either direction.

A Walsh diagram illustrates the evolution of frontier orbital energies as a function of $d_{16}$ [157, 167]. We plot these for molecule 1 and X=C(CN)$_2$ in Fig. 7-27. Molecules 1 and 2 belong to the $C_{2v}$ (or $2mm$) point group; the symmetry elements are the $\sigma_v(xz)$ and $\sigma_v'(yz)$ mirror planes, and a two-fold rotation axis ($C_2$) along the $z$ direction. Table 7.3 is the character table of the $C_{2v}$ point group. Each orbital belongs to one of the four irreducible representations, $A_1$, $A_2$, $B_1$, or $B_2$, all labelled in Fig. 7-27 (the symmetry of an orbital is actually written in lower case and the overall symmetry of the charge density is written in upper case). In the Walsh diagram, the orbitals with $a_2$ or $b_1$ symmetry have a $yz$ nodal plane; since the $C_1$-$C_6$ bond lies in that plane, the energy of those orbitals is not strongly affected by $d_{16}$. Orbitals with $a_1$ symmetry that have covalent character on the bond are strongly stabilized when $d_{16}$ is decreased whereas those with $b_2$ symmetry show an opposite trend, since $b_2$ orbitals have an $xz$ nodal plane, and this plane bisects the bond.
Figure 7-27: Walsh diagram showing the orbital energy evolution as a function of $d_{16}$ in the case of molecule 1 of X=C(CN)$_2$. The annotations denote the irreducible representation of each orbital. The side figures show the isosurfaces of charge density.

Table 7.3: Character table of point group $C_{2v}$ [172].

<table>
<thead>
<tr>
<th>$C_{2v}$</th>
<th>$E$</th>
<th>$C_2$</th>
<th>$\sigma_v(xz)$</th>
<th>$\sigma_v'(yz)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_1$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$A_2$</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>$B_1$</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>$B_2$</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
</tr>
</tbody>
</table>

Both photochemical excitations or electrochemical reduction or oxidation can populate or depopulate those frontier orbitals which strongly favor the open or closed form, and can modulate the bond distance, and thus conductance, in real time. As a proof of principle, the possibility of interconversion driven by optical excitation is
Figure 7-28: Potential energy surface as a function of $d_{16}$ for the ground and for singlet excited states of molecule 1 (a) and 2 (b) of X=C(CN)$_2$. The atomic coordinates are fixed to those in the ground state. This molecule belongs to the $C_{2v}$ point group and the annotations denote the irreducible representation of each excited state.

tested. Vertical excitation energy to singlet excited states ($S_n$) have been estimated from time-dependent density-functional theory (TDDFT). These calculations have been performed within the Tamm-Dancoff approximation as implemented in CPMD.
The potential energy surface of the excited states for 1 and 2 of X=C(CN)$_2$ are plotted in Fig. 7-28. In the case of 1 (panel (a)), the interconversion is not likely to happen. $S_1$ state in the open form ($B_1$) is mainly derived from the excitations

$$a_2(\text{HOMO}) \rightarrow b_2(\text{LUMO} + 1) \text{ and } a_1(\text{HOMO} - 1) \rightarrow b_1(\text{LUMO}),$$

and the Walsh diagram predicts that those excitations stabilize the open form even more; the slope of the $S_1$ state at the open minimum supports the reasoning. The $S_1$ state for the closed form will not trigger an interconversion either. The excited state with $A_2$ symmetry that might induce an interconversion is unfortunately forbidden [175]. The situation changes in the case of molecule 2 (panel (b)). The $S_1$ state in the open form ($B_1$) is derived from excitations of orbitals with the reversed symmetry,

$$b_2(\text{HOMO} - 1) \rightarrow a_2(\text{LUMO}) \text{ and } b_1(\text{HOMO}) \rightarrow a_1(\text{LUMO} + 2),$$

that will drive the system to the closed conformation.

A similar conclusion is drawn from experimental observations. For 1 with X=CH$_2$, 1o is stable in both the ground state and the $S_1$ excited energy surface [176]. On the other hand, 2o of X=CH$_2$ is stable in the ground state, while the closed 2c is presumed to be the stable tautomer for the $S_1$ energy surface (Fig. 7-29). Our TDDFT calculations for X=CH$_2$ agree with these experimental results (the potential energy surfaces are similar to the case of X=C(CN)$_2$). Though this preliminary result is encouraging, simulations of excited states require careful validations. Density-functional theory cannot predict excited states and TDDFT energy surfaces need to be checked with other theoretical methods. In addition, structural relaxations on the excited energy surface and the coupling with vibronic modes which may allow symmetrically-forbidden excitations should be taken into account. Application to carbon nanotubes will be even more challenging and we leave these as future work.

Temperature will also play an important role in deciding the ratio between the two tautomers since an upper bound for the energy barrier is only $\sim 0.1$ eV (Fig. 7-22).

---

$^6$Norm-conserving pseudopotentials and a planewave cutoff of 90 Ry are used.
Figure 7-29: Isomerization pathway of $2^o$ of $X=CH_2$ by photochemical excitations (figure from Ref. [155]).

Figure 7-30: Photochemical and thermal interconversion between several isomeric forms of a carbene functionalized fullerene (figure from Ref. [178]).

$^{13}$C NMR spectroscopy and X-ray data have captured the temperature-dependent equilibrium of similar fluxional systems: e.g., a higher temperature stabilizes the $1^o$ form of $X=C(CN)(CH_3)$, while destabilizing it in the case of $X=C(CH_3)_2$ [152, 177].

Electrochemical, photochemical and thermal conversion from [5,6] open to [6,6] closed isomer of functionalized fullerenes have been demonstrated (Fig. 7-30) [178–181]. We believe that the similar considerations would apply to the case of carbon nanotubes.
Conclusions

In the present work, we have developed an optimal, large-scale first-principles method to calculate electronic structure and quantum conductance of realistic nanoscale materials. The application of the present method to the case of covalently-functionalized carbon nanotubes has highlighted the accuracy and the efficiency of the method and at the same time has uncovered the central role of orbital rehybridizations in controlling and modulating the quantum conductance of carbon nanotubes.

At the core of our method lies our ability to derive exact tight-binding Hamiltonians from first-principles calculations, employing maximally-localized Wannier functions as explicit tight-binding orbitals. Application to the case of a (5,5) metallic and an (8,0) semiconducting carbon nanotube has demonstrated that the full band dispersions across the Brillouin zone are exactly reproduced by these tight-binding Hamiltonians constructed from the Γ-point. This result validates the whole procedure: from the construction of MLWFs by disentanglement and localization, to a reconstruction of the Hamiltonian matrix and Green’s functions in real-space, and highlights its general applicability to any system, irrespective of band gaps. The method can be applied to study much larger systems including thousands or tens of thousands of atoms, constructing the Hamiltonian matrix from smaller submatrices obtained from full first-principles calculations.

The energetics, electronic structure, and quantum conductance of functionalized armchair metallic carbon nanotubes have been extensively studied employing the present method.

The first class of functionalizing addends, represented by hydrogen and aryl groups, turns out to strongly damage the conductance of the tubes. In fact, these covalent
functionalizations replace $p_z$-MLWFs on the sidewall of a pristine tube by $sp^3$-bonding MLWFs whose on-site energy is lowered by several $eV$. Thus, the original $p_z$-MLWFs disappear from the $\pi$-manifold, which leads to much increased scattering. The chemical difference between the functional groups plays only a minor role since the major effect comes from $sp^3$ rehybridization, which is the common factor for this class of functionalizations. Functional groups are found to be more stable when they pair or cluster; this energetically favorable configuration slightly suppresses the perturbation to the conjugated network, results in higher conductance at the Fermi level than in the case of a single isolated group. Still, even when paired, single-bond ligands remain strong scatterers that decrease significantly the quantum conductance of armchair metallic tubes, as more and more pairs are attached.

The second class of functionalizing addends studied here, represented by carbenes and nitrenes, offers a radically different avenue. These addends bond to two neighboring sidewall carbon atoms, thereby creating a cyclopropane-like three-membered ring structure. On narrow-diameter tubes, the addend induces cleaving of the sidewall bond, restoring $sp^2$ hybridization for the two sidewall carbon atoms involved. The scattering effect of a single carbene or nitrene group then becomes extremely weak: the quantum conductance of the bond-cleaved tubes can approach the ideal limit of a pristine tube. This result suggests that cycloadditions are an ideal way of adding chemical functionalities to metallic CNTs, while preserving their excellent electronic transport properties.

Curvature plays a major role in determining bond stability. High curvature induces bond opening whereas low curvature favors bond closing. In graphene, the bond is always closed. However, between these two extremes, chemistry determines the critical curvature at which the closed-to-open transition takes place. The preference for bond opening or closing has been tested for many functional groups using molecular homologues of functionalized tubes. It is found that a subclass of addends, exemplified by dicyanocarbene, exhibits a bistable state for a narrow range of diameters. Most importantly, the quantum conductance in the open state is much higher than in the closed state. This result suggests an extremely promising way of con-
trolling the conductance of carbon nanotubes via interconversion between the two tautomeric forms. Optical, electrochemical, or thermal methods able to promote interconversion have been suggested, based on the vast body of knowledge accumulated on closely-related molecular homologues, from annulenes to fullerenes.
Bibliography


152


Y.-S. Lee and N. Marzari. Cycloaddition functionalizations to preserve or control the conductance of carbon nanotubes. submitted, 2006.


