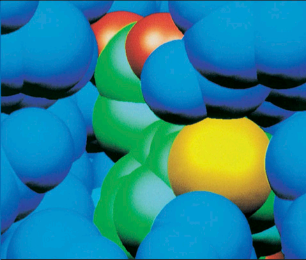
ELSEVIER

THEORETICAL AND COMPUTATIONAL CHEMISTRY 17



Molecular and Nano Electronics

Analysis, Design and Simulation

Jorge M. Seminario
editor

17

THEORETICAL AND COMPUTATIONAL CHEMISTRY

Molecular and Nano Electronics

THEORETICAL AND COMPUTATIONAL CHEMISTRY

Series Editors

Professor P. Politzer
Department of Chemistry
University of New Orleans
New Orleans, LA 70148, U.S.A.

VOLUME 1

Quantitative Treatments of Solute/Solvent Interactions
P. Politzer and J.S. Murray (Editors)

VOLUME 2

Modern Density Functional Theory: A Tool for Chemistry
J.M. Seminario and P. Politzer (Editors)

VOLUME 3

Molecular Electrostatic Potentials: Concepts and Applications
J.S. Murray and K. Sen (Editors)

VOLUME 4

Recent Developments and Applications of Modern Density Functional Theory
J.M. Seminario (Editor)

VOLUME 5

Theoretical Organic Chemistry
C. Párkányi (Editor)

VOLUME 6

Pauling's Legacy: Modern Modelling of the Chemical Bond
Z.B. Maksic and W.J. Orville-Thomas (Editors)

VOLUME 7

Molecular Dynamics: From Classical to Quantum Methods
P.B. Balbuena and J.M. Seminario (Editors)

VOLUME 8

Computational Molecular Biology
J. Leszczynski (Editor)

VOLUME 9

Theoretical Biochemistry: Processes and Properties of Biological Systems
L.A. Eriksson (Editor)

Professor Z.B. Maksić
Rudjer Bošković Institute
P.O. Box 1016,
10001 Zagreb, Croatia

VOLUME 10

Valence Bond Theory
D.L. Cooper (Editor)

VOLUME 11

Relativistic Electronic Structure Theory, Part 1. Fundamentals
P. Schwerdtfeger (Editor)

VOLUME 12

Energetic Materials, Part 1. Decomposition, Crystal and Molecular Properties
P. Politzer and J.S. Murray (Editors)

VOLUME 13

Energetic Materials, Part 2. Detonation, Combustion
P. Politzer and J.S. Murray (Editors)

VOLUME 14

Relativistic Electronic Structure Theory, Part 2. Applications
P. Schwerdtfeger (Editor)

VOLUME 15

Computational Materials Science
J. Leszczynski (Editor)

VOLUME 16

Computational Photochemistry
M. Olivucci (Editor)

VOLUME 17

Molecular and Nano Electronics: Analysis, Design and Simulation
J.M. Seminario (Editor)

VOLUME 18

Nanomaterials: Design and Simulation
P.B. Balbuena and J.M. Seminario (Editors)

17

THEORETICAL AND COMPUTATIONAL CHEMISTRY

Molecular and Nano Electronics: Analysis, Design and Simulation

Edited by

J. M. Seminario

Department of Chemical Engineering
and Department of Electrical and Computer Engineering
Texas A&M University
College Station, Texas, USA.

ELSEVIER

AMSTERDAM • BOSTON • HEIDELBERG • LONDON • NEW YORK • OXFORD
PARIS • SAN DIEGO • SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO

Elsevier
Radarweg 29, PO Box 211, 1000 AE Amsterdam, The Netherlands
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, UK

First edition 2007

Copyright © 2007 Elsevier B.V. All rights reserved

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means electronic, mechanical, photocopying, recording or otherwise without the prior written permission of the publisher

Permissions may be sought directly from Elsevier's Science & Technology Rights Department in Oxford, UK: phone (+44) (0) 1865 843830; fax (+44) (0) 1865 853333; email: permissions@elsevier.com. Alternatively you can submit your request online by visiting the Elsevier web site at http://elsevier.com/locate/permissions, and selecting *Obtaining permission to use Elsevier material*

Notice
No responsibility is assumed by the publisher for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions or ideas contained in the material herein. Because of rapid advances in the medical sciences, in particular, independent verification of diagnoses and drug dosages should be made

Library of Congress Cataloging-in-Publication Data
A catalog record for this book is available from the Library of Congress

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN-13: 978-0-444-52725-7
ISBN-10: 0-444-52725-7
ISSN: 1380-7323

For information on all Elsevier publications visit our website at books.elsevier.com

Printed and bound in Italy
07 08 09 10 11 10 9 8 7 6 5 4 3 2 1

Working together to grow
libraries in developing countries

www.elsevier.com | www.bookaid.org | www.sabre.org

ELSEVIER BOOK AID International Sabre Foundation

Contents

Preface

The new field of molecular and nano-electronics brings possible solutions for a post-microelectronics era. Microelectronics is dominated by the use of silicon as the preferred material and photo-lithography as the fabrication technique to build binary devices (transistors). Properly building such devices yields gates, able to perform Boolean operations and to be combined yielding computational systems capable of storing, processing, and transmitting digital signals encoded as electron currents and charges.

Since the invention of the integrated circuits, microelectronics has reached increasing performances by decreasing strategically the size of its devices and systems, an approach known as scaling-down, which simultaneously allow the devices to operate at higher speeds. However, as devices become faster and smaller, major problems have arisen related to removal of heat dissipated by the transistors and physical limitations to keep two well-defined binary states; these problems have triggered research into new alternatives using components fabricated by different procedures (self-assembly, chemical deposition, etc.), which may encode information using lower energetic means.

The goal of this book is to bring together the most active researchers in this new field, from the entire world. These researchers illustrate what is probably the only way for success of molecular and nano-electronics: a theory guided approach to the design of molecular- and nano-electronics. The editor thanks all the contributors for their kind collaboration, effort, and patience to put together this volume, and acknowledges the dedication of Ms Mery Diaz who helped compiling this camera-ready volume. The editor also thanks the continuous support of the US Army Research Office to the development of this new field.

Jorge M. Seminario
Texas A&M University

Molecular and Nano Electronics: Analysis, Design and Simulation
J. M. Seminario (Editor)

Chapter 1

Metal–molecule–semiconductor junctions: An ab initio analysis

Luis A. Agapito and Jorge M. Seminario

Department of Chemical Engineering and Department of Electrical and Computer Engineering, Texas A&M University, 3122 TAMU, College Station, TX 77843, USA.
seminario@tamu.edu

1. Introduction

The ability to calculate the current–voltage characteristics through a single molecule is essential for the engineering of molecular electronic devices [1, 2]. Because quantum-mechanical effects prevail at atomistic sizes, there is a need to implement and develop precise ab initio quantum chemistry techniques rather than using those originally developed for microscopic and mesoscopic systems.

In order to evaluate experimentally the use of single molecules as electronic devices, the usual approach is to attach them to macroscopic contacts to be able to measure their electrical properties. However, this is not a direct requirement for the design but just to help us to understand their electrical behavior and to make sure that we have the correct tools to model their behavior. In practice, molecular devices should not be connected to macroscopic contacts when they are components of a circuit. The whole advantage of having nanosized devices would be lost if they are connected to macroscopic or even microscopic contacts. Nevertheless, the presence of macroscopic contacts influences greatly the electrical properties of a single molecule [3–5]; thus our community tries to test the metal–molecule–metal junction as an independent unit instead of evaluating the isolated molecule. Experimentally, it has been challenging to measure metal–molecule–metal junctions with metallic contacts separated by a distance of ~20 Å or less. Only few experiments until now have claimed to have been able to address a single molecule between two macroscopic gold contacts [6].

Fortunately, quantum-chemistry techniques can be used to study precisely isolated [7, 8] and interconnected molecules. We use the Density Functional Theory (DFT) of a quantum-chemistry flavor [9] to determine the electronic properties of molecules; a mathematical formalism based on the Green function (GF) is used then to account for the effect of the contacts on the molecule keeping the realistic chemical nature of the sandwiched molecule. These techniques can also be used to study scenarios where the information is not coded in electron currents [10–12].

The electron transport in quantum chemistry is studied as a chemical reaction or as a state transition through junctions of atomic sizes and can also be approximately described in terms of mesoscopic physics models in a coherent regime, where the electrons travel with a given probability, sequentially one after the other through the molecule without electron–electron or phonon–electron interactions. This kind of transport is described by the Landauer formalism [13]. Here, we use our DFT-GF technique [14] to make an atomistic adaptation of the Landauer formalism for the calculation of current through molecular junctions.

Specifically, we focus our study on an oligo-phenylene-ethynylene (OPE) molecule, which has been proposed as a candidate for a molecular electronic device [15]. Similar OPE molecules, attached to gold contacts, have shown two distinctive states of conductance, namely a high- and a low- conductance state. Those states can be used to encode information as logic "0" and "1," hence, their importance. Switching between the two states of the molecule is mainly attributed to two different mechanisms: changes in charge state [15] and changes in conformational states [16].

We use our DFT-GF formalism to calculate the conductance through metal–nitroOPE–metal junctions in several charge and conformational states. Two different metallic materials are evaluated in this work: the commonly used gold and the promising carbon nanotube (CNT).

2. Electron transport at interfaces

From the computational viewpoint, primarily two types of molecular systems are involved in the work presented here: finite and extended systems. Finite systems refer to molecules or nanoclusters with a finite number of atoms whereas an extended system refers to a crystalline such as the contacts. The tools to study both types of systems are well-established in the computational chemistry field [1, 2, 17–20]. The Gaussian 03 [21] is capable of performing calculations of systems with periodic boundary conditions in one, two and three dimensions. However, systems that combine both a finite and an extended character represent a new and challenging area of research; this is the case for the study of a single molecule (finite) adsorbed to contact tips (modeled as an infinite crystal material).

The discrete electronic states of an isolate molecule are obtained by solving the Schrödinger equation; we solve that equation following a DFT approach. When the molecule is adsorbed on a contact tip, the continuous electronic states of bulk material modify the discrete electronic states of the molecule. In other words, electrons from the contacts leak into the molecule, modifying its electronic properties. A mathematical formalism based on the Green function is used to account for the effect of the bulk contacts.

2.1. Electronic properties of molecules and clusters

The electronic properties of a molecular system can be calculated from its auxiliary wavefunction, which is built as a determinant of molecular orbitals (MOs). MOs are linear combinations of atomic orbitals (AOs) from all the atoms composing the system. In other words, the atomic orbitals are the basis functions, χ, used to expand the MOs shown in Eq. (1).

2.1.1. Basis functions

Practical procedures represent the AO using linear combination of Gaussian functions also called primitives. Gaussian-type functions (GTFs) or primitives, which form a complete set of functions, are defined in their Cartesian form as:

$$g_{ijk} = K x_b^i y_b^j z_b^k e^{-\alpha r_b^2} \tag{1}$$

where i, j, k are nonnegative integers, α is a positive orbital exponent, x_b, y_b, z_b, are Cartesian coordinates and r_b is the radial coordinate. The subscript b indicates that the origin of the coordinates is at the nucleus b. K is a normalization constant.

The sum $l = x + y + z$ determines the angular momentum of an atomic orbital. Depending on whether l equals 0, 1, 2, . . . , the GTF is called s-, p-, d-, . . . type respectively. The principal quantum number n determines the range of the exponents for a particular function.

A basis function χ_r, also referred to as contracted Gaussian-type orbitals (GTOs), is defined as a normalized linear combination of GTFs (g_u) or primitives:

$$\chi_r = \sum_u d_{ur} g_u \tag{2}$$

where d_{ur} are called contraction coefficients. Basis sets published in the literature provide the values of α, Eq. (1), and d_{ur}, Eq. (2). A basis function is constructed to resemble a given AO. Throughout this work, two basis sets are used: the LANL2DZ, which includes an effective core potential and the 6-31G(d) also represented as 6-31G* in the specialized literature.

For instance, in the 6-31G(d) basis set, the inner shell 1s atomic orbital of carbon is formed by contracting six GTFs, as follows:

$$\chi_{1s} = \sum_{u=1}^{6} d_u g_{1s}(\alpha_u) \tag{3}$$

where the contraction coefficients d_u and the Gaussian exponents α_u are given in Table 1. For an s-type function, the GTF given in Eq. (1) simplifies to

$$g_{1s}(\alpha) = e^{-\alpha(x^2+y^2+z^2)} \tag{4}$$

where the normalization constant K, defined in Eq. (1), has been included in the contraction coefficients.

2.1.2. Density functional theory

For a polyatomic molecular system, the electronic non-relativistic Hamiltonian can be written as

$$\hat{H}_{el} = -\frac{1}{2}\sum_i \nabla_i^2 + \sum_i \sum_b \frac{Z_b}{r_{ib}} + \sum_i \sum_{j>i} \frac{1}{r_{ij}} \tag{5}$$

Table 1 Contraction coefficients and Gaussian exponents for the inner 1s atomic orbital of the carbon atom (Eq. 3) corresponding to the 6-31G(d) basis set

u	Contraction coefficients d_u	Gaussian exponents α_u
1	0.001834700	3047.52490
2	0.014037300	457.369510
3	0.068842600	103.948690
4	0.232184400	29.2101550
5	0.467941300	9.28666300
6	0.362312000	3.16392700

where i and j count over all electrons and b counts over all nuclei, and Z_b is the atomic number of the atom b. If the system contains n electrons then the wavefunction of the molecular system is a function of $3n$ spatial coordinates and n spin coordinates. Therefore, calculating the complete electronic wavefunction is computationally challenging simply because the wavefunction is a mathematical function that contains more information of the system than needed for specific applications.

The first Hohenberg–Kohn theorem [22] established that all the properties of a molecular system in the ground state are determined by the ground-state electron density $\rho_0(x, y, z)$, which is a function of only three variables. This theorem circumvents the use of the wavefunction; instead, the electron density function is used to calculate the properties of a molecular system. This theorem together with the constrain search of Levy [23] finally sets DFT on a formal basis.

In 1965 Kohn and Sham [24] published a method to determine the electron density without having to find the real wavefunction. They demonstrated that the electron density of a molecular system of interacting electrons can be represented with the electron density of an ideal or ficticious system of non-interactive electrons subjected to an effective potential ν_{s}. Therefore, the interacting many-electron problem is split into several non-interacting one-electron problems, which are governed by the following one-electron Kohn–Sham (KS) equations:

$$\hat{h}^{\mathrm{KS}}(r)\theta_i^{\mathrm{KS}}(r) = \varepsilon_i^{\mathrm{KS}}\theta_i^{\mathrm{KS}}(r) \tag{6}$$

where the one-electron KS Hamiltonian $\hat{h}^{\mathrm{KS}}$ is defined as:

$$\hat{h}^{\mathrm{KS}}(r) = -\frac{1}{2}\nabla_r^2 + \nu_s(r) \tag{7}$$

and the external potential for the fictitious electrons is defined as:

$$\nu_s(r) = -\sum_b \frac{Z_b}{|r - r_b|} + \int \frac{\rho(r')}{|r - r'|}dr' + \nu_{\mathrm{xc}}(r) \tag{8}$$

where ν_{xc} is the exchange-correlation potential

$$\nu_{xc}(r) \equiv \frac{\delta E_{\mathrm{xc}}[\rho(r)]}{\delta\rho(r)} \tag{9}$$

The external potential v_s is found by solving Eq. (6) self-consistently. The KS molecular orbitals (θ_i^{KS}), shown in Eq. (6), are expanded in terms of the GTOs defined in Eq. (2).

$$\theta_i^{\mathrm{KS}} = \sum_{r=1}^{B} c_{ri}\chi_r \tag{10}$$

where B is the number of basis functions of the molecular system. By inserting Eq. (10) in Eq. (6) and applying the variational principle, a Roothaan-type matrix equation is obtained. For example, the matrix equation for a molecular system that has only five basis functions is

$$\begin{pmatrix} h_{11}^{\mathrm{KS}} & h_{12}^{\mathrm{KS}} & h_{13}^{\mathrm{KS}} & h_{14}^{\mathrm{KS}} & h_{15}^{\mathrm{KS}} \\ h_{21}^{\mathrm{KS}} & h_{22}^{\mathrm{KS}} & h_{23}^{\mathrm{KS}} & h_{24}^{\mathrm{KS}} & h_{25}^{\mathrm{KS}} \\ h_{31}^{\mathrm{KS}} & h_{32}^{\mathrm{KS}} & h_{33}^{\mathrm{KS}} & h_{34}^{\mathrm{KS}} & h_{35}^{\mathrm{KS}} \\ h_{41}^{\mathrm{KS}} & h_{42}^{\mathrm{KS}} & h_{43}^{\mathrm{KS}} & h_{44}^{\mathrm{KS}} & h_{45}^{\mathrm{KS}} \\ h_{51}^{\mathrm{KS}} & h_{52}^{\mathrm{KS}} & h_{53}^{\mathrm{KS}} & h_{54}^{\mathrm{KS}} & h_{55}^{\mathrm{KS}} \end{pmatrix} C = \begin{pmatrix} S_{11} & S_{12} & S_{13} & S_{14} & S_{15} \\ S_{21} & S_{22} & S_{23} & S_{24} & S_{25} \\ S_{31} & S_{32} & S_{33} & S_{34} & S_{35} \\ S_{41} & S_{42} & S_{43} & S_{44} & S_{45} \\ S_{51} & S_{52} & S_{53} & S_{54} & S_{55} \end{pmatrix} E^{\mathrm{KS}} C \tag{11}$$

where h_{rs}^{KS} are matrix elements of the one-electron KS Hamiltonian operator $\hat{h}^{\mathrm{KS}}$. These matrix elements are defined as:

$$h_{rs}^{\mathrm{KS}} = \langle \chi_r | \hat{h}^{\mathrm{KS}} | \chi_s \rangle \tag{12}$$

The overlap integral S_{jk} between two basis functions is:

$$S_{jk} = \langle \chi_j | \chi_k \rangle \tag{13}$$

C is a matrix composed of the expansion coefficients c_{ri}, which are defined in Eq. (10). E^{KS} is a diagonal matrix composed of all the eigenvalues (energies) of the one-electron KS equation defined in Eq. (6).

$$E^{\mathrm{KS}} = \begin{pmatrix} \varepsilon_1^{\mathrm{KS}} & 0 & 0 & 0 & 0 \\ 0 & \varepsilon_2^{\mathrm{KS}} & 0 & 0 & 0 \\ 0 & 0 & \varepsilon_3^{\mathrm{KS}} & 0 & 0 \\ 0 & 0 & 0 & \varepsilon_4^{\mathrm{KS}} & 0 \\ 0 & 0 & 0 & 0 & \varepsilon_5^{\mathrm{KS}} \end{pmatrix} \tag{14}$$

The expansion coefficients, c_{ri}, of the molecular orbitals are found by solving iteratively Eq. (11) [25].

$$\rho = \sum_{i=1}^{n} |\theta_i^{\mathrm{KS}}|^2 \tag{15}$$

At all steps of the iteration, the expansion coefficients are updated. Consequently, new KS molecular orbitals, Eq. (10), and electron densities, Eq. (15), are obtained during the iterative process. When self-consistency is reached, the ground-state electron density and KS molecular orbitals can be evaluated. All properties for the molecular system can be extracted from the ground-state density, according to the Hohenberg–Kohn theorem.

2.1.3. Molecular electrostatic potential

A molecular system can be modeled as an electronic device, encapsulating all the chemistry of the system behind the electron density ρ. The equivalent electrostatic potential ϕ for such electronic device, measured at a point of space $r = (x, y, z)$, can be calculated as:

$$\phi(x, y, x) = \sum_{\alpha} \frac{Z_{\alpha}}{|r - R_{\alpha}|} - \iiint \frac{\rho(x', y', z')}{|r - r'|} \mathrm{d}x'\mathrm{d}y'\mathrm{d}z' \tag{16}$$

where the electron density ρ is defined in Eq. (15).

2.2. Electronic properties of crystalline materials

In the case of finite systems, atomic orbitals, Eq. (2), are used to build up the molecular orbitals. For infinite systems, Bloch functions $\phi_{\mu}(\vec{r}; \vec{k})$, are used to build up crystalline orbitals $\psi_i(\vec{r}; \vec{k})$:

$$\psi_i(\vec{r}; \vec{k}) = \sum_{\mu} c_{\mu i}(\vec{k}) \phi_{\mu}(\vec{r}; \vec{k}) \tag{17}$$

where $\vec{r}$ and $\vec{k}$ represent vectors in the direct and reciprocal space, respectively. Bloch functions are defined as follows

$$\phi_{\mu}(\vec{r}; \vec{k}) = \sum_{\vec{T}} \chi_{\mu}(\vec{r} - \vec{A}_{\mu} - \vec{T}) e^{i\vec{k}.\vec{T}} \tag{18}$$

where $\vec{T}$ represents all direct lattice vectors. χ_{μ} represents contracted GTOs as defined in Eq. (2). The subscript μ counts over all the basis functions used to expand the unit cell, $\vec{A}_{\mu}$ indicates the coordinates of the atom on which χ_{μ} is centered. The Bloch functions Eq. (18) are constructed to satisfy the Bloch theorem:

$$\phi_{\mu}(\vec{r} + \vec{T}; \vec{k}) = \phi_{\mu}(\vec{r}; \vec{k}) e^{i\vec{k}\vec{T}} \tag{19}$$

Bloch functions with different wavevectors, k, do not interact each other; therefore, a periodic system can be solved independently for each value of k.

A crystalline orbital Eq. (17) resembles the definition of an MO Eq. (10) in finite systems. The expansion coefficients for the crystalline orbitals $c_{\mu i}$, Eq. (17), are found analogously to the case of finite systems. The matrix $C(k)$, which contains the coefficients $c_{\mu i}$, is found by solving self-consistently Eq. (20) for each k point.

$$H^{\mathrm{KS}}(k) C(k) = S(k) C(k) E(k) \tag{20}$$

where $H^{\mathrm{KS}}(k)$ is the Kohn–Sham Hamiltonian matrix in reciprocal space

$$H_{\mu\nu}^{\mathrm{KS}}(k) = \left\langle \phi_{\mu}(\vec{r}; \vec{k}) | \hat{h}^{\mathrm{KS}} | \phi_{\nu}(\vec{r}; \vec{k}) \right\rangle = \sum_{\vec{T}} \left\langle \chi_{\mu}(\vec{r} - \vec{A}_{\mu} - \vec{0}) | \hat{h}^{\mathrm{KS}} | \chi_{\nu}(\vec{r} - \vec{A}_{\nu} - \vec{T}) \right\rangle e^{i\vec{k}\vec{T}}$$

$S(k)$ is the overlap matrix over the Bloch functions

$$S_{\mu\nu}(k) = \left\langle \phi_\mu(\vec{r};\vec{k}) | \phi_\nu(\vec{r};\vec{k}) \right\rangle = \sum_{\vec{T}} \left\langle \chi_\mu(\vec{r} - \vec{A}_\mu - \vec{0}) | \chi_\nu(\vec{r} - \vec{A}_\nu - \vec{T}) \right\rangle e^{i\vec{k}\vec{T}}$$

$E(K)$ is a diagonal matrix that contains the eigenvalues ε_i^k for a given point k. The number of eigenvalues per k point is equal to the number of basis functions of the unit cell, and $C(k)$ contains column-wise the coefficients of the crystalline orbitals.

The density of states (DOS) of the infinite system is found according to

$$\mathrm{DOS}(\varepsilon) = 2\sum_{i,k} \delta(\varepsilon - \varepsilon_i^k) = \frac{2}{V_{\mathrm{BZ}}} \sum_i \int_{\mathrm{BZ}} \delta(\varepsilon - \varepsilon_i^k) d^3k \tag{21}$$

where V_{BZ} is the volume of the first Brillouin zone. The software Crystal 03 [26] is used to calculated the DOS for the different crystalline materials that are used throughout this work.

2.2.1. DOS of Au and Pd crystals

The Au and the Pd crystals are modeled as FCC lattices with space group number 225. The lattice parameters for the conventional cells are $a = 4.078$ Å for gold and $a = 3.891$ Å for palladium (Figures 1 and 2). The primitive cell for both crystals contains one atom and is defined by the following primitive vectors

$$A_1 = \frac{1}{2}a\hat{y} + \frac{1}{2}a\hat{z},$$
$$A_2 = \frac{1}{2}a\hat{x} + \frac{1}{2}a\hat{z}$$

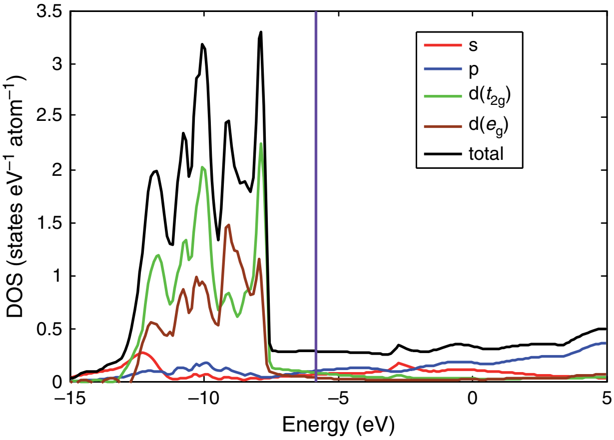


Figure 1 DOS for the Au crystal. Fermi level is at −5.83 eV using the B3PW91 functional with the LANL2DZ basis set and ECP

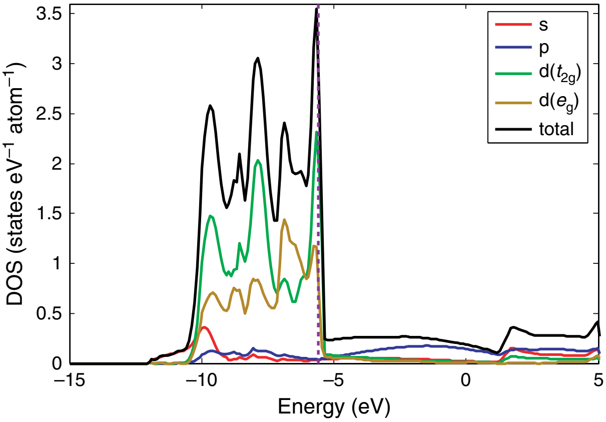


Figure 2 DOS for the Pd crystal obtained using the B3PW91 functional and LANL2DZ basis set and ECP. The Fermi level is at −5.59 eV

and

$$A_3 = \frac{1}{2}a\hat{x} + \frac{1}{2}a\hat{y}$$

The calculation of the electronic structure is performed at the B3PW91 level of theory combined with the LANL2DZ basis set. The total DOS for gold (black curve) is reported in Figure 1. We also compute the contribution of each type of basis function (s-, p-, or d-type) to the total DOS. For consideration to the symmetry of the d-type functions, their contribution is split into two groups: the contribution of the d_{xz}, d_{yz}, d_{xy} basis functions, which present t_{2g} symmetry and the contribution of the $d_z{}^{2*}$, $d_x{}^2{}_{-y}{}^2$ basis functions, which present e_g symmetry. For Au and Pd, most of the electrons available for conduction (at their Fermi level) have a d-character.

2.2.2. *DOS of silicon crystal*

Silicon presents a crystal structure of the diamond (point group number 227). The conventional cell has a lattice parameter $a = 5.42$ Å. The primitive cell is defined by the following primitive vectors

$$A_1 = \frac{1}{2}a\hat{y} + \frac{1}{2}a\hat{z},$$

$$A_2 = \frac{1}{2}a\hat{x} + \frac{1}{2}a\hat{z}$$

and

$$A_3 = \frac{1}{2}a\hat{x} + \frac{1}{2}a\hat{y}$$

---

[*] The orbital d_{z^2} is more properly referred as $d_{2z^2-x^2-y^2}$.

with two atoms per each primitive cell, the basis vectors for these atoms are

$$B_1 = -\frac{1}{8}A_1 - \frac{1}{8}A_2 - \frac{1}{8}A_3 = -\frac{1}{8}a\hat{x} - \frac{1}{8}a\hat{y} - \frac{1}{8}a\hat{z}$$
$$B_2 = +\frac{1}{8}A_1 + \frac{1}{8}A_2 + \frac{1}{8}A_3 = +\frac{1}{8}a\hat{x} + \frac{1}{8}a\hat{y} + \frac{1}{8}a\hat{z}$$

The crystal is calculated using the B3PW91 level of theory. Two sets of calculations, using different basis sets, are carried out.

The full-electron 6-31G(*d*) basis set uses four *s*-type, nine *P*-type, and six *d*-type Gaussian functions to represent the electrons of a and Si atom. The total DOS and the s, p, and d projections obtained using that basis set are shown in Figure 3. The states around the Fermi level have mostly a p-character and a bandgap of 0.72 eV; the midgap is at −2.22 eV.

The LANL2DZ basis set supports elements with large atomic numbers, such as gold. Whenever the molecule under study contains gold atoms, the system is calculated using the LANL2DZ basis set. Therefore, for compatibility purposes, the DOS of Si using the LANL2DZ basis set is also obtained. LANL2DZ is not a full-electron basis set for Si; only the four valence electrons are considered in the calculations; the remaining ten core-electrons are modeled by an effective core potential (ECP). The Si DOS using this basis set is reported in Figure 4; notice that there is not d projection of the total DOS since no d-type polarization functions are used for Si in the LANL2DZ basis set.

2.2.3. DOS of the (4, 4) CNT

Single-walled carbon nanotubes SWCNTs are one-dimensional crystals with interesting mechanical and electrical properties (See for instance [27]). The geometry and

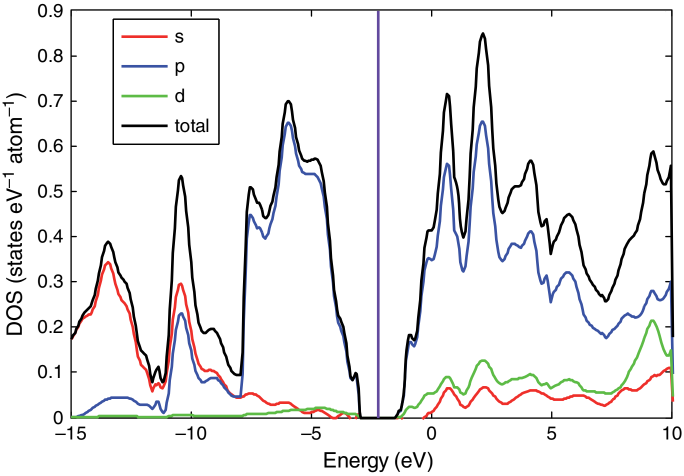


Figure 3 DOS for a silicon crystal calculated using the B3PW91 method and the 6-31G(d) basis set

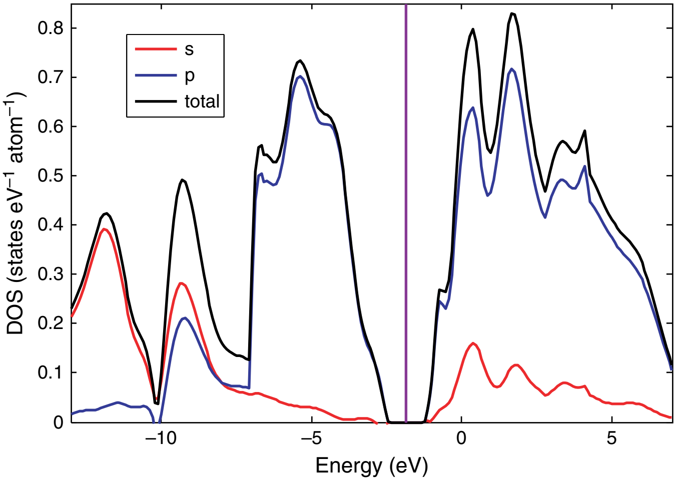


Figure 4 DOS for a silicon crystal calculated using the B3PW91 DFT method and the LANL2DZ basis set. The Fermi level for the material (purple line) is at −1.85 eV. The calculated bandgap is 1.11 eV

the electrical behavior of an SWCNT are defined by a pair of integers (m, n). It is known [28] that

Condition	Type	Bandgap
$n - m = 3q$	Semimetallic	∼m eV
$n - m = 3q$	Metallic	0 eV
$n - m = 3q$	Semiconductor	0.5–1 eV

where q is a non-zero integer. Recent breakthroughs in synthetic chemistry [29] have opened the possibility of using metallic CNTs as contacts to organic molecules. We use the (4, 4) CNT, which is a metal according to the above table, to explore the electrical characteristics of CNT–nitroOPE–CNT molecular junctions. The DOS of the (4, 4) CNT, Figure 5, is calculated using the B3PW91 DFT functional and the 6-31G basis set. Despite the presence of a gap in the CNT DOS at ∼3.50 eV, the absence of gaps at the Fermi level confirms the metallic character of this material. The calculated DOS is in agreement with previous experimental [30, 31] and theoretical [32–35] findings. A unit cell of the (4, 4) CNT is modeled by 16 carbon atoms.

2.3. *Combined DFT-GF approach to calculate the DOS of a molecule adsorbed on macroscopic contacts*

An isolated molecule has discrete electronic states, which are precisely calculated from the Schrödinger equation. When the molecule is attached to macroscopic contacts, the continuous electronic states of the contacts modify the electronic properties of the molecule. A technique that combines the Density Functional Theory and the Green

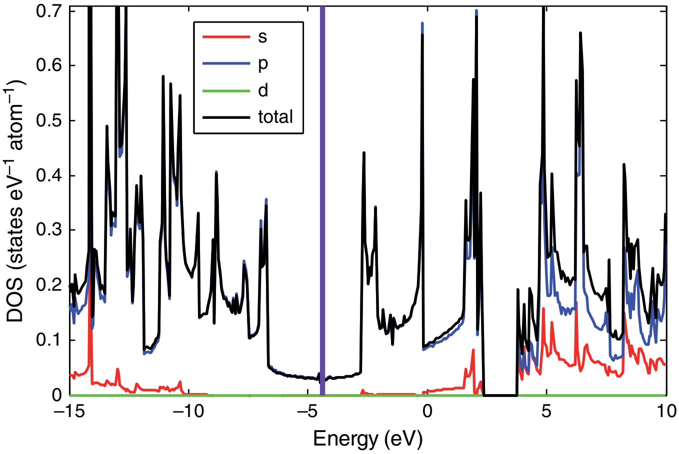


Figure 5 DOS for the metallic (4, 4) CNT, which is calculated using the B3PW91/6-31G method and basis set. The Fermi level (purple vertical line) is at −4.39 eV

function (DFT-GF) [36, 37] is used to account for the effect of the contacts on the electronic states of adsorbed molecule.

In a real system, molecules are chemically attached to real contacts, made of atoms, and not to ideal surfaces. Therefore, information about the interface, obtained at the molecular level, needs to be provided. This is accomplished through coupling matrices obtained from quantum-mechanical calculations of the extended molecule (i.e., the molecule attached to a few atoms from the contacts). The Gaussian 03 [21] software is used for the quantum-mechanical calculations of all the finite systems throughout this work. Thus, our calculations consider explicitly the chemistry of the attachment of the molecule to the contacts instead of unrealistic simulations of a molecule attached to perfect or ideal surfaces.

For a hypothetical molecular system that has only five basis functions (χ), the elements of the Kohn–Sham Hamiltonian matrix ($\mathbf{H}^{\text{KS}}$) are given by

$$\mathbf{H}^{\text{KS}} = \begin{pmatrix} \langle\chi_1|\hat{h}^{\text{KS}}|\chi_1\rangle & \langle\chi_1|\hat{h}^{\text{KS}}|\chi_2\rangle & \langle\chi_1|\hat{h}^{\text{KS}}|\chi_3\rangle & \langle\chi_1|\hat{h}^{\text{KS}}|\chi_4\rangle & \langle\chi_1|\hat{h}^{\text{KS}}|\chi_5\rangle \\ \langle\chi_2|\hat{h}^{\text{KS}}|\chi_1\rangle & \langle\chi_2|\hat{h}^{\text{KS}}|\chi_2\rangle & \langle\chi_2|\hat{h}^{\text{KS}}|\chi_3\rangle & \langle\chi_2|\hat{h}^{\text{KS}}|\chi_4\rangle & \langle\chi_2|\hat{h}^{\text{KS}}|\chi_5\rangle \\ \langle\chi_3|\hat{h}^{\text{KS}}|\chi_1\rangle & \langle\chi_3|\hat{h}^{\text{KS}}|\chi_2\rangle & \langle\chi_3|\hat{h}^{\text{KS}}|\chi_3\rangle & \langle\chi_3|\hat{h}^{\text{KS}}|\chi_4\rangle & \langle\chi_3|\hat{h}^{\text{KS}}|\chi_5\rangle \\ \langle\chi_4|\hat{h}^{\text{KS}}|\chi_1\rangle & \langle\chi_4|\hat{h}^{\text{KS}}|\chi_2\rangle & \langle\chi_4|\hat{h}^{\text{KS}}|\chi_3\rangle & \langle\chi_4|\hat{h}^{\text{KS}}|\chi_4\rangle & \langle\chi_4|\hat{h}^{\text{KS}}|\chi_5\rangle \\ \langle\chi_5|\hat{h}^{\text{KS}}|\chi_1\rangle & \langle\chi_5|\hat{h}^{\text{KS}}|\chi_2\rangle & \langle\chi_5|\hat{h}^{\text{KS}}|\chi_3\rangle & \langle\chi_5|\hat{h}^{\text{KS}}|\chi_4\rangle & \langle\chi_5|\hat{h}^{\text{KS}}|\chi_5\rangle \end{pmatrix} \quad (22)$$

The atoms of the molecular system can be classified as belonging to the contact 1, the contact 2, or the molecule (M). For illustration, the atoms conforming the contact 1, the contact 2, and the molecule are modeled by the χ_2; χ_3, χ_4; and χ_1, χ_5 basis functions, respectively. After reordering and partitioning $\mathbf{H}^{\text{KS}}$ into submatrices we have:

$$\mathbf{H}_{11} = (\langle\chi_2|\hat{h}^{\text{KS}}|\chi_2\rangle) \quad (23)$$

$$\mathbf{H}_{22} = \begin{pmatrix} \langle\chi_3|\hat{h}^{\text{KS}}|\chi_3\rangle & \langle\chi_3|\hat{h}^{\text{KS}}|\chi_4\rangle \\ \langle\chi_4|\hat{h}^{\text{KS}}|\chi_3\rangle & \langle\chi_4|\hat{h}^{\text{KS}}|\chi_4\rangle \end{pmatrix} \quad (24)$$

$$\mathbf{H}_{MM} = \begin{pmatrix} \langle \chi_1 | \hat{h}^{\mathrm{KS}} | \chi_1 \rangle & \langle \chi_1 | \hat{h}^{\mathrm{KS}} | \chi_5 \rangle \\ \langle \chi_5 | \hat{h}^{\mathrm{KS}} | \chi_1 \rangle & \langle \chi_5 | \hat{h}^{\mathrm{KS}} | \chi_5 \rangle \end{pmatrix} \tag{25}$$

$$\mathbf{H}_{1M} = \left(\langle \chi_2 | \hat{h}^{\mathrm{KS}} | \chi_1 \rangle \quad \langle \chi_2 | \hat{h}^{\mathrm{KS}} | \chi_5 \rangle \right) \tag{26}$$

$$\mathbf{H}_{M1} = \begin{pmatrix} \langle \chi_1 | \hat{h}^{\mathrm{KS}} | \chi_2 \rangle \\ \langle \chi_5 | \hat{h}^{\mathrm{KS}} | \chi_2 \rangle \end{pmatrix} \tag{27}$$

$$\mathbf{H}_{2M} = \begin{pmatrix} \langle \chi_3 | \hat{h}^{\mathrm{KS}} | \chi_1 \rangle & \langle \chi_3 | \hat{h}^{\mathrm{KS}} | \chi_5 \rangle \\ \langle \chi_4 | \hat{h}^{\mathrm{KS}} | \chi_1 \rangle & \langle \chi_4 | \hat{h}^{\mathrm{KS}} | \chi_5 \rangle \end{pmatrix} \tag{28}$$

$$\mathbf{H}_{M2} = \begin{pmatrix} \langle \chi_1 | \hat{h}^{\mathrm{KS}} | \chi_3 \rangle & \langle \chi_1 | \hat{h}^{\mathrm{KS}} | \chi_4 \rangle \\ \langle \chi_5 | \hat{h}^{\mathrm{KS}} | \chi_3 \rangle & \langle \chi_5 | \hat{h}^{\mathrm{KS}} | \chi_4 \rangle \end{pmatrix} \tag{29}$$

where $\mathbf{H}_{MM}$ is the submatrix representing the isolated molecule (restricted molecule). The other submatrix represents the couplings between the molecule (subscript M) and the atoms of the contact (subscripts 1 and 2).

Then, we create an ordered Hamiltonian matrix (**H**) and the respective overlap matrix (**S**) in the following way:

$$\begin{gathered} \mathbf{H} = \begin{pmatrix} \mathbf{H}_{11} & \mathbf{H}_{1M} & \mathbf{H}_{12} \\ \mathbf{H}_{M1} & \mathbf{H}_{MM} & \mathbf{H}_{M2} \\ \mathbf{H}_{21} & \mathbf{H}_{2M} & \mathbf{H}_{22} \end{pmatrix} \\ \mathbf{S} = \begin{pmatrix} \mathbf{S}_{11} & \mathbf{S}_{1M} & \mathbf{S}_{12} \\ \mathbf{S}_{M1} & \mathbf{S}_{MM} & \mathbf{S}_{M2} \\ \mathbf{S}_{21} & \mathbf{S}_{2M} & \mathbf{S}_{22} \end{pmatrix} \end{gathered} \tag{30}$$

This Hamiltonian matrix for the extended molecule (H) is also recalculated as the bias electrical field is applied to the junction in order to account for the reorganization of the molecular electronic structure due to the presence of such field. This allows us to study among others the effects of the external bias potential on charge transfer between the molecule and the contacts, the shift of molecular levels and the shape changes of the molecular orbitals, which have a direct effect on the conductance of the junction. These effects are needed to explain the nonlocal behavior of molecular systems presenting highly nonlinear features such as rectification, negative differential resistance, memory hysteresis, etc. Notice that the molecule itself does not have an integer charge in any of the charge states of the extended molecule because the charge distributes between the isolated molecule and the metal atoms. Charge transfers between the molecule and contact occur even at zero bias voltage and also as a result of an externally applied field. Certainly, this charge transfer is determined by the metal atoms attached to the molecule; these metal atoms together with the continuum define specific tip. It is clearly demonstrated from theoretical as well as experimental information [3, 14, 38] that the connection of the molecule to the metal is only through one or two metal atoms as concluded in [3]. However, the effect of local interactions with the atoms located beyond these nearest neighbors on the actual molecule is very small and usually truncated; this constitutes the strongest approximation of our procedure only if the molecule were realistically connected directly to a continuum. Fortunately, there is strong evidence that it is an acceptable approximation because it precisely considers the chemistry and physics of the actual local attachment or bonding of the molecule to the surface atoms [3, 14]. Methods such as the so-called "non-equilibrium" for instance are shown to include only the Hartree response of the system, thus missing important physics of the problem [39].

The coupling between atoms of the contact and those of the molecule yields the self-energy term, $\mathbf{\Sigma}_j$:

$$\mathbf{\Sigma}_j = \mathbf{H}_{Mj} g_j \mathbf{H}_{jM} \qquad j = 1,\ 2 \tag{31}$$

which depends on the complex Green function, g_j, describing the contact j. The complex g_i can be obtained from any source as long as it can be represented in matrix form of the appropriate dimensions; it provides the information from the contact to the DFT-GF formalism. We choose to generate the Green function for the conctacts using Crystal 03 since it allows obtaining a high-level electronic structure of a bulk system of any shape using DFT. This complex function is defined as:

$$\mathbf{g}_j(E) = -\pi\sqrt{-1}\begin{pmatrix} g_j^1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & g_j^{n_j} \end{pmatrix} \qquad j = 1,\ 2 \tag{32}$$

where each value of the diagonal matrix is proportional to the local density of states DOS, which has been calculated in Section 2.2.

$$\mathbf{g}_j^k(E) = \begin{pmatrix} \mathrm{DOS}_j s^k(E) & 0 & 0 & 0 \\ 0 & \mathrm{DOS}_j p^k(E) & 0 & 0 \\ 0 & 0 & \mathrm{DOS}_j d^k_{t_{2g}}(E) & 0 \\ 0 & 0 & 0 & \mathrm{DOS}_j d^k_{e_g}(E) \end{pmatrix} \tag{33}$$

In order to keep consistency in the matrix dimensions of Eq. (32), the index k runs over all the interfacial atoms that represent contact $j(k = 1, \ldots, n_j)$. Each diagonal term of Eq. (33) is again another diagonal matrix, in such a way that the size of DOS $_j s^k(E)$ is equal to the number of s-type basis functions used to model the electronic structure of the type of atom that composes contact j.

The coupling of the molecule to the contacts is obtained from molecular calculations (H_{iM} and H_{Mi} shown in Eq. (30)) that consider the atomistic nature of the contact–molecule interface. The interaction terms defined in Eq. (31) are added to the molecular Hamiltonian to account for the effect of the contact on the molecule:

$$\mathbf{H}_e = \begin{pmatrix} \mathbf{H}_{11} & \mathbf{H}_{1M} & \mathbf{H}_{12} \\ \mathbf{H}_{M1} & \mathbf{H}_{MM} + \Sigma_1 + \Sigma_2 & \mathbf{H}_{M2} \\ \mathbf{H}_{21} & \mathbf{H}_{2M} & \mathbf{H}_{22} \end{pmatrix} \tag{34}$$

To account for the non-orthogonality of the basis set, the overlap matrix S modifies the Hamiltonian into:

$$\mathbf{H}'_e = \mathbf{S}^{-1}\mathbf{H}_e = \begin{pmatrix} \mathbf{H}'_{11} & \mathbf{H}'_{1M} & \mathbf{H}'_{12} \\ \mathbf{H}'_{M1} & \mathbf{H}'_{MM} & \mathbf{H}'_{M2} \\ \mathbf{H}'_{21} & \mathbf{H}'_{2M} & \mathbf{H}'_{22} \end{pmatrix} \tag{35}$$

This modified Hamiltonian is used to obtain the Green function for a molecule attached to two contact tips:

$$\mathbf{G}_M(\mathbf{E}) = (\mathbf{E}1 - \mathbf{H}'_{MM})^{-1} \tag{36}$$

Finally, the DOS of the molecule subjected to the effect of the two contacts is calculated as:

$$\mathrm{DOS} = \frac{\sqrt{-1}}{2\pi} Trace(\mathbf{G}_M - \mathbf{G}_M^{\dagger}) \tag{37}$$

Within the Green function formalism, two separated and independent calculations are needed. First is molecular calculations on the molecule of interest plus a few atoms like those in the contact. Second is a calculation of the DOS of each contact; those calculations can be performed at any level of theory; however, it is desirable to choose *ab initio* methods known to provide chemical accuracy, such as DFT using generalized gradient approximation or better.

3. Electron transport in molecular junctions

We model our molecular system as a generic two-port network, shown in Figure 6. The bias voltage is defined as: $V = V_1 - V_2$. Thus, contact 1 is considered as the positive electrode and contact 2 the negative one. At contact 1, we define i_1^- as the current flowing from contact 1 towards the molecule and i_1^+ as the backscattered current, which flows from the molecule to the contact. Likewise, at contact 2, we have i_2^+ flowing from contact 2 to the molecule, and i_2^- flowing from the molecule to contact 2. For a detailed description of the original procedure the reader may refer to [13] and references therein.

The associated scattering matrix for such two-port network is:

$$\begin{pmatrix} i_1^+ \\ i_2^- \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} i_1^- \\ i_2^+ \end{pmatrix} \tag{38}$$

where the elements of the scattering matrix are defined as:

$$S_{21} = \left.\frac{i_2^-}{i_1^-}\right|_{i_2^+=0} \tag{39}$$

$$S_{12} = \left.\frac{i_1^+}{i_2^+}\right|_{i_1^-=0} \tag{40}$$

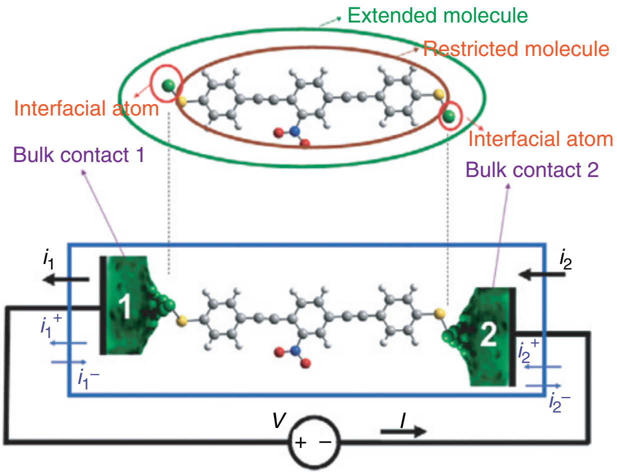


Figure 6 Terminology used in our electron transport calculations. The *bulk* contacts are pictorial representations of the two macroscopic tips that approach the molecule. A *restricted* molecule corresponds to the model under study itself; it includes the *alligator* atoms, such as sulfur, if they are present. *Interfacial* atoms correspond to atoms of the type belonging to the bulk contacts. The *extended* molecule is composed of the restricted molecule and some atoms of the type belonging to each contact material. Arrows show the convention used for the direction of the currents and the polarity of the bias voltage

$$S_{22} = \left.\frac{i_2^-}{i_2^+}\right|_{i_1^-=0} \tag{41}$$

$$S_{11} = \left.\frac{i_1^+}{i_1^-}\right|_{i_2^+=0} \tag{42}$$

From Eq. (39), s_{21} is interpreted as the number of electrons that can reach contact 2 (considering contact 2 as reflectionless, $i_2^+ = 0$), per each electron that is injected from contact 1. In other words, it is the probability for an electron to cross the molecular junction from contact 1 to contact 2. Analogously, s_{12} represents the probability for an electron to cross the junction from contact 2 to contact 1. At equilibrium, the probability for a particle to tunnel through a barrier would be the same whether it crosses the barrier from left to right or from right to left. We define this quantity as the “transmission probability”, $\overline{T}$.

$$S_{21} = S_{12} = \overline{T} \tag{43}$$

From Eq. (41), s_{22} is the number of backscattered electrons per each electron that goes through contact 1, considering no reflection at contact 2. Then it is the probability for an electron injected through port 2 to be reflected, which is the complement of the transmission probability:

$$S_{11} = S_{22} = 1 - \overline{T} \tag{44}$$

Then, Eq. (38) becomes

$$\begin{pmatrix} i_1^+ \\ i_2^- \end{pmatrix} = \begin{pmatrix} 1-\overline{T} & \overline{T} \\ \overline{T} & 1-\overline{T} \end{pmatrix} \begin{pmatrix} i_1^- \\ i_2^+ \end{pmatrix} \tag{45}$$

Equation (45) ensures the conservation of total current in the two-port network, i.e., $I = i_1 = i_2$, where:

$$i_1 = i_1^+ - i_1^-$$
$$i_2 = i_2^+ - i_2^-$$

At a given energy E, the current per mode per unit energy (as a result of an occupied state in one contact leaking into the molecule) is given by $2e/h$[*]. For a partially occupied state, such current needs to be corrected by the Fermi distribution factor (f) of the contact. The total current leaking from contact 1 into the molecule is given by:

$$i_1^-(E) = \frac{2e}{h} M(E) f_1(E) \mathrm{d}E \tag{46}$$

where $M(E)$ is the number of transmission modes allowed for the molecule at the energy E. Analogously, the amount of total current leaking from contact 2 into the molecule before reaching equilibrium is:

$$i_2^+(E) = \frac{2e}{h} M(E) f_2(E) \mathrm{d}E \tag{47}$$

When a small bias voltage $(V \neq 0)$ is applied between the contacts of the junction, the molecular system is taken out of equilibrium and the electrons flow. The application of a positive bias voltage between the contacts shifts down the Fermi level of contact 1 and shifts up the Fermi level of contact 2. In both cases, the shifts are by an equal amount of $0.5eV$ with respect to the equilibrium Fermi level of the extended molecule (μ_{EM}) [14], in the following way

$$\mu_2^* = \mu_{\mathrm{EM}} + \frac{1}{2} eV \tag{48}$$

$$\mu_1^* = \mu_{\mathrm{EM}} - \frac{1}{2} eV \tag{49}$$

Consequently, the Fermi distribution functions of both contacts are shifted whenever a bias voltage (V) is applied to the junction; this makes the Fermi distributions dependent on the applied bias voltage.

$$f_2\left(E - \mu_{\mathrm{EM}} - \frac{1}{2} eV\right) = \frac{1}{1 + e^{\frac{E-\mu_2^*}{kT}}} \tag{50}$$

$$f_1\left(E - \mu_{\mathrm{EM}} + \frac{1}{2} eV\right) = \frac{1}{1 + e^{\frac{E-\mu_1^*.}{kT}}} \tag{51}$$

---

[*] e refers to the charge of a proton $+1.602177 \times 10^{-19}$.

Combining Eqs. (45), (46), (47), (50), and (51), we obtain:

$$i(E,V)=\frac{2e}{h}M(E)\overline{T}(E)\left[f_2\left(E-\mu_{\mathrm{EM}}-\frac{1}{2}eV\right)-f_1\left(E-\mu_{\mathrm{EM}}+\frac{1}{2}eV\right)\right]\mathrm{d}E \tag{52}$$

Defining the transmission function as $T(E)=M(E)\overline{T}(E)$ and integrating over energy, the total current of electrons flowing between the contacts is:

$$I(V)=\frac{2e}{h}\int_{-\infty}^{+\infty}T(E,V)\left[f_2\left(E-\mu_{\mathrm{EM}}-\frac{1}{2}eV\right)-f_1\left(E-\mu_{\mathrm{EM}}+\frac{1}{2}eV\right)\right]\mathrm{d}E \tag{53}$$

The transmission function, T, is obtained from the chemistry of the molecular junction. It is defined as [40]:

$$T(E,\ V)=\frac{1}{N}Trace(\mathbf{\Gamma}_2\mathbf{G}_M\mathbf{\Gamma}_1\mathbf{G}_M^{\dagger}) \tag{54}$$

where N is the number of basis functions used to represent the restricted molecule and V is the bias voltage applied between the contacts. The consideration of the bias voltage affecting all the matrices equation (54) was of paramount importance to convert this early mesoscopic procedure into a molecular one [41]. The coupling ($\mathbf{\Gamma}_j$) between the molecule and the contact j is defined as:

$$\mathbf{\Gamma}_j=\sqrt{-1}(\mathbf{\Sigma}_j-\mathbf{\Sigma}_j^{\dagger}) \qquad j=1,\ 2 \tag{55}$$

where the self-energy term, Σ_j, Eq. (31), depends on the Green function, of the contacts. The Green function, $g(E)$, Eq. (32), depends on the Fermi level of the contact, which varies with the applied voltage according to Eqs. (48) and (49). Consequently, the Green function of each contact, the self-energy terms $\mathbf{\Sigma}_j$, the coupling terms $\mathbf{\Gamma}_i$, and the transmission function are a function of the applied voltage, i.e, $T(E,\ V)$.

4. Metal–molecule–metal junctions

4.1. Metal–benzene–metal junction

We aim to study the conductance of the nitroOPE molecule, which is composed of three benzene rings, attached to metallic CNT tips. We start the analysis with a simpler case, a single benzene molecule between two CNT tips (CNT–benzene–CNT junction). The Au–S–benzene–S–Au junction has been studied before [36, 42]; in those calculations, the adsorption of benzene to the gold contacts is possible by use of a sulfur atom connecting a carbon and a gold atom (thiol bond). Recent research has shown the possibility of direct attachment of benzene to carbon nanotubes [43–46]; this opens the possibility of employing metallic CNTs as contacts to organic molecules.

The first step in simulating the benzene connected to two infinitely long CNT contacts is the inclusion of interfacial carbon atoms, representing the CNT contacts, in the extended molecule (see terminology in Figure 6). It is known that an infinitely long (4, 4) CNT shows metallic behavior [28], but small pieces of (4, 4) CNT need not necessarily show a metallic character. Therefore, the CNT has to be modeled by an adequate number of atoms such that metallic behavior is reached. The second step is to include the effect of the continuum of electronic states provided by the infinitely long nature of the (4, 4) CNT contacts; this is accomplished by the use of the DFT-GF approach described in Section 2.3.

We test several junctions in which each CNT contact is modeled by 40, 48, 56, 64, 72, and 80 carbon atoms, corresponding to parts A, B, C, D, E, and F of Figure 7, respectively. The DOS for the (4, 4) CNT, which is shown in Figure 5, and the electronic structures of all the molecular junctions are calculated using the B3PW91 DFT method combined with the 6-31G basis set. The calculation of the I-Vs (Figure 8) shows that all the junctions present consistently similar values of current, indicating that even 40 carbon atoms are suffice to model each CNT contact. Moreover, in a previous work [47], we demonstrated that small pieces of CNT, composed of 80 atoms, did behave as expected for their infinitely long counterparts, i.e., metallic character for the (4, 4) and the (9, 0), and semiconducting character for the (8, 0) CNT.

All the junctions show ohmic behavior, with a constant resistance of $\sim 2\,\mathrm{M}\Omega$, for small bias voltages ($< \sim 3\,\mathrm{V}$). The ohmic behavior at low bias voltages agrees with the theoretical calculations reported by Derosa [36] and Di Ventra [42].

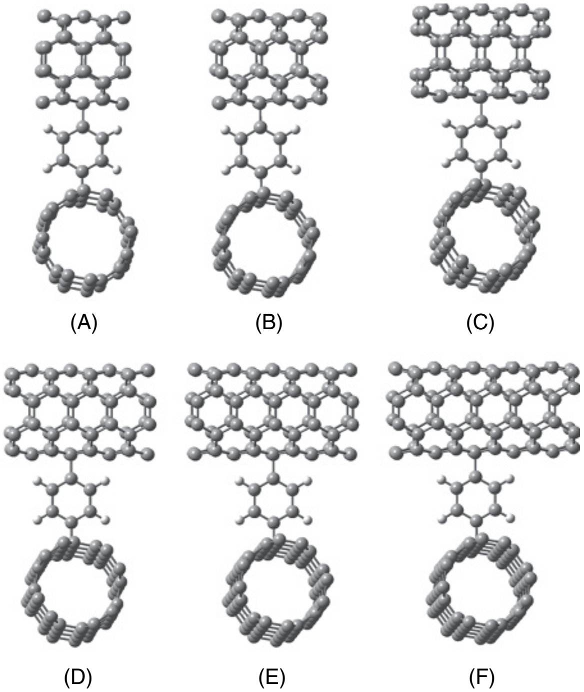


Figure 7 Molecular junctions of the type metal-benzene-metal. The pieces of (4, 4) CNTs (metal) are shown above and below the benzene. A ring of the metallic CNT is defined to be composed of 8 carbon atoms. The total number of atoms belonging to the top and bottom CNTs is increased progressively, both contacts are constructed to have the same number of carbon atoms. (A) is composed of 5 rings in the top and also 5 rings in the bottom contact, (B) of 6, (C) of 7, (D) of 8, (E) of 9, and (F) of 10

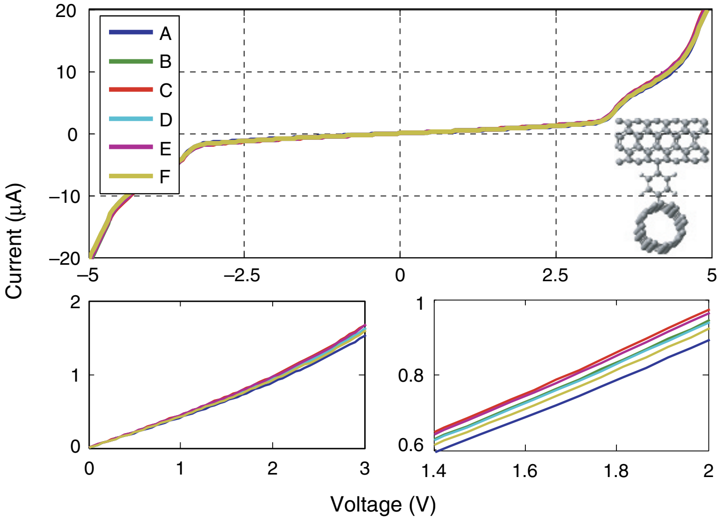


Figure 8 Current–voltage characteristics for six junctions of the form CNT–benzene–CNT (shown in Figure 7). In each junction a different number of carbon atoms is used to model the CNT contact (40, 48, 56, 64, 72, and 80 carbon atoms corresponding to the junctions A, B, C, D, E, and F, respectively). The inset shows the geometry of the junction F. The plots on the bottom part are two amplifications of the ohmic region

Gold has more electrons available for conduction per atom than the metallic (4, 4) CNT; ~10 times higher at their Fermi levels as shown in Figures 1 and 5, which in principle should make the Au–S–benzene–S–Au junction more conducting that the CNT–benzene–CNT junction. However, the current in the CNT–benzene–CNT junction, at 2 V, is found to be higher than the values reported theoretically [36] and experimentally [6] for the Au–S–benzene–S–Au junction. We attribute that higher current to the better (seamless) chemical bond between the benzene and the CNT than the Au–S–benzene bond.

4.2. Metal–nitroOPE–metal junction

We calculate junctions containing the nitroOPE molecule under metallic contacts such as Au and the (4, 4) CNT. These results are considered as references for subsequent calculations, which include semiconducting contacts.

Gold has been the preferred contact material for the experiments on molecular conduction either as a vapor-deposited top contact, such as in a nanopore device, or as the tip of an STM [48]. Here, we study two cases in which the nitroOPE is bonded to gold contacts, the Au_6–nitroOPE–S–Au_1 and the Au_1–S–nitroOPE–S–Au_1 junction.

In the Au_6–nitroOPE–S–Au_1 junction, the bottom contact is modeled by one interfacial gold atom. The nitroOPE is bound to the gold atoms by a thiol bond (C-S-Au). The top contact is modeled by six interfacial gold atoms that are not chemically bonded to the nitroOPE. This type of physical bond is expected to be found in experimental measurements of molecular I-V that use an STM tip as top contact. The geometry of the extended molecule is shown in the lower right corner of Figure 9B. The neutral, the

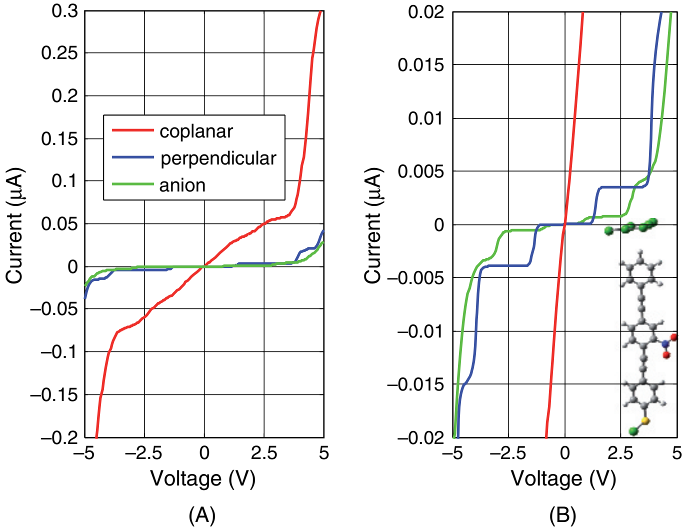


Figure 9 (A) Current–voltage characteristic for the Au_6–nitroOPE–S–Au_1 junction. (B) Amplification of the low-current region of (A). The coplanar conformation of the molecular junction is shown in the lower part of (B). The C, H, S, N, O, and Au atoms are colored grey, white, yellow, blue, red, and green, respectively

first charge state (anion), and the perpendicular conformational state are calculated for this Au_6–nitroOPE–S–Au_1 junction.

Also, two different and possible geometrical conformations are calculated. In the coplanar conformation, the three phenyl rings in the nitroOPE are lying in the same plane; however, in the perpendicular conformation, the middle phenyl ring is perpendicular to the other two. The calculation establishes the coplanar conformation as more stable than the perpendicular conformation, with a rotational barrier of −0.20 eV (−4.7 kcal/mol) for the middle phenyl ring. The current–voltage calculations for the coplanar, perpendicular, and anion states are reported in Figure 9A.

In the Au_1–S–nitroOPE–S–Au_1 junction, one gold atom is used to represent each contact. The attachment of the nitroOPE molecule to both gold atoms is through thiol bonds. The geometry of this junction is shown in the lower right corner of Figure 10A. The current–voltage characteristic for the coplanar, perpendicular, and anion states for these junctions are shown in Figure 10.

For both junctions, the Au_6–nitroOPE–S–Au_1 and the Au_1–S–nitroOPE–S–Au_1, two distinct states of conductance are observed, high conductance (red curve) and low conductance (green and blue curves). The neutral molecule (charge = 0) presents high conductance whereas the anion (charge = −1) and the perpendicular state show low conductance. Moreover, the high-conductance state of the junctions shows ohmic behavior at low bias voltage, which agrees with previous results reported for similar molecules [15, 16, 37].

The Au_1–S–nitroOPE–S–Au_1 junction allows for significantly higher current (~5 times) than the Au_6–nitroOPE–S–Au_1. The physical bond, present between the nitroOPE and the six-gold plane, is a gap of atomistic size that obstructs the flow of electrons. It is

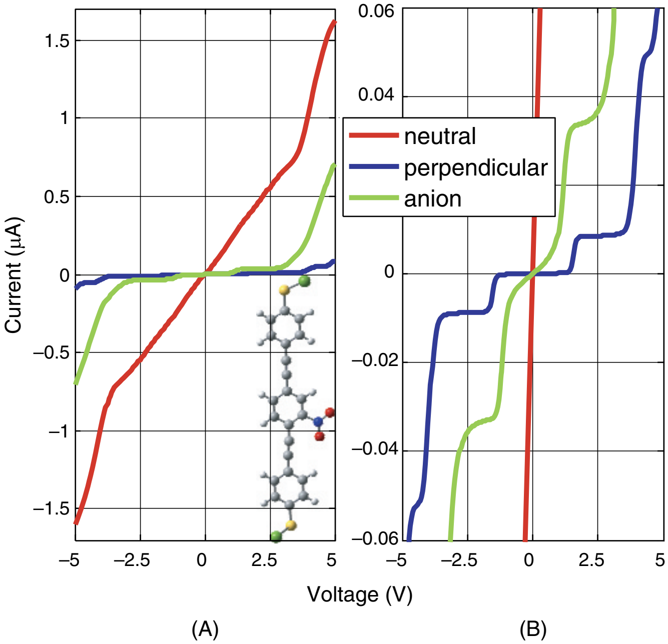


Figure 10 (A) Current–voltage characteristic for the nitroOPE under two gold tips (green). Sulfur atoms (yellow) have been included too. (B) Amplification of (A). The coplanar conformation of the molecular junction is shown in the lower right corner of (A)

effectively a thin tunneling barrier for the electrons to overcome. The thiol bond in the top contact of the Au_1–S–nitroOPE–S–Au_1 junction allows more transfer of electrons than the physical bond in the top contact of the Au_6–nitroOPE–S–Au_1 junction. In this regard, Cui *et al.* have experimentally demonstrated [49] a difference of four orders of magnitude between the current in a chemisorbed junction ("glued" by covalent bonds) and the current in a physisorbed junction ("glued" by physical bonds).

We study metallic CNTs as prospective contacts for molecular junctions. Eighty carbon atoms are used to model a piece of the (4, 4) CNT. The geometry of the coplanar conformation of the CNT–nitroOPE–CNT junction is shown in the lower right corner of Figure 11. The calculated current–voltage characteristics for the coplanar, perpendicular, and anion states are reported in Figure 11.

Similar to the case when having gold contacts, the two distinct states of conductance attributed to the nitroOPE molecule are still found for the CNT–nitroOPE–CNT junction. The coplanar conformation (red) exhibits high conductance whereas the perpendicular and anion states (blue and light green respectively) exhibit low conductance.

Despite the fact that gold has more electrons per atom available for conduction than the metallic (4, 4) CNT has, the CNT–nitroOPE–CNT allows higher current than the Au_6–nitroOPE–S–Au_1 junction. This is a consequence of the tunneling gap around the top contact of the Au_6–nitroOPE–S–Au_1 junction, which obstructs the flow of electrons.

Although thiol bonds are chemically easy to work with, they present a disadvantage from the electrical point of view. Thiol bonds are highly polar, and polar bonds introduce

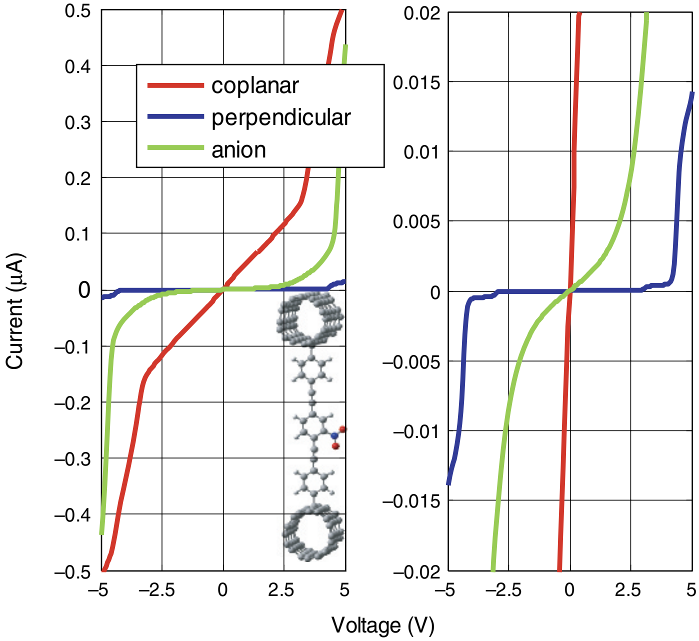


Figure 11 Left: Current–voltage characteristic for the coplanar, perpendicular, and anion states of the CNT–nitroOPE–CNT junction. The coplanar conformation of the molecular junction is shown in the lower right corner. Right: Amplification of the low-current region

undesirable capacitive effects that restrict the flow of electrons. Vondrak *et al.* used two-photon photoelectron spectroscopy to show that the S atom, in C-S-Cu thiol bonds, acts as insulators, obstructing the flow of electrons [50]. Thiol bonds should be considered as thin tunneling barriers. The Au crystal has ~10 times higher density of states than the (4, 4) CNT crystal does; however, the Au–S–nitroOPE–S–Au junction exhibits only ~3 times higher current than the CNT–nitroOPE–CNT junction does. This is an indication that the C-C bonds in the CNT–nitroOPE–CNT are electrically superior to the thiol bonds in the Au_1–S–nitroOPE–S–Au_1 junction.

From another point of view, the Au_1–S–nitroOPE–S–Au_1 and the Au_6–nitroOPE–S–Au_1 junctions can be thought of as a nitroOPE isolated by two thin tunneling barriers at each end, resembling the particle-in-a-box problem. The quantum confinement preserves the discrete nature of the molecular electronic states. The DOS of the perpendicular Au_1–S–nitroOPE–S–Au_1 junction shows the presence of an isolated and narrow peak in the proximity (a channel for conduction) of its Fermi level. The junction is not conducting until enough bias voltage (energy) is applied to reach the energy of that channel; electron transport takes place by resonant tunneling using that isolated channel. Moreover, the current does not change with the increase of voltage until another molecular channel is reached. This phenomenon is reflected in the steplike shape of the I–V curve (Figure 9B). The perpendicular Au_1–S–nitroOPE–S–Au_1 junction shows steplike I–V characteristic (Figure 10B), too. The steplike variation of the current has also been experimentally observed in molecular systems [51, 52].

In summary, metal–nitroOPE–metal junctions are found to have isolated and narrow DOS peaks, which are reflected in steplike I–V curve, whenever they meet two conditions: first, they are in a state of low-conductance (perpendicular conformation or anion); and second, they contain tunneling barriers (physical or thiol bonds). Junctions in states of high conductance (coplanar conformations) and junctions that do not contain tunneling barriers (CNT–nitroOPE–CNT) do not show steplike I–V curve.

The transport of current through a molecular junction comprises the study of a molecular system that presents both a finite and an infinite character. The finite part (single molecule) is calculated precisely from the fundamental Schrödinger equation. The effect of macroscopic contacts (infinite part) is included following the DFT-GF approach. At the scale of the molecular junctions considered in this work, the transport of electrons is described by the Landauer formalism.

A DFT-GF implementation of the Landauer formalism is used to calculate the I–V of metal–nitroOPE–metal junctions in different conformational and charge states. Gold and the (4, 4) CNT are tested as metallic contacts, and in both cases the metal–nitroOPE–metal junction presents high conductance when the nitroOPE is in its coplanar conformation. The calculations predict low conductance for the perpendicular conformation and for the charge states (anion, dianion, trianion) of the nitroOPE. It is observed that the states of high conductance exhibit ohmic I–V at low bias voltage.

The CNT–nitroOPE–CNT junction has values of current similar to the junctions containing gold contacts, despite the fact that CNT has $\sim$10 times lower DOS than gold. This result encourages the use of CNT as an alternative to gold in molecular devices; however, technological challenges remain regarding the manipulations of single CNTs. The rationale for the high conductance of the junction containing CNT is the direct C$-$C bond between the CNT and the nitroOPE; instead, the thiol bonds (Au–S–C) in the Au_1–S–nitroOPE–S–Au_1 junction behave as undesired interfacial capacitors at the interfaces, isolating the nitroOPE from the contacts. Moreover, the calculation shows that the gold atoms at the top contact of the Au_6–CNT–S–Au_1 junction form a physical bond with the nitroOPE. The physical bond is effectively a tunneling gap, which deters even more the flow of electrons. For the Au_6–CNT–S–Au_1 junction, the current is lower than for the CNT–nitroOPE–CNT junction.

5. Metal–molecule–semiconductor junctions

The semiconductor industry entered the nanometer regime (<100 nm) in 2000 and continues today to be in the race for miniaturization. The first commercial single molecule–based device is most likely to be built around Si.

At sizes approaching the quantum-confinement regime, the electrical properties of silicon, and any other material, diverge from the bulk properties. For example, studies have shown the increase of the bandgap with the decrease of the size of the semiconducting nanostructure [53–55]. For silicon nanowires (SiNWs), theoretical calculations have shown that the quantum effects are substantial at diameters below 3 nm [56–61]. Quantum-mechanical calculations of the type presented in this work are necessary for devices containing Si nanostructures in the quantum-confinement regimen.

In the previous section, we have described the distinctive impedance states of the metal–nitroOPE–metal junctions. Advances in synthetic chemistry have allowed the direct attachment of organic molecules on Si substrates [62, 63], opening the door for hybrid organic-semiconducting devices. In this section, we consider the effect of Si contacts on the bistable properties of the nitroOPE.

A Schottky diode, which is formed when a metal and a semiconductor are in intimate contact, acts as a current rectifier. Therefore, in a macroscopic metal–device–semiconductor junction, the simultaneous use of a semiconducting and a metallic contact implies a tremendous change in the properties of the device. In other words, the electrical behavior of the device may be overruled by the rectifying behavior of the contacts. The challenge is to use Si as one of the contacts in metal–nitroOPE–Si molecular junctions without destroying the bistable characteristics attributed to the nitroOPE molecule. The rectifying behavior has been experimentally observed to vanish as the size of the metal–semiconductor junction approaches the nanometer regime: i.e., ultrasmall Schottky diodes [64–66]. This gives hope for using Si as a righteous contact material in single molecule–based electronic devices; we perform quantum-mechanical calculations to assess the ability of metal–nitroOPE–Si junctions to keep the high- and low-impedance states found in metal–nitroOPE–metal junctions. Our study considers the different charge states (neutral, anion, dianion, and trianion) as well as the coplanar and perpendicular conformations of the nitroOPE molecule. Both gold and (4, 4) CNT are tested as metallic contacts.

5.1. Significance of the electronic chemical potential (Fermi level) for a single molecule

The electrochemical potential is a property traditionally defined, for macroscopic systems, as the variation of the total energy with respect to the number of particles in the ensemble. This concept needs to be extended to be able to determine the Fermi level of a single molecule.

The Fermi level for a molecule is synonymous with minus the electronegativity, which is defined as the average of the ionization potential (IP) and the electron affinity (EA) (Mulliken electronegativy):

$$\mu = -\frac{\mathrm{IP}+\mathrm{EA}}{2} \tag{56}$$

where the electron affinity (EA) is defined as the amount of energy needed by the molecule (or atom), in its neutral state, to accept an extra electron. The ionization potential (IP), also called ionization energy, is the energy needed to strip out one electron from the molecule (or atom). EAs and IPs can be calculated computationally as the difference between the self-consistent field (SCF) energies of the charge states of the molecule.

$$\mathrm{EA} = E_{\mathrm{anion}} - E_{\mathrm{neutral}}$$

$$\mathrm{IP} = E_{\mathrm{cation}} - E_{\mathrm{neutral}}$$

This approach is called ΔSCF; recent studies show that DFT methods are able to achieve 0.1–0.2 eV of accuracy to calculate EAs and IPs [67–69].

A more direct approach to calculate the molecular Fermi level is based on a quantum-mechanical extension of the traditional definition of chemical potential [70–72].

According to the first Hohenberg–Kohn theorem [22], the ground-state energy (E_0) is a functional of the density, $\rho_0(r)$:

$$E_0 = E_\nu[\rho_0(r)] = \int \rho_0(r)\nu(r) + F[\rho_0(r)] \tag{57}$$

where $\rho_0(r)$ is the ground-state electron density. For any other trial electron density, $\rho(r)$, we get another energy:

$$E' = E_\nu[\rho(r)] = \int \rho(r)\nu(r) + F[\rho(r)] \tag{58}$$

The second Hohenberg–Kohn theorem establishes that the energy of Eq. (58) cannot be lower than the energy in Eq. (57). In other words, the minimization of Eq. (58) with respect to variations of the electron density, $\rho(r)$, yields the ground-state energy (E_0) and ground-state electron density, $\rho_0(r)$

$$\int \rho(r)\mathrm{d}\tau = N \tag{59}$$

with N being the number of electrons in the molecular system. This minimization problem is commonly tackled by the introduction of a Lagrange multiplier μ

$$\delta\left\{E_\nu[\rho(r)] - \mu\int \rho(r)\mathrm{d}\tau\right\} \tag{60}$$

which results in the following Euler–Lagrange equations:

$$\mu = \left[\frac{\delta E[\rho(r)]}{\delta\rho(r)}\right]_{\rho=\rho_0} \tag{61}$$

$$\mu = \frac{\partial E}{\partial N} \tag{62}$$

The similarity of Eq. (62) to the traditional, thermodynamical definition of chemical potential in macroscopic systems reassures that the chosen Lagrange multiplier constant, μ, corresponds indeed with the Fermi level [72]. Moreover, Eq. (62) gives a direct relationship of the Fermi level with the total energy functional and the ground-state electron density.

Combining Eqs. (56) and (61), we obtain the following relation:

$$\mu = \left[\frac{\delta E[\rho(r)]}{\delta\rho(r)}\right]_{\rho=\rho_0} = -\frac{\mathrm{IP}+\mathrm{EA}}{2} \tag{63}$$

Based of Eq. (61), Perdew *et al.* [70, 71] have shown that the ionization potential (IP) is exactly minus the energy of the highest occupied Kohn–Sham molecular orbital energy ($\varepsilon_{\mathrm{HOMO}}$), for the exact energy functional.

$$\mathrm{IP} = -\varepsilon_{\mathrm{HOMO}} \tag{64}$$

The ground-state energy of a molecule varies continuously with fractional variations in the number of electrons in the system. For integer variations on the number of electrons, the exact exchange-correlation potential component of the total energy jumps by a constant, i.e. it has a derivative discontinuity at any integer number of electrons. However, the exact exchange-correlation functional may never be found and the approximations that are in use cannot account for these discontinuities. The smoothing of the curve introduces errors that make $\varepsilon_{\mathrm{HOMO}}$ deviate from the ideal relation in Eq. (64). Surprisingly, it has been shown [73] that the accumulation of errors makes the energy of the HOMO tend to the average of the IP and the EA instead of to the IP. Then

$$\varepsilon_{\mathrm{HOMO}} \approx -\frac{\mathrm{IP}+\mathrm{EA}}{2} \tag{65}$$

Combining Eqs. (64) and (65), we finally get an expression to find the Fermi level of a molecule as approximately the energy of the Kohn–Sham HOMO.

$$\mu \approx \varepsilon_{\mathrm{HOMO}} \tag{66}$$

5.2. *"Fermi-level alignment" in metal–semiconductor interfaces*

One of the paramount issues in the study of metal–semiconductor junctions relates to the electronic equilibration of charges across the interface. When having two materials with different Fermi levels in direct contact, electrons flow from the material with higher Fermi level to the one with lower Fermi level until equilibrium is reached. At equilibrium, it is said that the junction has a unique Fermi level throughout the two materials, this is called the "Fermi-level alignment" rule.

The rearrangement of charges produces a built-in electric field at the interface, which helps to maintain the equilibrium at the interface. The distribution of charges is expressed as a built-in electrostatic potential profile $V_{bi}(x)$ across the junction. This potential modifies the original Fermi level to produce an effective Fermi level, μ^*, in the following way

$$\mu^* = \mu + eV_{bi}(x) \tag{67}$$

Then "Fermi-level alignment" refers strictly to the alignment of the effective Fermi levels of the materials conforming the junction, not to the alignment of the Fermi levels.

Our method of studying the interfaces is schematized in Figure 12. Zones I and V correspond to the regions of the junction where both contacts (contact 1 and contact 2) behave as bulk materials and their effect on the junction is accounted using the Green function method. The critical part of the junction is the region where both bulk materials are in direct contact; the formation and breakage of molecular bonds takes place in this region, resulting in a new material (material 3) that is neither contact 1 nor contact 2 (see nomenclature in Table 2). The electronic properties of the junction depend mostly on the character of this interface; thus, a high degree of accuracy is needed in modeling this region. This region is treated as a separate new molecule, which is the extended molecule defined in Figure 6, and calculate quantum-mechanically. The extended molecule is

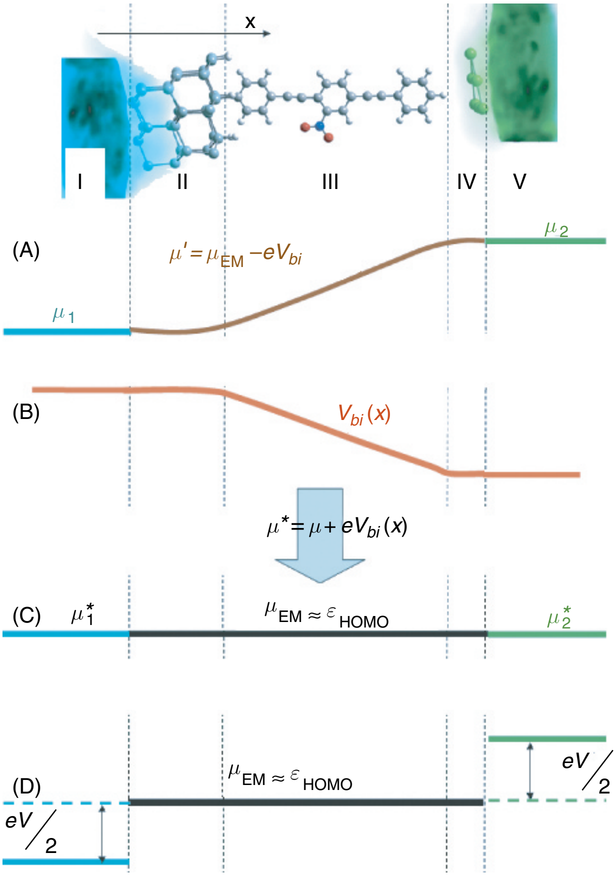


Figure 12 (A) Schematic of the electrochemical potential (Fermi level) distribution, $\mu(x)$, along the x axis, perpendicular to the junction. Zone I corresponds to bulk CNT, zone V to the silicon bulk, and zones II, III, and IV to the extended molecule. (B) Spatial distribution of the electrostatic potential (ESP), $V_{bi}(x)$, for the CNT–molecule–Si junction. (C) Spatial distribution of the effective electrochemical potential (effective Fermi level), $\mu^*(x)$, across the junction. (D) Shifting of the effective electrochemical potential (effective Fermi level) across the junction upon the application of an external bias voltage V

comprised of the zones II, III, and IV (Figure 12). Several atoms belonging to the contacts (interfacial atoms, zones II and IV) are included as part of the extended molecule.

In other words, our model considers the original two-contact junction as a junction composed of three different materials: material 1 (the contact 1), material 3 (the extended molecule), and material 2 (the contact 2). These three distinct materials reach and stay in equilibrium. Their effective Fermi levels are aligned to the value of the Fermi level of the extended molecule, as shown in Figure 12C. According to Eq. (66), the Fermi level of the extended molecule corresponds to the energy of the Kohn–Sham HOMO.

In order to read/write information from/in the molecule, an external bias voltage, V, needs to be applied between the contacts. Upon the application of the external voltage, the junction gets out of equilibrium. As a first approximation, the effective Fermi levels

Table 2 Parallel between several equivalent names given to the components of a junction. The extended molecule is composed of the interfacial atoms and the restricted molecule

Junction	Components	Molecule	Figure 12A	Figure 6
Material 1	Contact 1	Au DOS	I	Bulk contact
	Interface	Au atoms	II	Interfacial atoms
Material 3	Interface	nitroOPE	III	Restricted molecule
	Interface	Si atoms	IV	Interfacial atoms
Material 2	Contact 2	Si DOS	V	Bulk contact

of both contacts are affected by the external voltage as shown in Figure 12. This gradient of effective Fermi levels along the junction produces a flow of electrons between the contacts, i.e. current.

5.3. Quantum-mechanical calculation

5.3.1. Gold contact

The Au–nitroOPE–Si junction (Figure 13C) is composed of 6 interfacial Au atoms, which model the top contact, and 38 Si atoms, which model the bottom contact. The geometry for this extended molecule is obtained by performing quantum-mechanical optimizations of the top and bottom components of the junction separately.

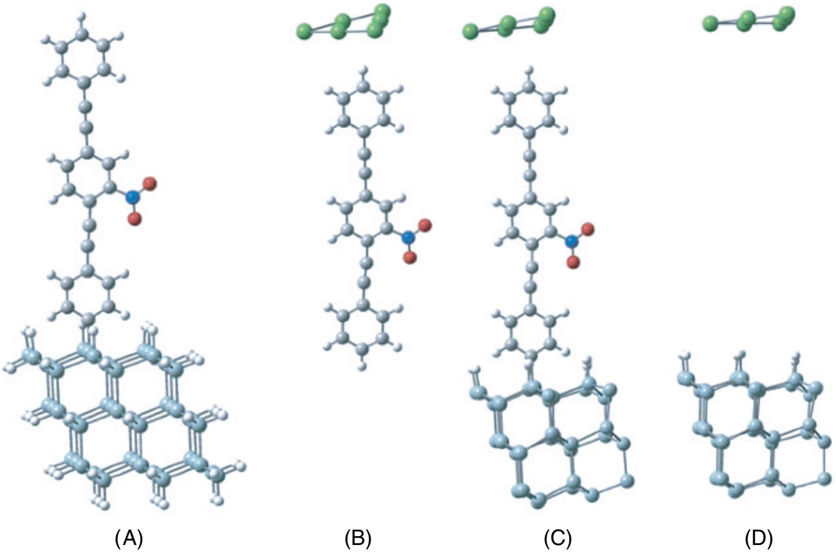


Figure 13 (A) Optimization of the bottom past of the junction. (B) Optimized geometry corresponding with the top part of the junction. (C) Final assembly of the Au–nitroOPE–Si junction. (D) Associated Au-Si tunneling junction. For higher compatibility all calculations are performed under the same DFT method and basis set (B3PW91/LANL2DZ)

To find an appropriate geometry for the bottom part of the junction, we optimize the nitroOPE molecule perpendicularly bonded to a hydride-passivated Si (111) surface, which is modeled by 52 silicon atoms (Figure 13A). Hydrogen atoms are added to saturate the boundary Si atoms. This molecule presents a total dipole moment of 5.08 D (+2.72 D in the direction of the junction). The optimized C−Si bond length is 1.913 Å.

The top part of the junction is found by optimizing the nitroOPE molecule and six gold atoms (Figure 13B). We run several calculations with increasing number of Au atoms (from 1 to 6); those geometry optimizations show that the gold atoms tend to a planar conformation and that there is no chemical bond between the gold atoms and the nitroOPE molecule. For compatibility, the optimization of the top (Figure 13B) and bottom (Figure 13A) parts of the junction is performed using the same level of theory, B3PW91, and basis set, LANL2DZ.

Figure 13C shows the final assembly of the Au–nitroOPE–Si junction from the optimized bottom and top parts. For practical reason to confront the computationally challenging nature of the geometry optimizations, the assembled geometry of the junction (Figure 13C) is kept fixed (not fully optimized) for all subsequent calculations. Also, notice that the number of total silicon atoms is reduced to 38 with respect to Figure 13A. The total dipole moment for this junction is 9.03 D (+7.8 D in the direction the junction).

We also calculated an alternative geometry, the perpendicular conformation. In that conformation, the middle phenyl ring, which contains the nitro group, is rotated 90° with respect to the plane of the other two phenyl rings. If the opposite is not stated explicitly, the default conformation corresponds to "coplanar", where all the phenyl rings are contained in a plane, as seen in Figure 13C.

The calculations of both conformations, shown in Table 3, predict that the Au–nitroOPE–Si junction is more stable in the perpendicular conformation than in the coplanar conformation, with an energetic barrier of 0.19 eV (4.3 kcal/mol, $\sim 7\,kT$) for rotation of the middle phenyl ring.

5.3.2. (4, 4) CNT contact

Recently, several procedures have been reported for attaching covalently aromatic hydrocarbons (arenes) to CNTs [44–46]. Manipulation of CNTs has been limited since they are synthesized as bundles or ropes. Because of the tendency to agglomerate, CNTs present low solubility and dispersion when placed in polymer matrices [74]. The ability to attach arene "handles" to CNTs allows direct manipulation of this amazing form of carbon, opening new possibilities of using individual CNTs as molecular devices.

Table 3 Summary of the calculation for the Au–nitroOPE–Si junction

	Coplanar conformation	Perpendicular conformation
Calculation type	single point	single point
Calculation method	UB3PW91	UP3PW91
Basis set	LANL2DZ	LANL2DZ
Total electronic energy	−2014.57326 Ha	−2014.58015 Ha
Dipole moment	9.03 D	9.52 D

Moreover, several functionalization techniques have been reported to react faster in metallic CNTs rather than in semiconducting ones [43, 75, 76], which has allowed the separation of CNTs based on their electronic properties, i.e., metallic from semiconducting [43]. The advances have opened the possibility of using metallic CNTs as tips for contacting organic molecules.

On the other hand, the synthesis of nitroOPE molecules perpendicularly assembled on a hydride-passivated Si (111) substrate, with the top end covalently attached to a metallic CNT, i.e., the mettalic CNT–nitroOPE–Si junction shown in Figure 14A, has been reported recently [29]. Computationally, the use of atoms with smaller atomic number, such as carbon instead of gold, has the advantage of allowing a full-electron study of the system, which leads to a more precise calculation.

We optimize the geometry of the (4, 4) CNT–nitroOPE–Si junction by parts. The top part of the geometry is obtained by optimizing a piece of (4, 4) CNT with a benzene ring covalently bonded to it. The piece of the armchair (4, 4) CNT is composed of 10 carbon rings, each ring containing 8 carbon atoms. The positions of the CNT atoms away from the benzene-CNT bond are kept fixed. The bottom part is obtained as described for the Au–nitroOPE–Si junction. The geometry of the assembled (4, 4) CNT–nitroOPE–Si junction is shown in Figure 14A. Due to computational restrictions, this geometry is kept fixed for all subsequent calculations.

We calculated the coplanar (Figure 14A) and the perpendicular (Figure 14B) conformations. Contrary to the case when having a gold top contact, the coplanar

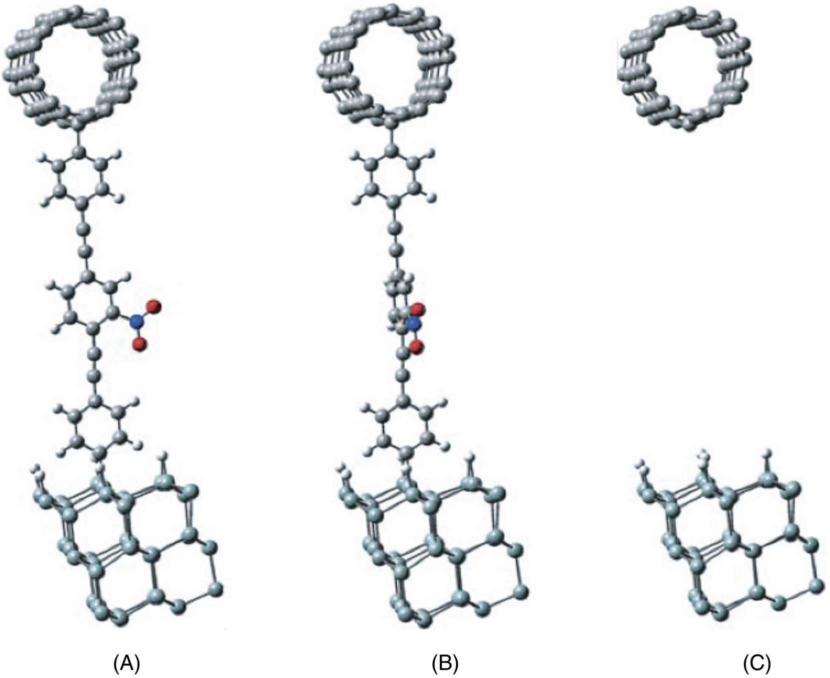


Figure 14 (A) Coplanar and (B) perpendicular configuration of the (4, 4) CNT–nitroOPE–Si junction. (C) The (4, 4) CNT–Si tunneling junction has the interfacial atoms in the same position as (A) and (B), but without the nitroOPE molecule between them

Table 4 Summary of the calculation for the (4, 4) CNT–nitroOPE–Si junction

	Coplanar configuration	Perpendicular configuration
Calculation type	single point	single point
Calculation method	UB3PW91	UP3PW91
Basis set	6-31G(d)	6-31G(d)
Total electronic energy	−15097.75352 Ha	−15097.74713 Ha
Dipole moment	130.46 D	133.37 D

conformation of the CNT–nitroOPE–Si junction turns out to be slightly more stable than the perpendicular conformation (Table 4), with a rotational barrier of 0.17 eV (4 kcal/mol).

The calculated total dipole moment is 130.46 D (−130.16 D in the direction of the junction) for the coplanar configuration. The perpendicular configuration presents a similar dipole moment: 133.37 D (−133.03 D in the direction of the junction).

Because of the larger spatial extension of d-electrons over p-electrons, the wavefunction of gold can tunnel farther into the vacuum than the wavefunction of a CNT contact. The variation of the ESP (Figure 15B and 15D) along metal–Si tunneling junctions (Figure 15A and C) corroborates the fact that the wavefunction of gold can tunnel farther, yielding higher tunneling currents. Gold would apparently be a superior choice for metallic contact than the (4, 4) CNT would for a nitroOPE-based molecular device.

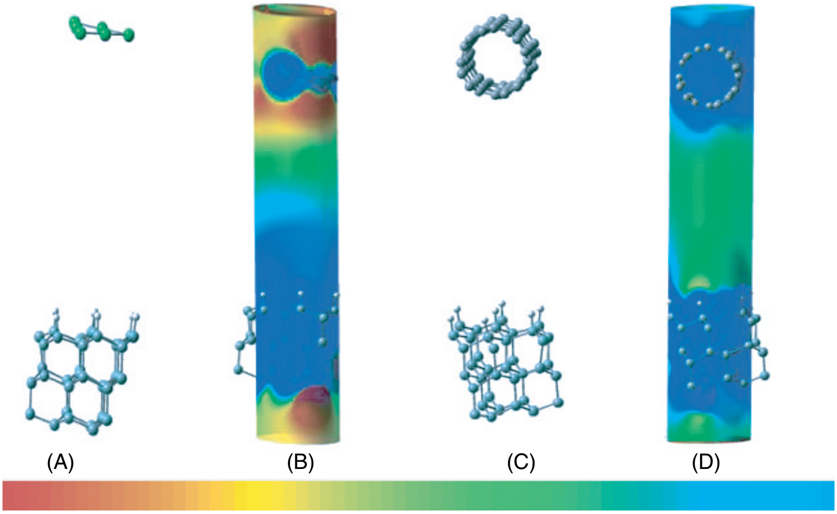


Figure 15 (A) Geometry of the Au–Si tunneling junction. The position of the gold and silicon atoms are kept the same as in the Au–nitroOPE–Si junction. (B) Distribution of the ESP for (A). The spatial region corresponds to the same cylindrical surface shown in Figure 21C. (C) CNT–nitroOPE–Si junction. (D) Distribution of the ESP for (C). The spatial region for all the figures corresponds to a cylinder of radius 4 Å. The color scale for all the figures ranges from −0.1 V (red) to 0.1 V (blue)

However, the CNT has the advantage of forming a covalent bond with the nitroOPE-Si whereas gold forms a physical bond.

5.3.3. Mulliken charges

The distribution of charges is computed based on the recipe given by Mulliken [77–80], which starts from the calculated wavefunction of the molecular system.

The analysis of Mulliken atomic charges (Table 5) shows a strong rearrangement of charges taking place in the (4, 4) CNT–nitroOPE–Si junction; for the coplanar neutral state of this junction, the CNT loses 0.97*e*, the Si loses 1.09*e* and the nitroOPE gains 2.06*e*. The high rearrangement of charges accounts for the high dipole moment of the (4, 4) CNT–nitroOPE–Si junction, −130.16 D in the direction of the junction.

The optimization of the gold atoms in the top contact of the Au–nitroOPE–Si junction shows a gap between the plane of gold atoms and the nitroOPE molecule (Figure 13B). This gap obstructs the free displacement of charges between the Au contact and the rest of the junction, explaining the very low charge rearrangement throughout the Au–nitroOPE–Si junction (Table 6). Most of low charge transfer takes place between the nitroOPE and the Si contact with an almost-null transfer between the nitroOPE and the Au contact, 0.03*e*. This also explains the relatively low dipole moment that is found for the Au–nitroOPE–Si junction (7.80 D in the direction of the junction).

The metal–Si junctions (Figure 15A and C) present a gap of ~20 Å, which is large enough to obstruct any transfer of charges between the contacts. The lack of charge displacement results in the negligible dipole moment found for the CNT-Si, 1.31 D in the direction of the junction, and the Au–Si tunneling junction, 2.85 D in the direction of junction.

Table 5 Distribution of Mulliken charges for the (4, 4) CNT–nitroOPE–Si junction in its coplanar conformation. The Si contact includes the hydrogen atoms adsorbed on it. The units of the charges are in *e*, the absolute value of the charge of an electron

	Neutral	Anion	Dianion	Trianion
CNT contact	0.97	0.19	0.00	0.00
nitroOPE	−2.06	−0.33	−0.98	−1.22
Si contact	1.09	−0.86	−1.02	−1.78
total charge	0	−1	−2	−3

Table 6 Distribution of Mulliken charges for the Au–nitroOPE–Si junction in its coplanar conformation. The Si contact includes the hydrogen atoms adsorbed on it. The units of the charges are in *e*, the absolute value of the charge of an electron

	Neutral	Anion	Dianion	Trianion
Au contact	−0.03	−0.03	−0.08	−1.64
nitroOPE	−0.10	−0.17	−0.19	−0.26
Si contact	0.13	−0.80	−1.73	−1.10
total charge	0	−1	−2	−3

5.4. Current–voltage calculation

The calculation of current assumes electrons being injected from the top contact (negative electrode) to the bottom contact (positive electrode). At zero bias voltage ($V = 0$), the most energetic electrons in the top and bottom bulk contacts have the same energy; therefore, the junction is in equilibrium, without net flow of electrons. This is called "Fermi-level alignment" as described in Section 5.2. The most intimate part of the junction is modeled by an extended molecule, which contains atoms representing both contacts; the Fermi level of the extended molecule gives an approximation of the Fermi level of the macroscopic junction. The quantum-mechanical calculations allow to find the Fermi level of the extended molecule, which corresponds to the energy of the HOMO, as discussed in Section 5.1.

The applied bias voltage (V) is defined such that the semiconducting contact is positively biased with respect to the metallic contact, $V = V_{\text{semic}} - V_{\text{metal}}$. Therefore, after applying a bias voltage between the contacts, the effective Fermi level of the metal is shifted up whereas the effective Fermi level of the Si contact is shifted down (by an equal amount of $0.5 \times e \times V$) with respect to the equilibrium Fermi level of the extended molecule (μ_{EM}), in the following way

$$\text{Metal:} \quad \mu_2^* = \mu_{\text{EM}} + \frac{1}{2}eV \tag{68}$$

$$\text{Semiconductor:} \quad \mu_1^* = \mu_{\text{EM}} - \frac{1}{2}eV \tag{69}$$

The values of current reported here refer to "current of electrons" and is defined as positive when flowing from the metal (contact 2) to the semiconductor (contact 1).

5.4.1. Gold contact

The Fermi levels for the Au-nitroOPE-Si (μ_{EM}) are calculated as the energy of the HOMO ($\varepsilon_{\text{HOMO}}$) of the extended molecule. The calculated values for the $\varepsilon_{\text{HOMO}}$ and the $\varepsilon_{\text{LUMO}}$ of several Au-nitroOPE-Si junctions are reported in Table 7.

The Green function, $g(E)$, for the metallic contact is based on the density of states for the FCC gold crystal, which is calculated under the same level of theory (B3PW91)

Table 7 Summary of the α-HOMO and α-LUMO energies for the different charge states and conformations of the Au–nitroOPE–Si. The calculations are performed using the B3PW91 method and the LANL2DZ basis set

Conformation	Charge (e)	μ_{EM} or $\varepsilon_{\text{HOMO}}$ (eV)	$\varepsilon_{\text{LUMO}}$ (eV)
Coplanar	0	−5.45	−4.75
	−1	−2.96	−2.86
	−2	−1.67	−1.04
	−3	0.36	0.50
Perpendicular	0	−5.37	−4.70
	−1	−3.00	−2.77
	−2	−1.47	−0.92

and basis set (LANL2DZ) as the junction. The DOS for Au used for the current–voltage calculations is shown in Figure 1. For compatibility, we use a Si DOS calculated using the combination B3PW91/LANL2DZ (Figure 4).

The calculation of the current–voltage characteristic for the junction with the nitroOPE in its coplanar conformation is shown in Figure 16A. Only the contribution of α-electrons is taken into account for the I-V calculation.

5.4.2. (4, 4) CNT contact

We test several (4, 4) CNT–nitroOPE–Si junctions, which include the coplanar, perpendicular, and charge states. The quantum-mechanical calculation of the values for the Fermi levels is performed using the B3PW91 method and the full-electron 6-31G(d) basis set. These values are reported in Table 8.

The Green functions, $g(E)$, for the Si and CNT contacts are based on B3PW91/6-31G(d) calculations of the DOS using the Crystal 03 software. The Si DOS is shown in Figure 3, it is calculated using the 6-31G(d) basis set. We point out that the Si DOS use for the Au–nitroOPE–Si junction is calculated using a different basis set, the LANL2DZ basis. The DOS of the (4, 4) CNT is shown in Figure 5.

Since the molecule is connected to one half of the Si bulk, we consider that only half of the total Si electronic states are available to leak into the molecule through the bottom contact. Then, a DOS factor (a scaling factor) of 0.5 is applied to the silicon DOS, mostly on intuitive geometrical grounds [36]. Interestingly, the atoms in a CNT are both surface and bulk atoms at the same time; therefore, DOS per atom of the bulk material corresponds to the DOS of the atom to which the nitroOPE molecule is adsorbed and no factor needs to be applied to the CNT DOS. The current–voltage characteristic for the coplanar (4, 4) CNT–nitroOPE–Si junction is reported in Figure 16B.

For both metal–nitroOPE–semiconductor junctions, we notice a flat region in the I-V curve at low bias voltage (approximately from −1.4 to 1.4 V). This flat region of approximately zero conductance has also been observed in experimental calculations of semiconductor–molecule–metal (metallic STM tip) junctions [81–84] and in insulator–molecule–metal junctions [85], Al–AlO_x–molecule–Ti–Al. The flat region in the I-V curve constitutes the most notorious difference with respect to the metal–nitroOPE–metal junctions (Figures 9 and 10).

Table 8 Summary of the α-HOMO and α-LUMO energies for the different charge states and conformations of the (4, 4) CNT–nitroOPE–Si junction

Conformation	Charge (e)	μ_{EM} or ε_{HOMO} (eV)	ε_{LUMO} (eV)
Coplanar	0	−4.83	−4.50
	−1	−3.06	−3.06
	−2	−1.58	−1.53
	−3	0.07	0.12
Perpendicular	0	−4.81	−4.48
	−1	−3.18	−3.01
	−2	−1.57	−1.52
	−3	−0.11	0.19

Despite the fact that gold is a metal with higher Dos than the (4, 4) CNT (~10 times higher at their respective Fermi levels), the current in the CNT–nitroOPE–Si junction is similar to or higher than the current in the Au–nitroOPE–Si junction. This indicates that a better chemical contact between the molecule and the metallic contact can make up for the lower density of electrons of the (4, 4) CNT.

Any junction composed of a metallic of a metallic and a semiconducting contact is a simple Schottky diode. The interface between the metal and the semiconductor gives rise to a potential barrier, which was first explained by Schottky. The Schottky barrier obstructs the transport of carriers in one direction of the junction, acting as an electrical rectifier. The rectifying behavior is one of the most characteristic features associated with macroscopic Schottky diodes; however, lack of rectifying behavior in Schottky diodes of nanometer sizes has been experimentally observed [64–66]. Likewise, no rectifying behavior is seen in our I-V calculations of molecular metal–molecule–semiconductor junctions (Figure 16).

Three mechanisms for electron transport can take place in a junction: thermionic emission, tunneling, and diffusion. Thermionic emission and tunneling are the most

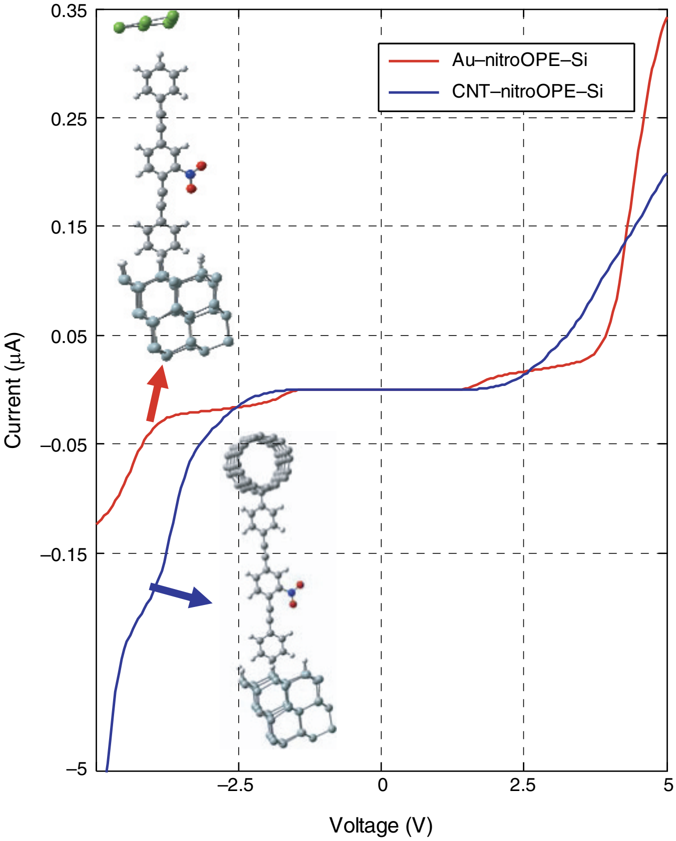


Figure 16 (A) Current–voltage characteristic for the Au–nitroOPE–Si in linear scale. (B) I-V for the (4, 4) CNT–nitroOPE–Si junction

important mechanisms in Schottky diodes. The thermionic transport depends mostly on the height of the Schottky barrier whereas the tunneling transport depends on both the height and the thickness of the barrier. The thermionic contribution to the current is given by the following equation:

$$I \propto T^2 e^{\frac{\phi_B}{kT}} \left[e^{\frac{eV}{kT}} - 1 \right] \tag{70}$$

where ϕ_B is the height of the barrier and T, temperature. The tunneling current, for a triangular barrier, is given by:

$$I \propto \Theta; \qquad \Theta = e^{\frac{4}{3}\frac{\sqrt{2m\phi_B}}{\hbar}L} \tag{71}$$

where Θ is the tunneling probability, m is the effective mass, L is the thickness of the barrier, and h is the reduced Planck constant ($h = h/2\pi$).

Smit *et al.* used a theoretical model based on the Poisson equation to track the behavior of the Schottky barrier for diodes of arbitrary sizes, from macroscopic to ultra-small dimensions. Following that top–down methodology, they demonstrated that for diodes smaller than a characteristic length (associated with the doping level of the semiconductor), the thickness of the potential barrier no longer depends on the concentration of the dopant in the semiconductor but on the size and shape of the diode [86, 87]. "Molecular Schottky diodes" exhibit thin potential barriers; therefore, the tunneling contribution to conduction outweighs the thermionic contribution [64, 65]. Then, contrary to macroscopic diodes, the I-V of molecular diodes does not exhibit the characteristic diode-like shape of Eq. (70).

Because of the small length of molecular junctions, the electron transport is coherent. Our DFT-GF is built upon the Landauer formalism, which deals with coherent transport. In coherence transport, electrons travel non-interactively from one contact to the other in a single quantum-mechanical process whose probability can be calculated directly from the fundamental equations. The transport in molecular junctions can be seen as a probability for an electron to cross from one side of the molecule to the other, which is a tunneling process. Equation (53), used in our formalism to calculate the current–voltage characteristic, reflects the fact that tunneling is the main mechanism for electron transport through molecular junctions.

It is interesting to see how Smit *et al.*, by using a top-down approach, reached the conclusion that electron transport in ultra-small Schottky diodes is predominantly by tunneling. This conclusion arises naturally when using an atomistic (bottom-up) approach, such as our DFT-GF interpretation of the Landauer formalism.

5.5. Changes in the conformation and charge states

This section very importantly discusses how to determine the current-voltage characteristics of the junctions. We will analyze the gold and the (4, 4) CNT contacts as well as the electrostatic potential distribution along the junction and their molecular orbitals.

5.5.1. Gold contact

We also analyze the perpendicular conformation of the Au–nitroOPE–Si junction. For that, we rotate by 90° the atoms of the middle phenyl ring and keep all the other simulation parameters the same as for the coplanar conformation.

Our result predicts a drastic difference in conductance between the coplanar and perpendicular conformations (Figure 17). Such change in conductance has been attributed to the rupture of the π-orbital network [88]. When the phenyl rings are coplanar, they form a conducting path across the molecule; however, when the phenyl rings are perpendicular to each other, this conducting path is broken, decreasing tremendously the conductance.

This huge change of conductance was observed by Donhouser [48] in STM electrical measurement of several types of OPE molecules, including the nitroOPE, over time.

The charge trapped in the molecular junction has been thought to cause a strong change in the conductance of molecular junctions [15, 85]; we investigate that by calculating the I-V of Au–nitroOPE–Si junctions containing an extra number of electrons (anion, dianion, and trianion). In the coplanar conformation, we observe a strong decay in conductance when the extended molecule gets negatively charged (Figure 18A). The anion and dianion states present similar low conductance, but the trianion is seen to have the lowest conductance. At 3.3 V the neutral junction conduces 23.6 nA, which is 34.6, 75.3, and 23.6 times the current conduced by the anion, the dianion, and the trianion states respectively. At 1.0 V the ratios are even better, $I_{\text{neutral}}/I_{\text{anion}} = 88.5$, $I_{\text{neutral}}/I_{\text{dianon}} = 123.4$, $I_{\text{neutral}}/I_{\text{trianion}} = 246.2$, where $I_{\text{neutral}} = 5.5 \times 10^{-13}$ A.

High ratios of the currents ($I_{\text{neutral}}/I_{\text{charged}}$) are encouraging for the design of a bistable electronic device. Although the absolute current values are too small ($\sim 10^{-13}$ A) to be measured by present equipment (we point out that those values correspond to the conduction through a single molecule), in reality, thousands of molecules are expected to be self-assembled on parallel, with the net current being also thousand times higher.

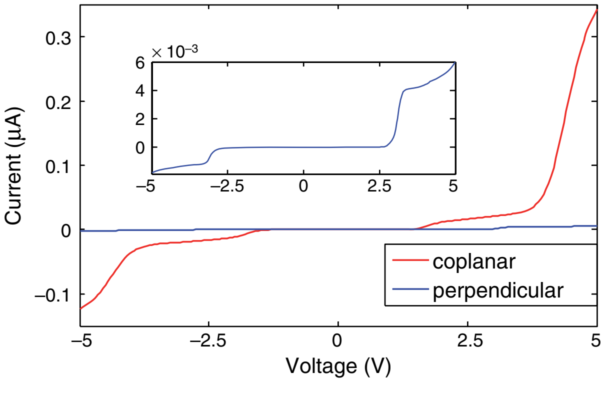
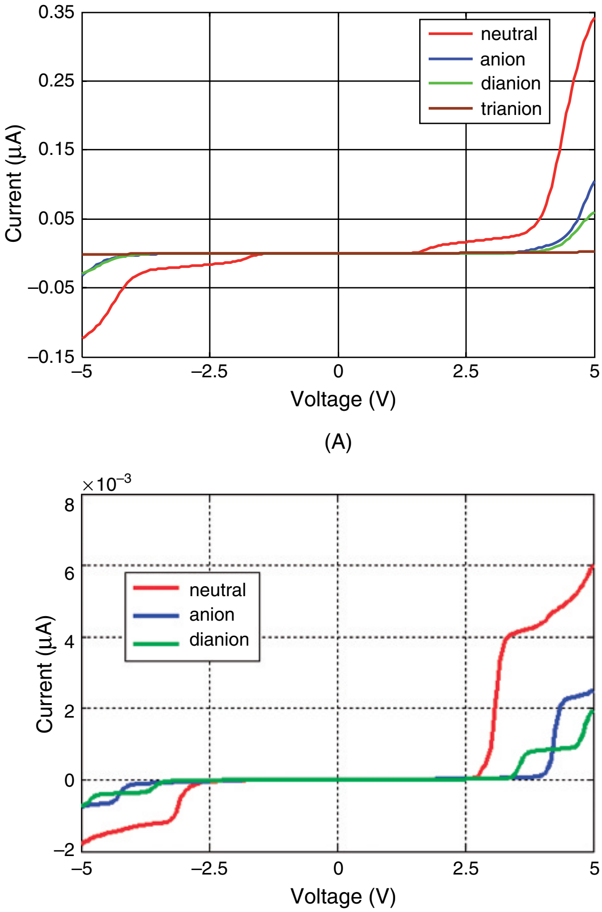


Figure 17 Current–voltage characteristic for two different geometrical conformations of the Au_6–nitroOPE–Si. The planar conformation is referred to when the three rings are coplanar, and the perpendicular conformation refers to the case when the ring containing the nitro group is perpendicular to the other two. The inset shows a zoomed view of the I-V characteristic for the molecule in its perpendicular conformation



(A)

(B)

Figure 18 (A) Current–voltage characteristic for different charge states of the Au–nitroOPE–Si junction in their coplanar conformation. Only the contribution of α electrons is shown. (B) Current–voltage characteristic for the different charge states of the Au–nitroOPE–Si junction in its perpendicular conformational state

In the perpendicular conformation, the junctions already have a much lower conductance for the neutral state with respect to the coplanar configuration, as seen in Figure 17. The addition of charge to the perpendicular conformation lowers even more the conductance as reported in Figure 18B.

5.5.2. (4, 4) CNT contact

We calculate the I-V for the perpendicular conformation of the (4, 4) CNT–nitroOPE–Si junction. Figure 19 shows a comparison of the conductance between the coplanar and the perpendicular conformations. The calculation confirms that the strong change in

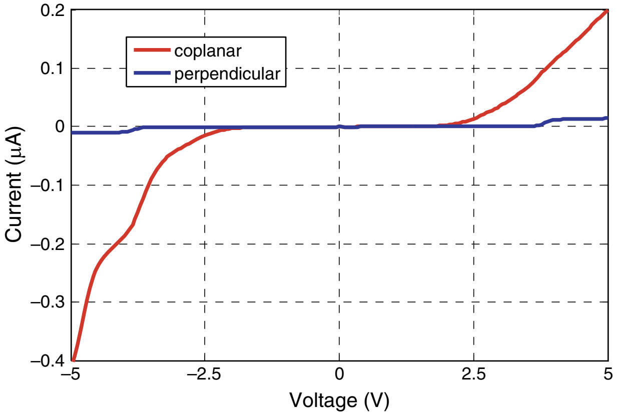


Figure 19 Comparison of the current–voltage characteristics of the coplanar and perpendicular configurations for the (4, 4) CNT–nitroOPE–Si junction

conductance between the two conformational states is still present if we use the metallic CNT instead of gold.

For instance, at 3.3 V the molecule in its coplanar conformation conducts 53 nA, which is 496 times the current that we find for the perpendicular conformation.

We calculate the I-V characteristic for several charge states (charges −1, −2, and −3) for the coplanar and perpendicular configurations of the (4, 4) CNT–nitroOPE–Si junction, as summarized in Figure 20. The geometries of the charge states are kept fixed to the geometry of the neutral molecule, only the wavefunctions are optimized. The optimization of all the conformational and charge states is performed using the combination of B3PW91 level of theory and the 6-31G(d) basis set.

Our I-V calculations for the negatively charged states of the coplanar (4, 4) CNT–nitroOPE–Si junction show a reduction in conductance with respect to the neutral (Figure 20A). The anion, dianion, and trianion present similar values of conductance, which is clearly distinctive from the neutral. For instance, at 3.3 V, the neutral molecule conducts 52.9 nA, the anion 17.4 nA, the dianion 5.9 nA, and the trianion 10 nA. The reduction in conductance is not as drastic as when using Au for the top contact.

For instance, at 3.3 V the molecule in its coplanar conformation conducts 53 nA, which is 496 times the current that we find for the perpendicular conformation.

We calculate the I-V characteristic for several charge states (charges −1, −2, and −3) for the coplanar and perpendicular configurations of the (4, 4) CNT–nitroOPE–Si junction, as summarized in Figure 20. The geometries of the charge states are kept fixed to the geometry of the neutral molecule, only the wavefunctions are optimized. The optimization of all the conformational and charge states is performed using the combination of B3PW91 level of theory and the 6-31G(d) basis set.

Our I-V calculations for the negatively charged states of the coplanar (4, 4) CNT–nitroOPE–Si junction show a reduction in conductance with respect to the neutral (Figure 20A. The anion, dianion, and trianion present similar values of conductance, which is clearly distinctive from the neutral. For instance, at 3.3 V, the neutral molecule

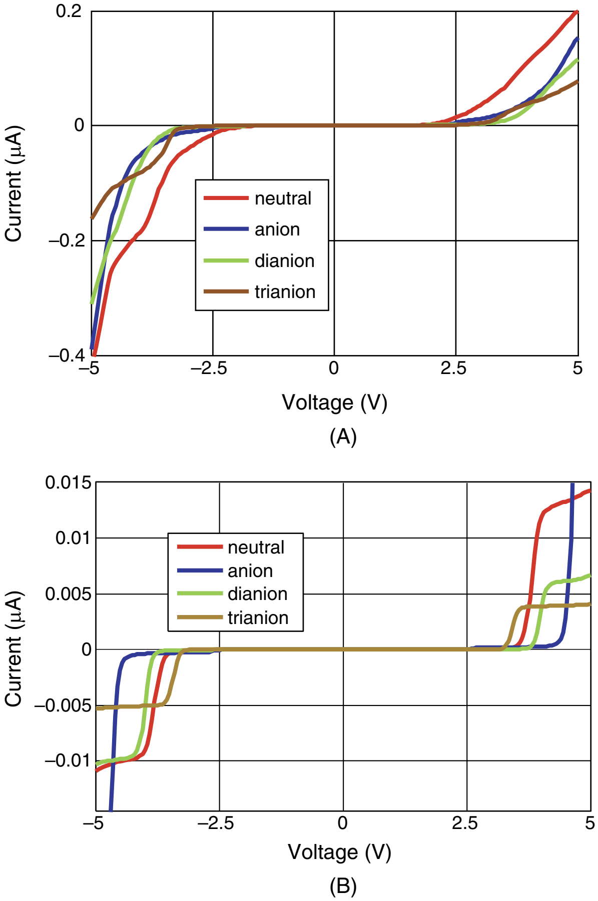


Figure 20 Current–voltage characteristic for different charge states of (4, 4) CNT–nitroOPE–Si (A) in its coplanar conformation. (B) in its perpendicular conformation. For all cases, only the contribution of α electrons is shown

conducts 52.9 nA, the anion 17.4 nA, the dianion 5.9 nA, and the trianion 10 nA. The reduction in conductance is not as drastic as when using Au for the top contact.

5.5.3. ESP distribution along the junction

We compare the ESP distribution between the neutral and anion states of the coplanar nitroOPE junction. The ESP distributions for the neutral states of the Au–nitroOPE–Si and the CNT–nitroOPE–Si are shown in Figure 21A and Figure 21C, respectively. The value for the ESP in the junction ranges from positive (blue) to negative (red) values; however, for the anion states (Figure 21, B and D) mostly regions of negative values are found within the junction. Because of electrostatic repulsion, negatively charged particles are scattered from regions of negative ESP (red). Therefore, the negatively charged junctions (negative ions) behave as nearly closed channels for electron transport,

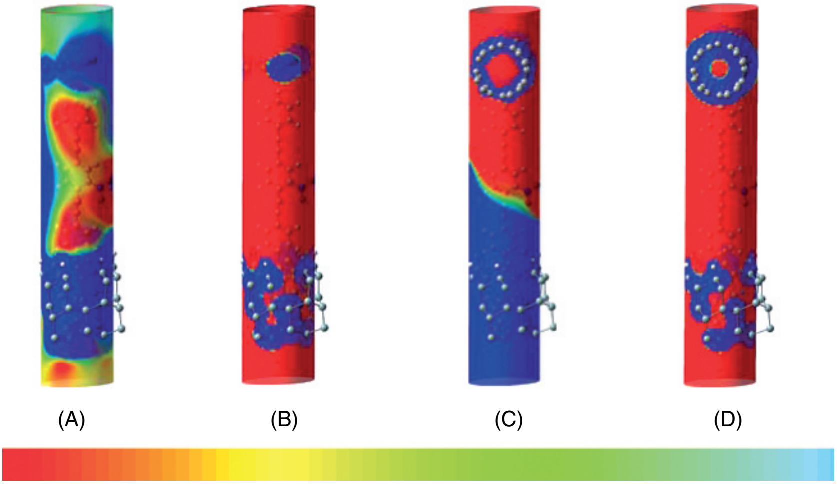


Figure 21 Distribution of the ESP for (A) the neutral and (B) anion states of the coplanar Au–nitroOPE–Si junction. Distribution of the ESP for (C) the neutral and (D) anion charge states for the coplanar (4, 4) CNT–nitroOPE–Si junction. The spatial region corresponds to a cylinder of radius 4 Å centered on the main C−C axis. The scale ranges from −0.1 V (red) to 0.1 V (blue)

which explains the noticeable reduction of current in the charged junctions as compared to their neutral states.

5.5.4. Analysis of the molecular orbitals

Molecular orbitals (θ_i^{KS}) are the mathematical solutions to the one-electron Kohn–Sham equation, given in Eq. (6). Despite controversies regarding their physical reality, MOs have extensively been used as important qualitative indicators of the conductance of molecular systems [15, 16, 89]. An MO that is delocalized throughout the molecular junction represents a conducting channel; an MO that is localized only on specific regions of the molecular junctions is not a good conducting channel. As a qualitative rule, the more delocalized an MO is, the more conducting the channel it represents.

In the resonant tunneling picture, an electron from one contact jumps into an unoccupied MO (HOMO, HOMO−1, . . .) then it jumps again from that MO to the other contact, freeing the way for another electron to repeat the process. The transport of electrons can also take place through the occupied MOs (LUMO, LUMO+1, . . .). In this case, the electron occupying the MO jumps first into one contact, with a given tunneling probability. Then, an electron from the other contact can jump into the available MO. The tunneling probability depends not only on the shape of the MO, which indicates how localized or delocalized the MO is, but also on the proximity of the energy of the MO to the Fermi level of the molecular junction. As a qualitative rule, the MOs whose energies are closest to the Fermi level of the molecule are more conducting than the ones whose energies farther apart.

A simple electrostatic explanation suffices to account for the change in conductance due to charge state, as discussed in section 5.5.3. However, we could not find a direct electrostatic explanation for the change in conductance observed between the coplanar and the perpendicular. The explanation rests more in the quantum-mechanical nature of the system. We use the MOs to explain the change in conductance between the coplanar and the perpendicular conformational states. Since the Fermi level of the molecular junction is taken as the energy of the HOMO, we expect most of the electron transport to take place through the HOMO.

For the Au_1–S–nitroOPE–S–Au_1 junction, the HOMO for the coplanar conformation is totally delocalized between the metallic contacts and the nitroOPE, as seen in Table 9. This explains the high conductance of the coplanar configuration and the low conductance of the perpendicular configuration.

Similar results are found for the Au_6–nitroOPE–S–Au_1, shown in Table 10. The HOMO for the coplanar configuration is delocalized across the nitroOPE and the bottom Au atom; however, it does not cover the top Au_6 contact, reflecting the absence of a chemical bond. Because of that, the conduction through the HOMO is not as high as for the Au_1–S–nitroOPE–S–Au_1 junction, where a chemical bond is present. However, this is compensated by the presence of three delocalized orbitals (the HOMO, HOMO$-$1, and HOMO$-$2) instead of only one. The frontier MOs for the perpendicular configuration are localized, accounting for the low conductance observed in the I-V calculation.

The MOs for the CNT–nitroOPE–CNT junction are shown in Table 11. The population of the MOs at both CNT–nitroOPE interfaces gives an indication of the seamless chemical attachment of an organic molecule to the CNT, which reflects an electronically superior contact. We observe four delocalized MOs (LUMO, HOMO, HOMO$-$3, HOMO$-$4) which enhance the conduction in the coplanar conformation. Conversely, the MOs, for the perpendicular conformation are localized.

For the coplanar Au_6–nirtoOPE–Si junction, the MOs (Table 12 close to the HOMO are localized but two delocalized MOs (HOMO$-$3, HOMO$-$4) are found at lower energies. This agrees with the flat region observed in the I-V curve of the coplanar junction, where current is found upon the application of higher voltages. All the MOs for the perpendicular conformation are localized.

The frontier MOs for the CNT–nitroOPE–Si junction are reported in Table 13. The MOs are localized for the perpendicular conformation and delocalized for the coplanar conformation, accounting for the difference in conductance found between the two conformational states.

6. Summary and conclusions

The miniaturization of conventional electronic devices can change drastically their electrical characteristics due to the predominance of quantum-mechanical effects at atomistic dimensions. Because of its technological importance, the use of silicon in molecular electronic devices is desirable. Macroscopically, metal–semiconductor junctions are known to behave as electrical rectifiers (Schottky diodes); this is apparently a drawback for the use of a semiconductor material as a contact for single molecule–based devices. However, our calculations of metal–nitroOPE–semiconductor junctions corroborated previous predictions that a "molecular Schottky diode," contrary to its macroscopic counterpart, does not present rectifying behavior. This is due to the predominance of the tunneling over the thermionic transport mechanisms at atomistic sizes.

Table 9 Molecular orbitals for the coplanar and the perpendicular conformations of the Au_1–S–nitroOPE–S–Au_1 junction. An isosurface of value 0.02 is used for all the plots

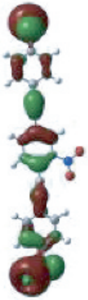
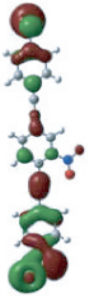
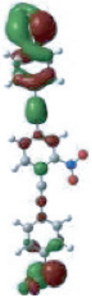
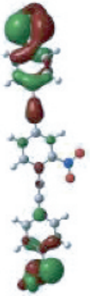
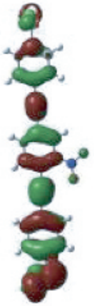
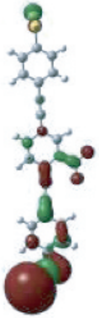
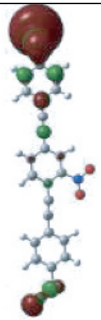
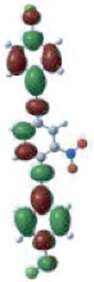
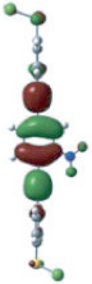
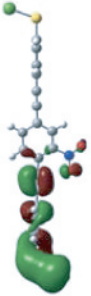
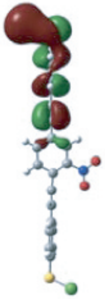
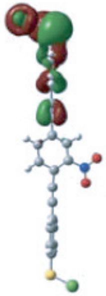
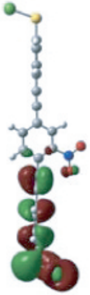
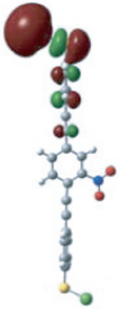
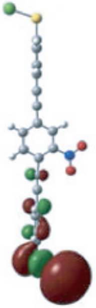
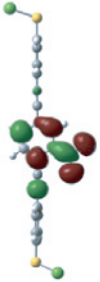
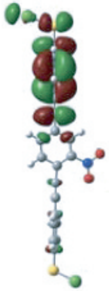
	HOMO–4	HOMO–3	HOMO–2	HOMO–1	HOMO	LUMO	LUMO+1	LUMO+2	LUMO+3	LUMO+4
neutral coplanar										
neutral perpendicular										

Table 10 Comparison of the molecular orbitals for the coplanar and the perpendicular conformations of the Au–nitroOPE–Au junction. An isosurface of value 0.02 is used for all the plots

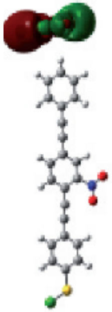
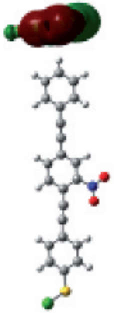
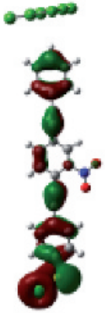
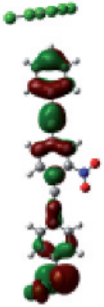
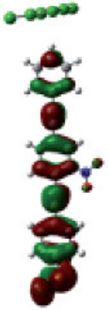
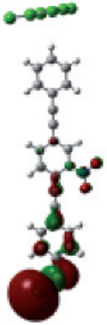
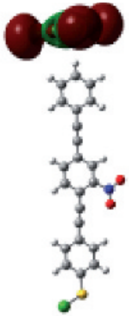
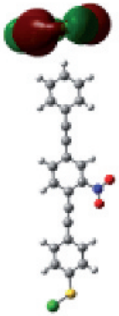
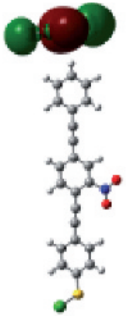
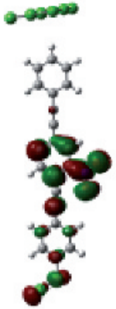
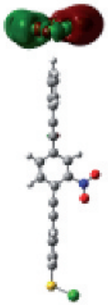
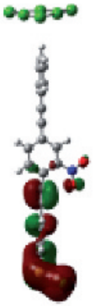
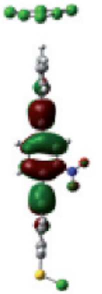
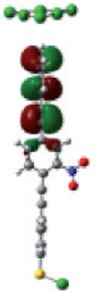
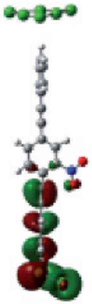
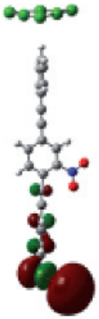
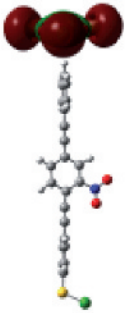
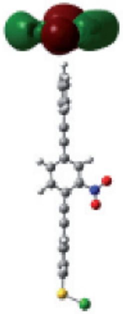
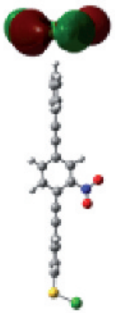
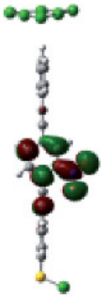
	HOMO–4	HOMO–3	HOMO–2	HOMO–1	HOMO	LUMO	LUMO+1	LUMO+2	LUMO+3	LUMO+4
neutral coplanar										
neutral perpendicular										

Table 11 Molecular orbitals for the coplanar and the perpendicular conformations of the CNT–nitroOPE–CNT junction. An isosurface of value 0.001 is used for all the plots

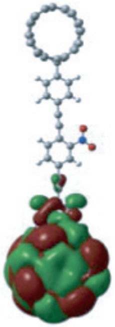
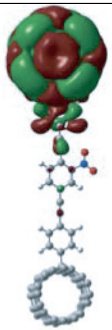
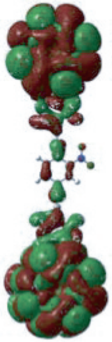
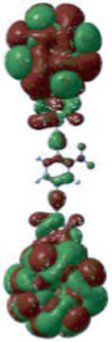
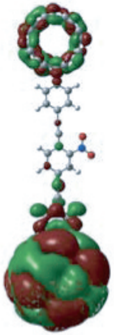
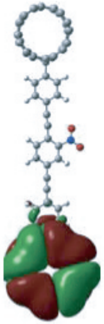
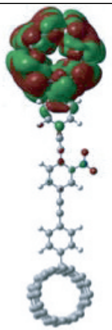
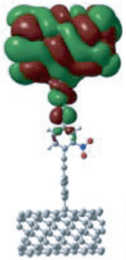
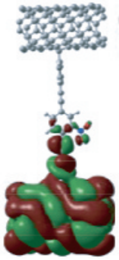
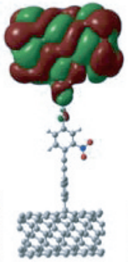
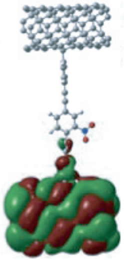
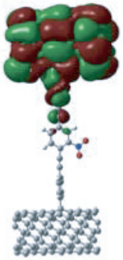
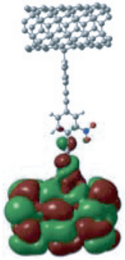
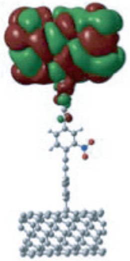
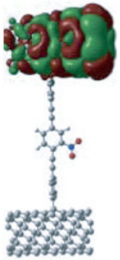
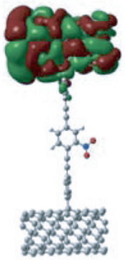
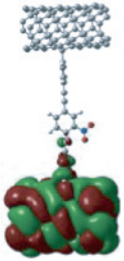
	HOMO–4	HOMO–3	HOMO–2	HOMO–1	HOMO	LUMO	LUMO+1	LUMO+2	LUMO+3	LUMO+4
neutral coplanar										
neutral perpendicular										

Table 12 Molecular orbitals for the coplanar and the perpendicular conformations of the Au_6–nitroOPE–Si junction. An isosurface of value 0.001 is used for all the plots

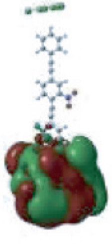
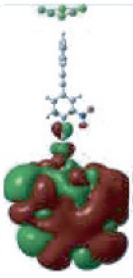
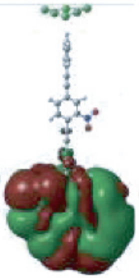
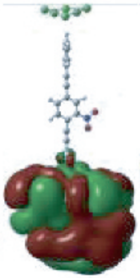
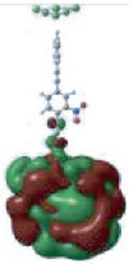
	HOMO–4	HOMO–3	HOMO–2	HOMO–1	HOMO	LUMO	LUMO+1	LUMO+2	LUMO+3	LUMO+4
neutral coplanar										
neutral perpendicular										

Table 13 Comparison of the molecular orbitals for the coplanar and the perpendicular conformations of the CNT–nitroOPE–Si junction. An isosurface of value 0.0001 is used for all the plots

	HOMO–4	HOMO–3	HOMO–2	HOMO–1	HOMO	LUMO	LUMO+1	LUMO+2	LUMO+3	LUMO+4
neutral coplanar										
neutral perpendicular										

We determined that the low- and high-conductance states, observed for metal–nitroOPE–metal junctions, could still be obtained from a metal–nitroOPE–semiconductor junction. The switch between the two states of conductance is because of the conformational and charge-state changes of the molecular junctions. However, the operating point of the device should be set at higher bias voltages since a flat region of zero current is predicted to appear at low bias voltages.

The change of conductance of the metal–nitroOPE–Si junctions due to charging effects can be accounted by a simple electrostatic explanation. For the negatively charged junctions, the ESP distribution along the junction reveals a region of negative electrostatic potential, which repels the electrons, thus, it obstructs the flow of carriers.

Changes in conductance between the coplanar and the perpendicular conformation of all the analyzed junctions can be explained by looking at the shape and spatial distribution of their molecular orbitals. The coplanar conformations have molecular orbitals that are more delocalized than the ones belonging to the perpendicular conformations. This indicates that the molecular orbitals for coplanar conformations are more conductive than those from the perpendicular conformation, therefore, explaining the higher conductance of the coplanar junctions.

Molecules can be used as electronic circuits to store binary information encoded in two molecular bistable states: for instance, a state of high conductance and a state of low conductance. Molecules can be arranged in crossbar architectures to create logic or memory electronic devices. In both cases, the desirable property of bistability depends not only on the molecule but also on the type and geometry of the contacts. Therefore, our calculations consider the molecule and few atoms of the contacts as an entwined unit, a molecular junction.

Advances in DFT and the increase in computational power of modern computers have allowed us to perform full-quantum-mechanical calculations of molecular junctions composed of up to a few hundred atoms. The continuum of electronic states, derived from the semi-infinite nature of the contacts, has been added to the molecular junction by using a mathematical formalism based on the Green function.

We have estimated the conductance of molecular junctions composed of the nitroOPE molecule and two contacts of materials such as Au, Si, and carbon nanotubes. These results are summarized in Figure 22. Two very well defined states of conductance, requisite for digital electronics, are obtained from the nitroOPE by changing the conformation and charge states of the molecule. The nitroOPE molecule is found to conduct higher current (logic “1”) when all its phenyl rings are coplanar rather than when the middle phenyl ring is perpendicular to the other two (logic “0”). In addition, bistable states can be obtained by charging the molecule. When the nitroOPE is in its neutral state, it conducts higher current (logic “1”) than when the molecule is negatively charged (logic “0”).

We determined that when two gold contacts are used to address the molecule, the bistable states of conductance are readily available at any bias voltage. However, when a combination of one semiconducting and one metallic contact is used, the bistable states are lost at low bias voltages. We found a nearly flat region in the current–voltage characteristic of metal–nitroOPE–semiconductor junctions in the low bias voltage region. Instead of presenting two distinguishable states, all the conformation and charge states of the metal–nitroOPE–semiconductor junctions (coplanar, perpendicular, anion, dianion, and trianion) showed almost indistinguishable zero conductance in the flat region. These

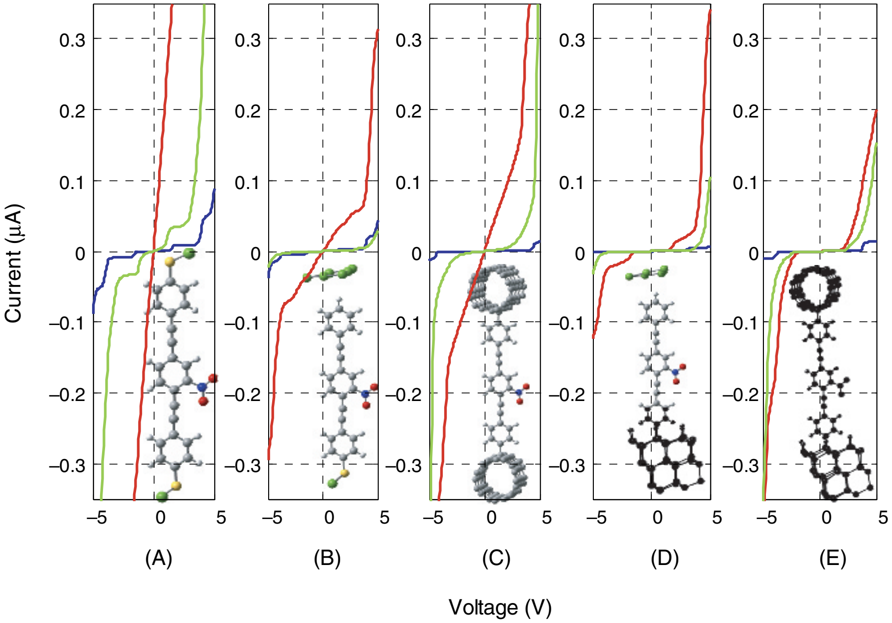


Figure 22 Summary of current–voltage characteristics for the following molecular junctions: (A) Au_1–S–nitroOPE–S–Au_1, (B) Au_6–nitroOPE–S–Au_1, (C) CNT–nitroOPE–CNT, (D) Au_6–nitroOPE–Si, and (E) CNT–nitroOPE–Si

indicate that when a semiconducting contact is used, we must move the operating point of the device to higher bias voltages.

Our calculations demonstrated that the recently synthesized CNT–nitroOPE–Si junction [29] presents the same potential as a bistable device that the Au–nitroOPE–Si junction does. However, for the former case, charging of the junction does not lead to two very distinguishable states; therefore, the two bistable states to consider should come from the neutral coplanar (logic “1”) and the neutral perpendicular (“0”).

Quantum-mechanical effects cannot be disregarded whenever dealing with electronic devices of atomistic dimensions. For instance, we corroborated that a “molecular Schottky diode” does not present the rectifying behavior associated with a macroscopic Schottky diode. It is most likely that future electronic devices will be purposely engineered to take advantage of the most exotic quantum-mechanical properties for applications such as quantum computation, quantum teleportation, and quantum cryptography. Consequently, computational tools derived from the fundamental equation of quantum mechanics, the Schrödinger equation, will be mandatory.

Acknowledgements

We thank the invaluable help of Mery Diaz Compos for her careful assistance in several aspects of this manuscript and we highly appreciate the support from the US Army Research Office (ARO), Defense Threat Reduction Agency (DTRA), and Defense

Advanced Research Projects Agency (DARPA) as well as the research funds from the Texas A & M University.

References

[1] J. M. Seminario, L. E. Cordova, and P. A. Derosa, “An ab initio approach to the calculation of current-voltage characteristics of programmable molecular devices,” *Proceedings of IEEE*, vol. 91, pp. 1958–1975, 2003.

[2] J. M. Seminario, “A theory guided approach to molecular electronics,” *Proceedings of IEEE Nanotechnology Conference*, vol. 3, pp. 75–78, 2003.

[3] J. M. Seminario, A. G. Zacarias, and P. A. Derosa, “Analysis of a dinitro-based molecular device,” *Journal of Chemical Physics*, vol. 116, pp. 1671–1683, 2002.

[4] J. M. Seminario, C. E. De La Cruz, and P. A. Derosa, “A theoretical analysis of metal–molecule contacts,” *Journal of American Chemical Society*, vol. 123, pp. 5616–5617, 2001.

[5] P. A. Derosa and J. M. Seminario, “Electron transport through single molecules: Scattering treatment using density functional and green function theories,” *J Physical Chemistry B*, vol. 105, pp. 471–481, 2001.

[6] M. A. Reed, C. Zhou, C. J. Muller, T. P. Burgin, and J. M. Tour, “Conductance of a molecular junction,” *Science*, vol. 278, pp. 252–254, 1997.

[7] J. M. Seminario, M. G. Maffei, L. A. Agapito, and P. F. Salazar, “Energy correctors for accurate prediction of molecular energies,” *Journal of Physical Chemistry A*, vol. 110, pp. 1060–1064, 2006.

[8] L. A. Agapito, M. G. Maffei, P. F. Salazar, and J. M. Seminario, “Performance of multiplicity-based energy correctors for molecular containing second-row elements,” *Journal of Physical Chemistry A*, vol. 110, pp. 4260–4265, 2006.

[9] R. G. Parr and W. Yang, *Density Functional Theory of Atoms and Molecules*. Oxford: Oxford University Press, 1989.

[10] J. M. Seminario, Y. Ma, and V. Tarigopula, “The nanocell: A chemically assembled molecular electronic circuit,” *IEEE Sensors*, in press.

[11] L. Yan, and J. M. Seminario, “Moletronics modeling towards molecular potentials,” *Int. J. Quantum Chem.*, vol. 106, pp. 1964–1969, 2006.

[12] L. Yan, Y. Ma, and J. M. Seminario, “Encoding information using molecular vibronics,” *Journal of Nanoscience Nanotechnology*, vol. 6, pp. 675–684, 2006.

[13] S. Datta, *Quantum Transport: Atom to Transistor*. New York: Cambridge University Press, 2005.

[14] P. A. Derosa, A. C. Zacarias, and J. M. Seminario, “Application of Density Functional Theory to the Study and Design of Molecular Electronic Devices; The Metal–Molecule Interface,” in K. D. Sen, ed., *Reviews in Modern Quantum Chemistry*. Singapore: World Scientific, 2002, pp. 1537–1567.

[15] J. M. Seminario, A. G. Zacarias, and P. A. Derosa, “Theoretical analysis of complementary molecular memory devices,” *Journal of Physical Chemistry A*, vol. 105, pp. 791–795, 2001.

[16] J. M. Seminario, A. G. Zacarias, and J. M. Tour, “Theoretical study of a molecular resonant tunneling diode,” *Journal of the American Chemical Society*, vol. 122, pp. 3015–3020, 2000.

[17] P. A. Derosa, V. Tarigopula, and J. M. Saminario, “Molecular electronics: Analysis and design of switchable and programmable devices using ab initio methods,” in *Encyclopedia of Nanoscience and Nanotechnology*. New York: Dekker, 2004.

[18] J. M. Seminario, “Practical aspects of electron transport through single molecules,” *Proc. IEEE Nanotech. Conf.*, vol. 4, pp. 518–520, 2004.

[19] J. M. Seminario and L. E. Cordova, "Theoretical interpretation of intrinsic line widths observed in inelastic electron tunneling scattering experiments," *Journal of Physical Chemirty A*, vol. 108, pp. 5142–5144, 2004.

[20] J. M. Seminario, C. De La Cruz, P. A. Derosa, and L. Yan, "Nanometer-size conducting and insulating scattering experiments," *Journal of Physical Chemirty B*, vol. 108, pp. 17879–17885, 2004.

[21] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, T. Vreven Jr, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Raga, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, *Gaussian 03, Revision C.02*, Wallingford CT: Gaussian, Inc., 2003.

[22] P. Hohenberg and W. Kohn, "Inhomogeneous electron gas," *Physical Review B*. vol. 136, pp. 864–871, 1964.

[23] M. Levy, "Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem," *Proceedings of the National Acadency of the United States of America*, vol. 76, pp. 6062–6065, 1979.

[24] W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlation effects," *Physical Review A*, vol. 140, pp. 1133–1138, 1965.

[25] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Mineola, N.Y.: Dover Publications, 1996.

[26] C. Roetti, "The crystal code," in C. Pisani, ed., *Quantum-Mechanical Ab-initio Calculation of the Properties of Crystalline Materials*, vol. 67, pp. 125–137, Berlin: Springer-Verlag, 1996.

[27] R. Saito, G. Dresselhaus, and M. S. Dresselhaus, *Physical Properties of Carbon Nanotubes*. London: Imperial College Press, 1998.

[28] R. Saito, M. Fujita, G. Dresselhaus, and M. S. Dresselhaus, "Electronic-structure of Chiral graphene tubules," *Applied Physics Letters*, vol. 60, pp. 2204–2206, 1992.

[29] A. K. Flatt, B. Chen, and J. M. Tour, "Fabrication of carbon nanotube-molecule-silicon junctions," *Journal of the American Chemical Society*, vol. 127, pp. 8918–8919, 2005.

[30] M. Ouyang, J. L. Huang, C. L. Cheung, and C. M. Lieber, "Energy gaps in 'metallic' single-walled carbon nonotubes," *Science*, vol. 292, pp. 702–705, 2001.

[31] M. Ouyang, J. L. Huang, C. L. Cheung, and C. M. Lieber, "Atomically resolved single-walled carbon nanotube intramolecular junctions," *Science*, vol. 291, pp. 97–100, 2001.

[32] R. A. Jishi, J. Bragin, and L. Lou, "Electronic structure of short and long carbon nanotubes from first principles," *Physical Review B*, vol. 59, pp. 9862–9865, 1999.

[33] C. D. Spataru, S. Ismail-Beigi, L. X. Benedict, and S. G. Louie, "Excitonic effects and optical spectra of single-walled carbon nanotubes," *Physical Review Letters*, vol. 92, p. 77402, 2004.

[34] C. D. Spataru, S. Ismail-Beigi, L. X. Benedict, and S. G. Louie, "Quasiparticle energies, excitionic effects and optical absorption spectra of small-diameter single-walled carbon nanotubes," *Applied Physics A: Materials Science & Processing*, vol. 78, pp. 1129–1136, 2004.

[35] P. V. Avramov, K. N. Kudin, and G. E. Scuseria, "Single wall carbon nanotubes density of States: comparison of experiment and theory," *Chemical Physics Letters*, vol. 370, pp. 597–601, 2003.

[36] P. A. Derosa and J. M. Seminario, “Electron transport through single molecules: Scattering treatment using density functional and green function theories,” *Journal of Physical Chemistry B*, vol. 105, pp. 471–481, 2001.

[37] J. M. Seminario and P. A. Derosa, “Molecular gain in a thiotolane system,” *Journal of the American Chemical Society*, vol. 123, pp. 12418–12419, 2001.

[38] J. M. Seminario, C. E. De la Cruz, and P. A. Derosa, “A theoretical analysis of metal–molecule contacts,” *Journal of the American Chemical Society*, vol. 123, pp. 5616–5617, 2001.

[39] M. Koentopp, K. Burke, and F. Evers, “Zero-bias molecular electronics: Exchange-correlation corrections to Landauer’s formula,” *Physical Review B*, vol. 73, pp. 121403(R)-1-4, 2006.

[40] S. Datta, *Electronic Transport in Mesoscopic Systems*. Cambridge, U.K.: Cambridge University Press, 1995.

[41] J. M. Seminario, A. G. Zacarias, and P. A. Derosa, “Theoretical analysis of complementary molecular memory devices,” *Journal of Physical Chemistry A*, vol. 105, pp. 791–795, 2001.

[42] M. Di Ventra, S. T. Pantelides, and N. D. Lang, “First-principles calculation of transport properties of a molecular device,” *Physical Review Letters*, vol. 84, pp. 979–982, 2000.

[43] M. S. Strano, C. A. Dyke, M. L. Usrey, P. W. Barone, M. J. Allen, H. W. Shan, C. Kittrell, R. H. Hauge, J. M. Tour, and R. E. Smalley, “Electronic structure control of single-walled carbon nanotube functionalization,” *Science*, vol. 301, pp. 1519–1522, 2003.

[44] C. A. Dyke and J. M. Tour, “Overcoming the insolubility of carbon nanotubes through high degrees of sidewall functionalization,” *Chemistry-A European Journal*, vol. 10, pp. 813–817, 2004.

[45] C. A. Dyke and J. M. Tour, “Covalent functionalization of single-walled carbon nanotubes for materials applications,” *Journal of Physical Chemistry A*, vol. 108, pp. 11151–11159, 2004.

[46] C. A. Dyke, M. P. Stewart, F. Maya, and J. M. Tour, “Diazonium-based functionalization of carbon nanotubes: XPS and GC-MS analysis and mechanistic implications,” *Synlett*, pp. 155–160, 2004.

[47] J. M. Seminario, L. A. Agapito, and H. P. Figueroa, “Towards the Design of Programmable Self-Assembled DNA-Carbon Nanotubes: An Approach to Nanobiotronics,” presented at IEEE-NANO 2002, Arlington, VA 22202, 2002.

[48] Z. J. Donhauser, B. A. Mantooth, K. F. Kelly, L. A. Bumm, J. D. Monnell, J. J. Stapleton, D. W. Price, A. M. Rawlett, D. L. Allara, J. M. Tour, and P. S. Weiss, vol. 292, pp. 2303–2307, 2001.

[49] X. D. Cui, A. Primak, X. Zarate, J. Tomfohr, O. F. Sankey, A. L. Moore, T. A. Moore, D. Gust, G. Harris, and S. M. Lindsay, “Reproducible measurement of single-molecule conductivity,” *Science*, vol. 294, pp. 571–574, 2001.

[50] T. Vondrak, H. Wang, P. Winget, C. J. Cramer, and X. Y. Zhu, “Interfacial electronic structure in thiolate self-assembled monolayers: Implication for molecular electronics,” *Journal of the American Chemical Society*, vol. 122, pp. 4700–4707, 2000.

[51] H. Sumi, “V-I characteristics of STM processes as a probe detecting vibronic interactions at a redox state in large molecular adsorbates such as electron-transfer metalloproteins,” *Journal of Physical Chemistry B*, vol. 102, pp. 1833–1844, 1998.

[52] R. Hasunuma, T. Komeda, and H. Tokumoto, “Indentation effects on atom manipulation on Si(111) surfaces investigated by STM,” *Applied Physics a-Materials Science & Processing*, vol. 66, pp. S689–S693, 1998.

[53] A. D. Yoffe, “Low-Dimensional systems–Quantum-size effects and electronic-properties of semiconductor microcrystallites (zero-dimensional systems) and some quasi-2-dimensional systems,” *Advances in Physics*, vol. 42, pp. 173–266, 1993.

[54] A. D. Yoffe, "Semiconductor quantum dots and related systems: electronic, optical, luminescence and related properties of low dimensional systems," *Advances in Physics*, vol. 50, pp. 1–208, 2001.

[55] A. L. Efros and M. Rosen, "The electronic structure of semiconductor nanocrystals," *Annual Review of Materials Science*, vol. 30, pp. 475–521, 2000.

[56] A. J. Read, R. J. Needs, K. J. Nash, L. T. Canham, P. D. J. Calcott, and A. Qteish, "1st-principles calculations of the electronic-properties of silicon quantum wires," *Physical Review Letters*, vol. 69, pp. 1232–1235, 1992.

[57] C. Delerue, G. Allan, and M. Lannoo, "Theoretical Aspects of the Luminescence of Porous Silicon," *Physical Review B*, vol. 48, pp. 11024–11036, 1993.

[58] J. B. Xia and K. W. Cheah, "Exciton states in isolated quantum wires," *Physical Review B*, vol. 55, pp. 1596–1601, 1997.

[59] J. B. Xia and K. W. Cheah, "Quantum confinement effect in thin quantum wires," *Physical Review B*, vol. 55, pp. 15688–15693, 1997.

[60] J. B. Xia and K. W. Cheah, "Electronic structure and optical transition of semiconductor nanocrystallites," *Journal of Physics: Condensed Matter*, vol. 9, pp. 9853–9862, 1997.

[61] B. Delley and E. F. Steigmeier, "Size dependence of band-gaps in silicon nanostructures," *Applied Physics Letters*, vol. 67, pp. 2370–2372, 1995.

[62] J. M. Buriak, "Organometallic chemistry on silicon and germanium surfaces," *Chemical Reviews*, vol. 102, pp. 1271–1308, 2002.

[63] M. P. Stewart, F. Maya, D. V. Kosynkin, S. M. Dirk, J. J. Stapleton, C. L. McGuiness, D. L. Allara, and J. M. Tour, "Direct covalent grafting of conjugated molecules onto Si, GaAs, and Pd surfaces from aryidiazonium salts," *Journal of the American Chemical Society*, vol. 126, pp. 370–378, 2004.

[64] P. Avouris, I. W. Lyo, and Y. Hasegawa, "Scanning tunneling microscope tip sample interactions – atomic modification of Si and nanometer Si Schottky diodes," *Journal of Vacuum Science & Technology A: Vacuum Surfaces and Films*, vol. 11, pp. 1725–1732, 1993.

[65] G. D. J. Smit, S. Rogge, and T. M. Klapwijk, "Enhanced tunneling across nanometer-scale metal–semiconductor interfaces," *Applied Physics Letters*, vol. 80, pp. 2568–2570, 2002.

[66] W. Wang, T. Lee, M. Kamdar, M. A. Reed, M. P. Stewart, J. J. Hwang, and J. M. Tour, "Electrical characterization of metal–molecule–silicon junctions," *Superlattices and Microstructures*, vol. 33, pp. 217–226, 2003.

[67] L. A. Curtiss, P. C. Redfern, K. Raghavachari, and J. A. Pople, "Assessment of Gaussian-2 and density functional theories for the computation of ionization potentials and electron affinities," *Journal of Chemical Physics*, vol. 109, pp. 42–55, 1998.

[68] G. de Oliveira, J. M. L. Martin, F. de Proft, and P. Geerlings, "Electron affinities of the first- and second-row atoms: Benchmark ab initio- and density-functional calculations," *Physical Review A*, vol. 60, pp. 1034–1045, 1999.

[69] V. N. Staroverov, G. E. Scuseria, J. M. Tao, and J. P. Perdew, "Comparative assessment of a new nonempirical density functional: Molecules and hydrogen-bonded complexes," *Journal of Chemical Physics*, vol. 119, pp. 12129–12137, 2003.

[70] J. P. Perdew and M. Levy, "Comment on "Significance of the highest occupied Kohn–Sham eigenvalue"," *Physical Review B*, vol. 56, pp. 16021–16028, 1997.

[71] J. P. Perdew, R. G. Parr, M. Levy, and J. L. Balduz, "Density-Functional Theory for Fractional Particle Number - Derivative Discontinuities of the Energy," *Physical Review Letters*, vol. 49, pp. 1691–1694, 1982.

[72] R. G. Parr, R. A. Donnelly, M. Levy, and W. E. Palke, "Electronegativity – Density functional viewpoint," *Journal of Chemical Physics*, vol. 68, pp. 3801–3807, 1978.

[73] O. A. Vydrov and G. E. Scuseria, "Ionization potentials and electron affinities in the Perdew-Zunger self-interaction corrected density-functional theory," *Journal of Chemical Physics*, vol. 122, p. 184107, 2005.

[74] J. L. Bahr, E. T. Mickelson, M. J. Bronikowski, R. E. Smalley, and J. M. Tour, "Dissolution of small diameter single-wall carbon nanotubes in organic solvents?," *Chemical Communications*, pp. 193–194, 2001.

[75] K. Kamaras, M. E. Itkis, H. Hu, B. Zhao, and R. C. Haddon, "Covalent bond formation to a carbon nanotube metal," *Science*, vol. 301, pp. 1501–1501, 2003.

[76] S. Banerjee and S. S. Wong, "Selective metallic tube reactivity in the solution-phase osmylation of single-walled carbon nanotubes," *Journal of the American Chemical Society*, vol. 126, pp. 2073–2081, 2004.

[77] R. S. Mulliken, "Electronic population analysis on Lcao-Mo molecular wave functions. 3. Effects of hybridization on overlap and gross Ao populations," *Journal of Chemical Physics*, vol. 23, pp. 2338–2342, 1955.

[78] R. S. Mulliken, "Electronic population analysis on Lcao-Mo molecular wave functions. 4. Bonding and antibonding in Lcao and valence-bond theories," *Journal of Chemical Physics*, vol. 23, pp. 2343–2346, 1955.

[79] R. S. Mulliken, "Electronic population analysis on Lcao-Mo molecular wave functions. 1.," *Journal of Chemical Physics*, vol. 23, pp. 1833–1840, 1955.

[80] R. S. Mulliken, "Electronic Population Analysis on Lcao-Mo Molecular Wave Functions. 2. Overlap Populations, Bond Orders, and Covalent Bond Energies," *Journal of Chemical Physics*, vol. 23, pp. 1841–1846, 1955.

[81] N. P. Guisinger, N. L. Yoder, and M. C. Hersam, "Probing charge transport at the single-molecule level on silicon by using cryogenic ultra-high vacuum scanning tunneling microscopy," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, pp. 8838–8843, 2005.

[82] N. P. Guisinger, R. Basu, A. S. Baluch, and M. C. Hersam, "Molecular electronics on silicon – An ultrahigh vacuum scanning tunneling microscopy study," *Molecular Electronics III*, vol. 1006, pp. 227–234, 2003.

[83] N. P. Guisinger, R. Basu, M. E. Greene, A. S. Baluch, and M. C. Hersam, "Observed suppression of room temperature negative differential resistance in organic monolayers on Si(100)," *Nanotechnology*, vol. 15, pp. S452–S458, 2004.

[84] N. P. Guisinger, M. E. Greene, R. Basu, A. S. Baluch, and M. C. Hersam, "Room temperature negative differential resistance through individual organic molecules on silicon surfaces," *Nano Letters*, vol. 4, pp. 55–59, 2004.

[85] C. A. Richter, D. R. Stewart, D. A. A. Ohlberg, and R. S. Williams, "Electrical characterization of Al/AlOx/molecule/Ti/Al devices," *Applied Physics A: Materials Science & Processing*, vol. 80, pp. 1355–1362, 2005.

[86] G. D. J. Smit, M. G. Flokstra, S. Rogge, and T. M. Klapwijk, "Scaling of micro-fabricated nanometer-sized Schottky diodes," *Microelectronic Engineering*, vol. 64, pp. 429–433, 2002.

[87] G. D. J. Smit, S. Rogge, and T. M. Klapwijk, "Scaling of nano-Schottky-diodes," *Applied Physics Letters*, vol. 81, pp. 3852–3854, 2002.

[88] J. M. Seminario and J. M. Tour, "Ab initio methods for the study of molecular systems for nanometer technology: Toward the first-principles design of molecular computers," *Molecular Electronics: Science and Technology*, vol. 852, pp. 68–94, 1998.

[89] J. M. Seminario and L. M. Yan, "Ab initio analysis of electron currents in thioalkanes," *International Journal of Quantum Chemistry*, vol. 102, pp. 711–723, 2005.

Molecular and Nano Electronics: Analysis, Design and Simulation
J. M. Seminario (Editor)

Chapter 2

Bio-molecular devices for terahertz frequency sensing

Ying Luo,[a] Boris L. Gelmont[a], and Dwight L. Woolard[b]

[a]*Electrical and Computer Engineering Department, University of Virginia, Charlottesville, Virginia 22904-4224, USA*
[b]*U. S. Army Research Laboratory, Army Research Office, PO Box 12211, RTP, North Carolina 27709-2211, USA. dwight.woolard@us.army.mil*

The rapidly developing field known as nanotechnology suggests that there is a vast potential for revolutionizing electronic systems as they are known today. The possibilities for enormous increases in functionality, integration, and speed offered by molecular-level systems have motivated a detailed investigation of nanoscale elements and components. Major scientific and engineering research challenges must be addressed today before a robust and fully functional nanoscale capability can be realized in the future. Hence, entirely new scientific approaches and innovations are clearly necessary for bridging the intellectual and technology gap between the nanoscopic and the microscopic where the full advantages of ultra-small, ultra-dense and ultra-complex electronic systems can be realized. One specific avenue for realizing the technological payoffs available at the nanoscale is to incorporate aspects of architecture and algorithmic functions that already exist in nature. Indeed, many biological molecules (e.g., DNA) demonstrate phenomenal levels of functionality and efficiency and may hold the ultimate key to revolutionary capabilities within human-engineered systems of the future.

While many fantastic molecular-level device concepts have been introduced in recent years, much fundamental research work remains to elevate biomolecular-based devices and systems to a practical and useful level. This chapter presents new results from an ongoing research effort [1–3] that seeks to define a new architectural paradigm whereby enhanced sensing of terahertz (THz)-frequency biological-signatures (bio-signatures) may be achieved. Here, the basic methodology is to incorporate the target biomolecular agents directly into the electronic architecture and to utilize their THz-frequency characteristics directly into the function of the sensor platform. This chapter will illustrate how new sensor platforms can utilize combinations of electrical (conductive), THz, and optical – or Electro-THz-Optical (ETO) – based communication and control to refine the process of collecting THz bio-signatures and to define approaches for increasing the available number of THz spectral features through a procedure that will be referred to as Multi-State Spectral Sensing (MS^3). The introduction of MS^{3}is of significant importance because the number of spectral signature features associated with any individual bio-agent in its natural state (i.e., ground state) is somewhat limited and new methods are needed to provide for adequate levels of agent discrimination [1–4].

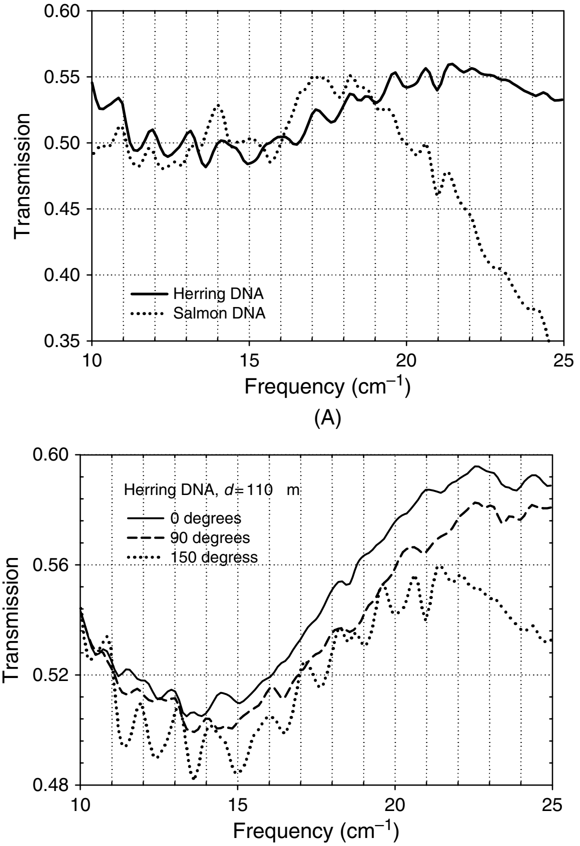
The first step in the realization of the proposed integrated sensor architecture is a detailed theoretical description of the target bio-molecules, and the associated ETO-based processes, that are required for defining communication and control at the nanoscale. Hence, this chapter will present modeling and simulation results for an initial set of bio-molecules (e.g., butene and retinal) that (a) illustrates a procedure through which optical excitation can be used to control the molecular conformational state, and (b) accurately predicts the THz spectral characteristics in both the natural (i.e., ground) state and the final metastable state. Therefore, this research is establishing a new scientific foundation of knowledge which can serve as a blueprint for enabling integrated bio-molecular sensing platforms of the future with enhanced capabilities for sensing and processing THz bio-signatures.

1. THz sensing science motivation and requirements

The new concepts for a biomolecular-based sensing architecture presented in this chapter were motivated almost exclusively by prior research into the utilization of THz-frequency spectroscopy as a potential tool for the detection of biological materials and agents. Indeed, by 1995, our group initiated new scientific studies into the fundamental interactions of THz radiation with biological materials at the molecular level [5]. This has been followed by a continuous and focused effort to establish a new THz sensing science and technology base that can be used to assess the detection, identification, and characterization capability of THz spectroscopic analysis [6, 7]. Obviously, these activities were made a high priority because THz sensing has important potential relevance to both military and homeland defense against bio-based threats. However, the spectral probing of the fine structural and electronic characteristics of organic molecules also has the potential to provide new insights into microscopic biological systems and therefore may also have broad ramifications to biological and medical science in the future [1, 8, 9]. Important details regarding this THz sensing research and the influence on guiding the definition of the integrated sensor methodology will now be briefly discussed.

1.1. THz sensing science issues for bio-systems

Scientific results generated from an ongoing research project conducted by our group [10–18] have already demonstrated the potential use of THz-frequency transmission spectroscopy as a technique for the general characterization of biological agents. Under initial exploratory research, experimental measurements were carried out and theoretical models were developed that enabled detailed investigations into the submillimeter-wave (i.e., approximately <1 THz, which is equal to 1000 GHz) spectra of DNA macromolecules [19]. While the THz-frequency regime had been predicted to be fairly rich with spectral features of DNA phonon modes that arise out of poorly localized motions spread over one or more base-pair units [20, 21], the new experimental studies performed by our group [15, 16, 19, 22] were the first to confirm the presence of this phenomenon in this spectra regime (i.e., ~0.3–3 THz). These long-wave absorption features, of the type as illustrated in Figure 1(A), where the THz spectra of Salmon DNA is compared to Herring DNA, were shown to be intrinsic properties of biological materials that arose as a result of phonon activity. In addition, DNA films prepared such that the molecular chains were partially aligned and oriented were shown to exhibit a significant dependence on the polarization of the measurement field (see Figure 1(B)). It is clear that these results are important because they reveal fundamental frequency-dependent properties of bio-molecules that are arising at the microscopic level. In particular, these results suggest that bio-electronic components might be used to selectively filter and control the transmission of electromagnetic signals with carrier frequencies in the THz band.



(A)

(B)

Figure 1 (A) Absorption spectra for Salmon & Herring DNA. (B) Spectral enhancement from Herring DNA sample with structural alignment [19]

In an effort to quantify various aspects of THz-frequency sensing, our prior research [19, 22] was used to show that unique submillimeter-wave resonance absorption features arise in both bio-molecular samples (e.g., DNA) and complete cellular materials (e.g., bacteria spores). These spectral signatures verified that the THz regime is potentially useful for the study, analysis, and identification of biological macromolecules when placed under highly controlled laboratory conditions. In addition, more recent system-level investigations conducted [1, 23, 24] suggested that remote detection of bio-particles is also viable using THz-frequency signatures. However, while adequate levels of sensitivity appear to be demonstrated even for remote detection applications, the viability of THz spectroscopy for biological sensing (i.e., point and remote) will ultimately hinge on the level of reliable discrimination it can provide. Two challenges arise within this context. First, there are a reasonably limited number of spectral signatures (i.e., $< \sim 100$) associated with any bio-agent in its natural state and, second, the strong atmospheric absorption limits the sensitivity of the approach (i.e., at remote-sensing ranges) at all but a few THz-band transmission channels. Therefore, it appears that THz spectral sensing techniques will not become practical unless new probing techniques are identified that allow for (1) introducing controlled perturbations at the microscopic level that increase the amount of phonon-mode information which is specific to the target bio-molecule, and (2) the precise extraction of all available dynamical phonon-mode characteristics.

While this might at first appear to be an insurmountable challenge, details regarding the phenomenology of bio-systems, and their response to external influences, can be used to define electronic architectural paradigms for achieving the prescribed goals.

1.2. Multiplication of THz phonon mode information in bio-systems

Based upon the discussion above, it may be argued that the main bottleneck to the effective utilization of THz spectral sensing is a technological limitation for interfacing to dynamical processes at the molecular level. For example, the THz-frequency phonon modes that arise in macromolecules will be strongly dependent on the specific type of bonding present and molecular structure. A number of research groups have already investigated relationships between bio-molecular conformation and the interactions of very long-wavelength radiation. This includes the use of microwave reflection to sense the thermal modulation of protein folding [25, 26] and the application of THz spectroscopy for sensing conformational changes in a number of bio-molecules (e.g., DNA, lysozyme, myoglobin, bacteriorhodopsin) as a function of temperature, hydration, and photoexcitation [27], just to name two examples. It is apparent from these examples that external stimulus can be used to alter the microscopic structure of bio-molecules such as to effect changes in the THz spectral signatures. Therefore, if an appropriate electronic interface could be defined, then in theory it should be possible to modify molecular structure such that new phonon mode information becomes available for characterization purposes. Fortunately, one can hypothetically envision nanoscale, integrated architectures where strategically designed electronic interactions are used to modify the state of embedded bio-molecules. If such controlled interfacing was possible, then the integrated system itself could be used to modify the electronic and/or conformational state of a target bio-molecule, and to define methodologies for extracting the spectral signature information. In addition, if numerous excited- or metastable-states could be defined and utilized in integrated architectures, then the amount of THz spectral information could be significantly multiplied.

1.3. Acquisition of THz phonon mode information from bio-systems

It is widely accepted that the phonon absorption characteristics obtained from biological materials will contain *weak and broadened* resonant features. Indeed, almost all the published experimental transmission and reflection spectra taken from solid-state biological samples routinely show extremely limited variations (i.e., ~1–5%) in absorption versus frequency characteristics within the THz regime [7]. Furthermore, before the very recent experimental demonstrations of THz signatures, the scientific community generally assumed that THz vibrations in weakly bonded bio-systems would possess low-level polarizability and be severely over-damped at very long wavelengths. However, a number of experimental results obtained by our group and others suggest that the interaction dynamics at the microscopic level may be significantly underestimated by many of the measurements made on concentrated macro-sized bio-samples. Specifically, it is important to note that most measurements made to date on bio-materials utilize large samples that possess mesoscopic structure (i.e., in between macroscopic and microscopic. As will now be discussed and demonstrated, macroscopic measurements can fail at revealing the true dynamical characteristics at microscopic levels.

As an initial example, consider the results given in Figure 2 that compares the experimentally derived absorption coefficient taken from a thin (i.e., 6.6 μm thick) sample of short-chain (12 base pairs) double-stranded Poly[C]-Poly[G] RNA to modeling results for equivalent layers of the isolated molecules. Here, two different values of molecular dissipation have been considered (i.e., $\gamma = 0.5$ and $1\,\mathrm{cm}^{-1}$) in the simulations. The

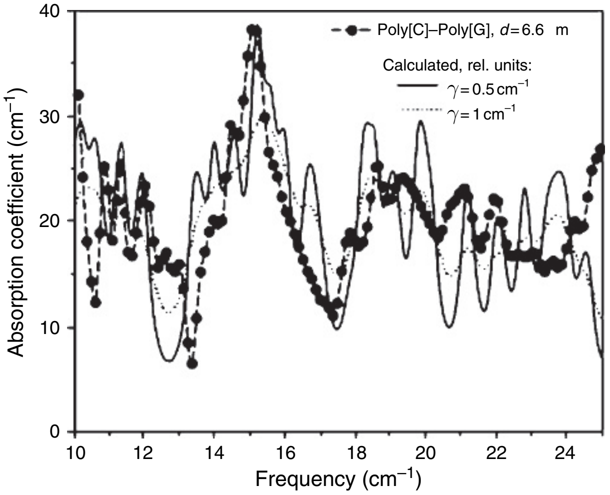


Figure 2 Experimental absorption coefficient taken from 6 μm thick Poly[C]–Poly[G] RNA films compared to modeling prediction for $\gamma = 0.5$ and $1.0\,\mathrm{cm}^{-1}$ [15]

experimental studies utilized free-standing dry films of the RNA material that were produced by drying gel solutions using a previously documented process [28]. Here, one can observe very good agreement between the experimentally derived spectral signature at 15 cm^{-1} (i.e., frequency location and magnitude) and that predicted by theory when a dissipation value of $\gamma = 0.5\,\mathrm{cm}^{-1}$ is employed. Hence, this experimental result is strong evidence for the existence of a very strong and optically active phonon resonance that is relatively isolated in frequency from any other modes. Most importantly, the intensity of this particular mode was found to be strongly dependent on the thickness of the sample used in the spectroscopic measurement. In fact, when the sample thickness is increased to 30 μm and above, the peak-to-valley variation in the phonon mode intensity was observed to decrease by a factor of 10, as shown in reference [28]. The strong dependence of absorption intensity on film thickness immediately suggests some type of mesoscale interaction effect. Specifically, these results suggest that interactions that are occurring at scales much larger than the size of the RNA molecules (i.e., perhaps between large clustering of molecules) to be responsible for leading to the very weak and broadened spectral signatures which are usually obtained from macroscopic measurements. Furthermore, one might speculate that the underlying molecular dynamics are best represented by the results obtained from the very thin samples given in Figure 2. If this theory is true, then there is a great opportunity at the nanoscale for gathering precise THz signature information associated with bio-molecules that will be useful for sensing and characterization applications.

As another example, consider the measured transmission results given in Figure 3(A) that were obtained from Salmon DNA gel samples that were produced by dissolving the DNA into distilled water at a mass ratio of 1:18, Salmon to water.

Here, the measurements were taken from gel samples of thickness ~100 microns that were prepared between two very thin layers (i.e., 13 microns) of polyethylene film. These results contain very sharp spectral features that may be attributed to absorption by the Salmon DNA in the solution. While one should expect influences on the DNA molecules by the surrounding solution (e.g., water mass loading) that will perturb the absorption spectra, it is clear that the relative intensity of the spectral signature are much more pronounced when the bio-molecules under test are subject to a less confining and rigid environment (i.e., in comparison to the results given in Figure 1 for dry films). This viewpoint is further supported by Figure 3(B) where strong peak-to-valley variations in a Herring DNA gel (initially prepared at a 1:12 ratio) are seen to persist (and even intensify) as the water content evaporates over time and the background absorption is reduced. While these results are suggestive of strong dynamical activity in semi-isolated bio-molecules, results recently published by another group provide much more compelling evidence [29]. Indeed, experimental studies performed upon bio-particles (i.e., bacterial endospores) have generated evidence that if the samples are diluted (i.e., such that individual spores are isolated from one another) then the resulting THz spectral signatures are greatly enhanced.

This section has presented credible experimental results that suggest there are opportunities at the nanoscale for the precise collection and controllable multiplication of THz-frequency bio-signatures. The implementation of integrated systems that can take advantage of this sensing modality will require the identification of novel methods for interfacing to bio-molecules at the microscopic level. The next section will introduce a new concept for a biomolecular-inspired electronic architecture for enhanced THz

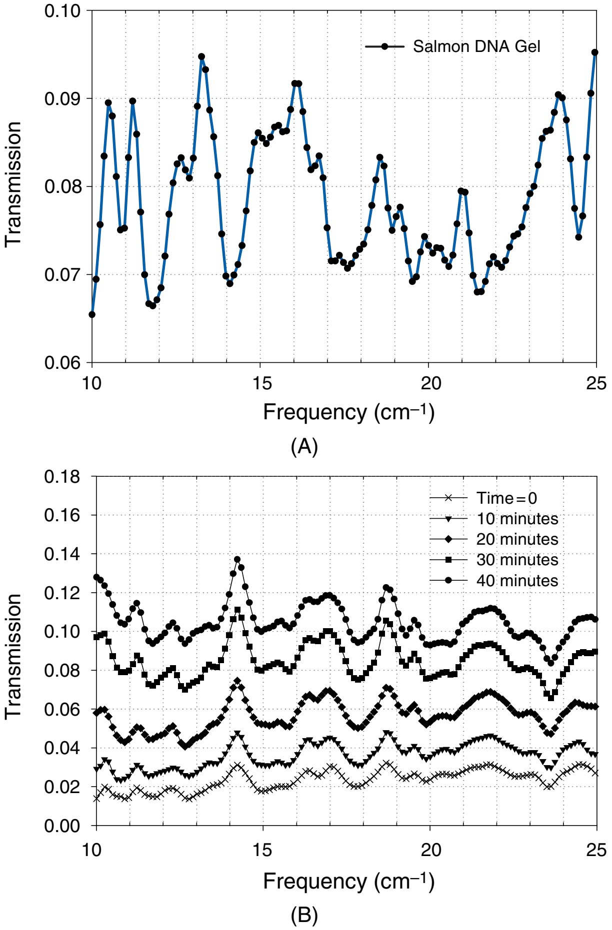


Figure 3 (A) Absorption spectra for Salmon DNA gel prepared in 1:18 mass ratio with water, (B) Absorption spectra from Herring DNA gel (1:12 mass ratio with water) as a function of drying time [1]

sensing and present a theoretical analysis for some prototype bio-components that will be needed for implementing this type of system.

2. Bio-molecular architectural concept and bio-component studies

The novelty of the bio-inspired architecture to be presented here lies in the strategic use of integrated biological elements to achieve higher-level function and spectral data processing within a nanoscale and molecular-level architecture. As discussed at length in Section 1, fundamental absorption/emission properties present in known biological materials (e.g., DNA, RNA, etc.) can provide new insight for a novel approach to nanoscale device functionality and integrated molecular-level sensing and data processing. Specifically, DNA and RNA macromolecules have been shown to exhibit spectral absorption characteristics with multiple absorption peaks that might be used to selectively filter and control transmission frequency channels at very long wavelengths. In addition, the absorption characteristics of such bio-molecules are known to be strongly dependent on molecular conformations. As all bio-molecules can be elevated to excited-state conformations through optical or THz-frequency excitation, this process can be used to define ETO molecular function. Therefore, it should be possible to utilize existing and the future artificially designed bio-molecular elements to realize optical or THz-frequency controlled filters of long wavelength electromagnetic signals.

2.1. *Bio-molecular inspired architecture for sensing*

A very simple example of the proposed concept is illustrated in Figure 4 where a *hypothetically* tunable DNA-based filter (i.e., the THz-frequency spectral absorption peak is influenced by the photonic emission at frequency f_2 as depicted in the inset) can be used to establish feedback, and as will be shown later, clocking/register function. As conceptualized here, the ETO-based architecture utilizes two emission sources (i.e., at frequencies f_1 and f_2), a single detector (i.e., sensitive to radiation at frequency f_1) and a direct-current driven circuit. This particular bio-electronic element, as defined, will allow for defining functionality through the coupling of multiple-frequency channels that in turn control a direct-current pulse that periodically flows through the circuit connecting detector D_1 to emitter E_2. One important advantage of this architecture is that it is possible to realize gating and feedback within densely packed bio-molecular components while at the same time providing for isolation between other system elements (i.e., by

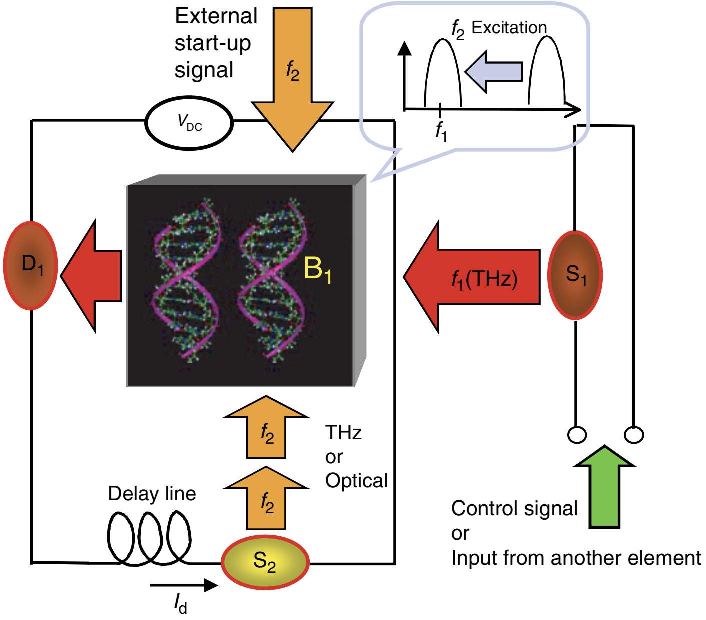


Figure 4 An electro-THz-optical architecture for illustrating functional bio-sensing [1]

utilizing alternative transmission frequencies for nearby units). Also note that these types of bio-molecular components might also be integrated with other inorganic molecules and to traditional semiconductor devices to enable functional control and data processing at the nanoscale. Most importantly, as this basic architecture is intrinsically linked to species-specific spectral absorption features, these bio-electronic systems can be engineered into sensing arrays that detect the presence of target bio-agents. This approach also solves a critical sensing problem, as this type of bio-architecture allows for the nanoscale modification of the electronic and/or conformational state of the target agent – here it is assumed that the target molecule will be captured into the system via engineered ligand bonding sites. Therefore, this bio-electronic architecture allows for a new MS^3 approach that can expand the amount of available spectral signature information and greatly improve detection, identification, and characterization capabilities.

The control element of the nano-circuit defined in Figure 4 is to utilize the ETO characteristics of bio-molecules – e.g., the THz-frequency spectral absorption will be modified through optical excitation. For example, consider an arbitrary DNA fragment that exhibits a conformation and typical absorption profile around a phonon-induced resonance as shown in Figure 5(A). Here, the resulting spectral absorption is defined by the ground-state conformation of the molecule as shown in the top of Figure 5(A). When subjected to an adequate external excitation (i.e., energy and polarization), the molecule can experience a charge (or potential) redistribution that is accompanied by changes in the conformation and spectral characteristics. If the proper modeling tools are available, it should be possible to identify naturally occurring or artificially engineered bio-molecules that under proper excitations assume excited-state conformations with modified spectral characteristics of the type as shown at the bottom of Figure 5(A). Indeed, simulation studies [30] on small test molecules have already been used to demonstrate this basic effect in the *cis*-2- and *trans*-2-isomers of butene. Figure 5(B) shows these two isomers (i.e., which can occur via a rotation around the center bond) along with the discrete (no broadening) THz-frequency absorption spectrum which displays the dramatic changes of the type needed for defining the new bio-molecular architecture. This ETO effect will allow for defining multi-frequency communications and for defining bio-molecular architectures that enable logic and signal processing type functionality.

2.2. Bio-component modeling and simulation

The successful implementation of bio-molecular architectures of the type described in the last sub-section requires the identification of bio-molecules that can be controllably induced (i.e., through optical and/or THz radiation) into alternative geometric conformations. Furthermore, these molecules must exhibit significant variations in their THz-frequency absorption characteristics, as to make the approach useful for processing and collecting bio-signatures. The sub-sections that follow will present a theoretical analysis of two bio-molecules that demonstrate these required ETO-based characteristics. Specifically, the investigations that follow will consider the optically induced isomers of butene and retinal. Here, the results will show that optical control of spectral characteristics is possible, with the smaller butene molecule exhibiting infrared absorption signatures and the larger retinal molecule that demonstrates similar properties in the THz regime.

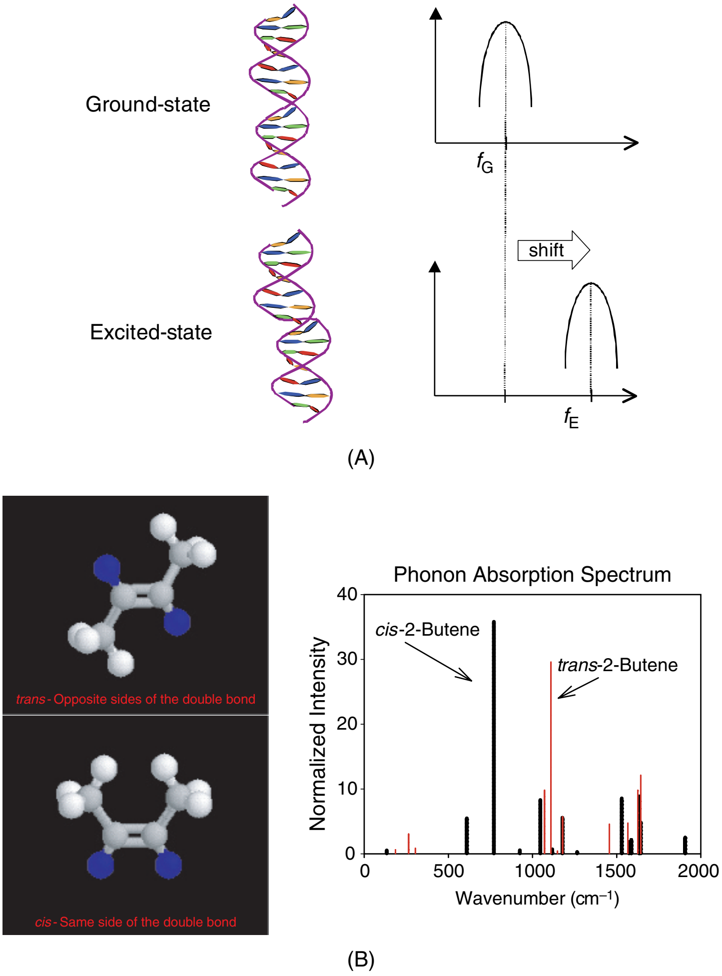


Figure 5 Basic ETO effect: (A) hypothetical DNA example, (B) isomers of butane and simulated spectra [1]

2.2.1. Simulation and analysis of butene isomers

In the initial study, the *cis*-*trans* isomers of butene have been chosen because they are very well known and simple examples of a ground- and metastable-state pair with significantly different geometries. The *cis*-*trans* isomers are geometric isomers, a type of stereoisomerism in which atoms or groups display orientation differences around a double bond, such as *trans*-2- and *cis*-2-butene as was shown earlier in Figure 5(B). This work demonstrated molecular isomers that yield different sets of vibrational frequencies. In this case to be studied, photo-induced transitions bring about the conversion from one geometric shape (i.e., *trans*-2-butene) to another (i.e., *cis*-2-butene) through rotation about a double bond and the conversion is also called isomerization.

The analysis presented here includes calculations of the optimized energies, vibrational frequencies, and infrared intensities that were carried out at the Hartree–Fock

(HF) level within the split valence polarized 6-31G(d) basis set in the Gaussian 98 package [31]. The energies of the excited states were found within the single configuration interaction approach (CIS), where one models the first excited state as combinations of a single substitution from the highest occupied molecular orbital (HOMO) to the lowest unoccupied molecular orbital (LUMO). Vibrational frequencies were then computed by determining the second derivatives of the potential energy with respect to the Cartesian nuclear coordinates and then transforming to mass-weighted coordinates. Due to the fact that electron correlation is neglected, the frequencies computed using the HF approximation are known to be overestimated by approximately 10–12%. Furthermore, because a medium-sized basis set was used, the derived values can be expected to deviate even more from experiments, i.e., by approximately 15% in total [32]. Therefore, a scaling was performed on the originally calculated frequencies by an empirical factor of 0.893 to eliminate known systematic errors in the physical model.

To monitor the process of isomerization from *trans*-2-butene to *cis*-2-butene, the torsional angle θ, defined as the dihedral angle of the plane of C1–C2–C3 and the plane of C2–C3–C4 around the double bond in the C1–C2=C3–C4 chain associated with the molecule given in Figure 6(A), was taken to be the reaction coordinate. The potential energy (PE) curves associated with the isomerization process for the ground (S_0) and the first excited singlet state (S_1) of 2-butene are shown in Figure 6(B), along

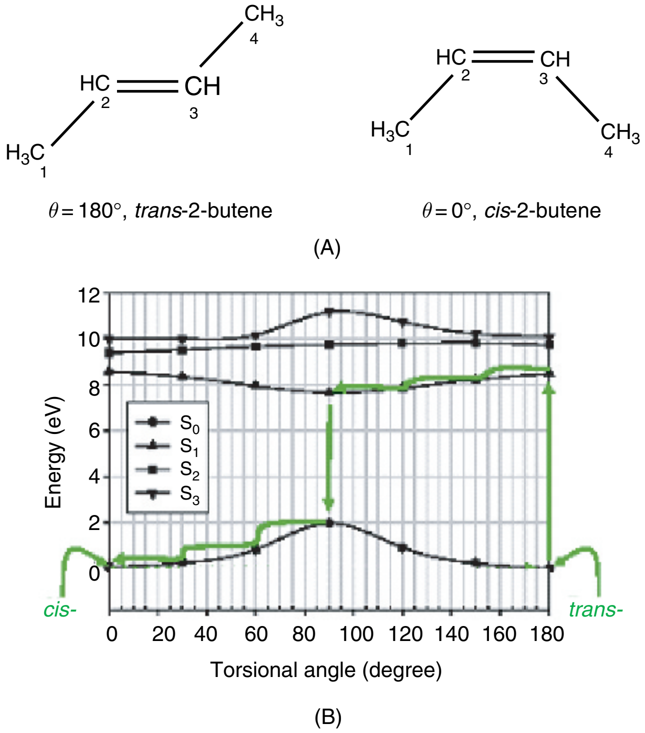


Figure 6 (A) Geometrical structure of *trans*-2-butene and *cis*-2-butene; (B) Potential energy curves associated with the isomerization process which involve the ground-state, S_0, and the first excited-state, S_1. Arrows depict the energy-path of the molecules

with the next two excited states S_2 and S_3. Here it is important to note that the *trans*-2 geometry (i.e., $\theta = 180°$) is the natural-state and the *cis*-geometry (i.e., $\theta = 0°$) is the metastable-state. Without external excitations, the large PE barrier 1.98 eV in the ground state prevents the transition from *trans*-2-butene to *cis*-2-butene. It is also clear from Figure 6(B) that there is no barrier in the first excited singlet state S_1, which is ideal for an ultra-fast switching between the *trans*- and *cis*- isomers.

Potential energy curves can be used to illustrate this optically induced isomerization process by following the superimposed arrows. The process begins with a required optical excitation (~8.46 eV, 146 nm) of the "electronic-state" of the *trans*-geometry from S_0 to S_1. This is followed by a non-radiative decay to the S_1 PE valley minimum, which corresponds to a 90° rotation about the reaction coordinate. At this point, the molecule undergoes an "electronic" radiative decay from S_1 to S_0. This is followed by a second non-radiative decay to the *cis*-geometry, which corresponds to a second 90° rotation about the reactive coordinate. At this point, butene will remain in the metastable *cis*-2 geometry until thermal relaxation of the system back to the ground state.

To estimate the probability that the excited *trans*-2-butene will follow the isomerization process described above, instead of relaxing back to its own ground state, we investigate the molecular dynamics of butene in the first singlet excited state S_1 by combining the theory of Newton's Dynamics and standard quantum chemistry software Gaussian 98. We began the investigation from *cis*-geometry ($\theta = 0°$) and calculated the time for the excited *cis*-geometry to achieve the energy minimum ($\theta = 90°$) and the excited trans-geometry ($\theta = 180°$) of S_1. The calculation procedure is as follows: (i) we calculated the potential energy of the ground state and the excitation energy of S_1 of the optimized *cis*-butene using Gaussian 98 package and by adding the two energies we got the potential energy of the excited *cis*-butene. Note that we took this potential energy as the total energy of butene (i.e. we assume the kinetic energy of excited cis-butene is 0). To initiate the rotation, we actually started from a very small θ with a very small initial velocity. (ii) For every 10°, we calculated 20 points (i.e., we performed Gaussian 98 calculation every 0.5°). For every point, we calculated the ground-state energy without optimization and the corresponding excitation energy; thus we got the potential energy of the excited state. By subtracting the potential energy from the total energy, we got the kinetic energy. Then we used the energies of 20 points to calculate the time for butene to rotate 10° by integrating Newton's equations.

The relation between the potential energies of excited states and time is given in Figure 7, and for demonstrative purposes, the PE curves of the ground state S_0 is presented on the same graph. We can see that the time for excited *trans*-2-butene to relax to the valley minimum is about 100 fs and symmetrically it takes almost the same time to go from the minimum to *cis*-2-butene. On the other hand, according to the spontaneous emission theory the time for excited *trans*-2-butene to relax back to ground state is about 5.9×10^3 s. So the isomerization process will be significantly probable. In Figure 8, we demonstrate how the dihedral angle of C1–C2=C3–C4 double bond, or the geometric structure of butene, will change with time.

The lowest vibrational frequencies and the associated IR intensities that were calculated for *trans*-2-butene and *cis*-2-butene are compared in Figure 9. These results indicate a significant difference in spectral signatures of the two molecular conformations and one that has the general quantitative characteristics needed by the previously discussed bio-molecular architecture.

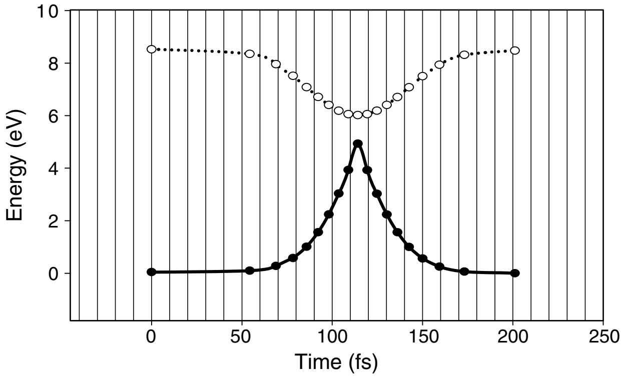


Figure 7 Molecular dynamics of butene: Energies of first singlet excited state vs. time

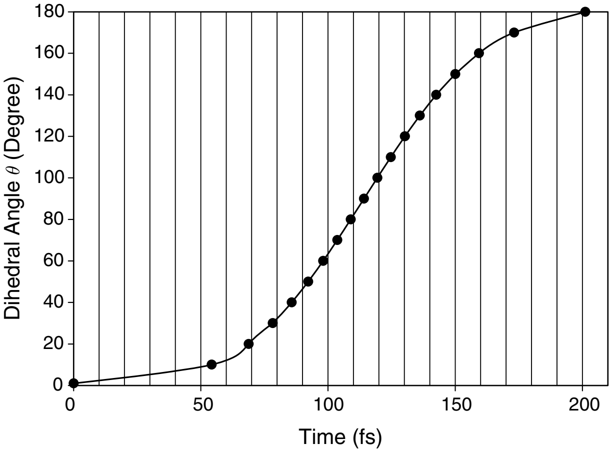


Figure 8 Molecular dynamics of butene: Dihedral angle vs time

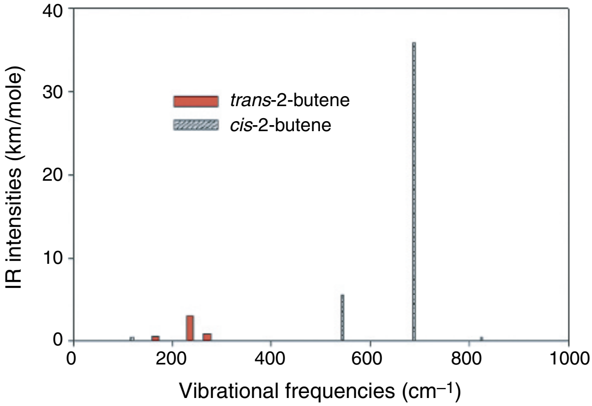


Figure 9 The calculated characteristics of the lowest spectral signatures of *trans*-2-butene and *cis*-2-butene [1]

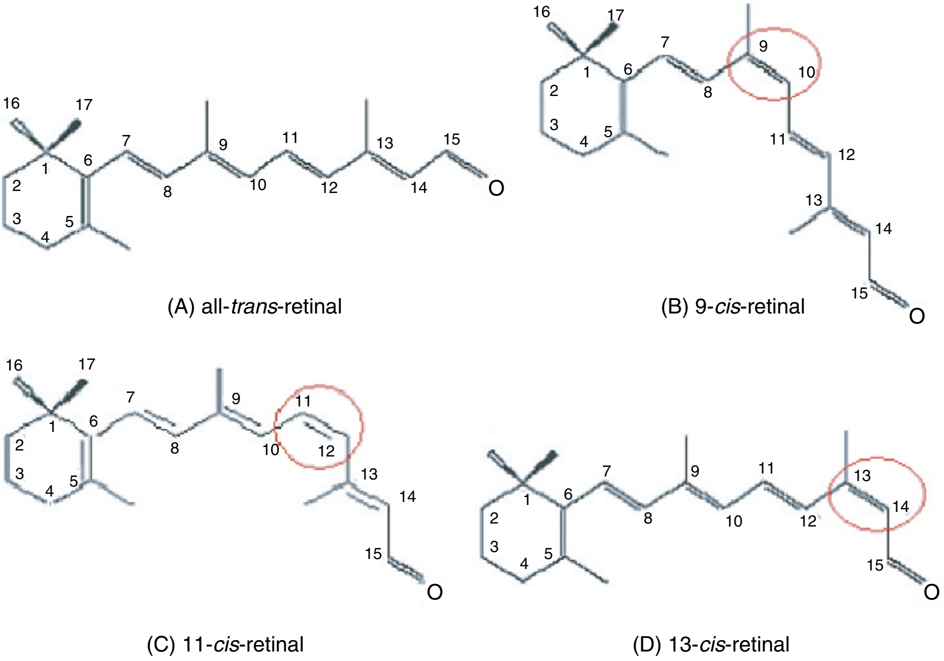
Table 1 A comparison of calculated spectral signatures with measurements of McKean *et al.* [33]

cis-2-butene		*trans*-2-butene	
McKean *el al.* [33]	Calculated in this work	McKean *et al.* [33]	Calculated in this work
566	544.2	239	236.0
686	687.8	284	271.2
865	934.2	964.3	956.8
970.9	997.3	978.8	990.7
1009.1	1051.6	1046.8	1025.4
1036.9	1129.6	1063.6	1052.9
1268.5	1366.3	1303	1300.6
1383.8	1405.1	1382.2	1398.9
1408.4	1417.3	1447.4	1451.9
1444.5	1458.1	1461	1466.9
1456.2	1460.2	2835.6	2848.9
1458.9	1466.9	2871	2890.0
1669	1702.9	2890	2890.1
2893.2	2852.9	2932.5	2917.6
2900.8	2855.8	2949.4	2957.3
2914.5	2891.9	2973	2961.9
2929.5	2932.5		
2947.5	2951.4		
2954.1	2955.6		
2979.2	2979.8		

Table 1 compares the calculated frequencies to prior experimental measurements of McKean *et al.* [33]. One can see that all the absorption frequencies calculated in this work are reasonably-good agreeable (i.e., within ~3%) with the experimental results.

While this study yielded qualitative bio-molecular function of the type needed for the proposed architecture, the required excitation from S_0 to S_1 is far too large (i.e., ~8.5 eV) to be practical, and the associated spectral bio-signatures are well above the THz regime where species-specific information is expected to be present. Hence, the alternative molecule retinal will now be considered, which is known to be visible-frequency light sensitive, and which is more complex and can be expected to yield much lower frequency vibrational modes.

All-*trans*-retinal is the chromophore of bacteriorhodopsin, the light transducing protein in the purple membrane of *Halobacterium salinarium*. The isomerization from all-*trans* to 13-*cis* conformation takes places very quickly in approximately 500 fs [34] induced by absorption of a 568 nm photon [35]. 11-*cis*-retinal is the chromophore of rhodopsin. The ultra-fast (in ca. 200 fs [34]) photoisomerization from 11-*cis*-retinal to all-*trans*-retinal conformation has been elucidated as induced by absorption of a 498 nm photon [35]. Besides 11-*cis*-retinal and 13-*cis*-retinal, 9-*cis*-retinal is also a geometric isomer of all-*trans*-retinal and derives from the rotation of all-*trans*-retinal around C9=C10 double bond by 180° although 11-*cis*-retinal and 13-*cis*-retinal derive from the rotation of all-*trans*-retinal around C11=C12 and C13=C14 double bonds.



(A) all-*trans*-retinal

(B) 9-*cis*-retinal

(C) 11-*cis*-retinal

(D) 13-*cis*-retinal

Figure 10 Geometric structures of retinal isomers

The geometric structures of all-trans-retinal and its isomers are shown in Figure 10. The retinal has two properties which make it very interesting for our studies. First, the electrons in retinal readily absorb photons in the visible range of wavelengths (400–700 nm) which means the molecule can be optically induced into excited-states. Second, as this study will show, the lowest vibrational frequencies are in THz region.

2.2.2. Simulation and analysis of retinal isomers

The vibrational spectra were calculated at Hartree–Fock method within split valence polarized 6-31G(d) basis set using Gaussian 98 package. All calculated frequencies were multiplied by empirical factor 0.893 to eliminate the known systematic errors. The very far infrared (FIR) spectra (under 30 cm^{-1}) of four isomers are shown in Figures 11–13.

These results illustrate that the various metastable-state conformation yields spectra fidelity (e.g. 22 cm^{-1} in 9-*cis*, 12 cm^{-1} in 11-*cis*, 26 cm^{-1} in 13-*cis*) that can be used in defining multiple communication channels and for realizing the MS^3 approach to expand the amount of bio-signature information. We compared our theoretical vibrational frequencies with the density functional calculations of Gervasio [36], and also with the experimental data measured by THz time-domain spectroscopy at 10 K [37] and by FTIR spectrometer at 15 K [36]. As shown in Table 2, the frequencies predicted in our simulation are in very good agreement with the calculations of Gervasio and comparable to experimental data above 40 cm^{-1}. It is noteworthy that vibrational modes under 40 cm^{-1} exist in both simulations but are not available in experimental

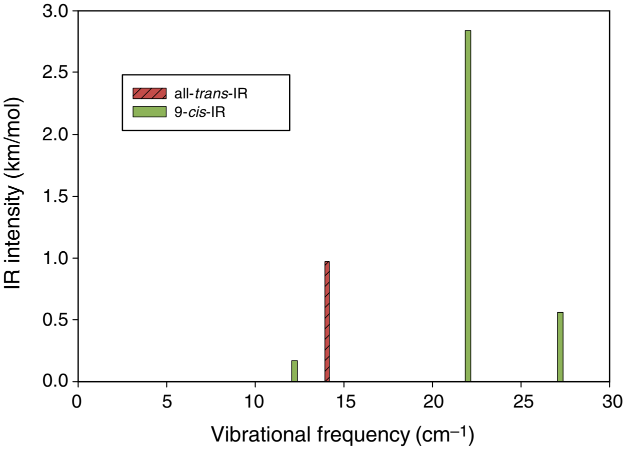


Figure 11 FIR spectra of 9-*cis*-retinal and all-*trans*-retinal ($<30\,\mathrm{cm}^{-1}$) [2]

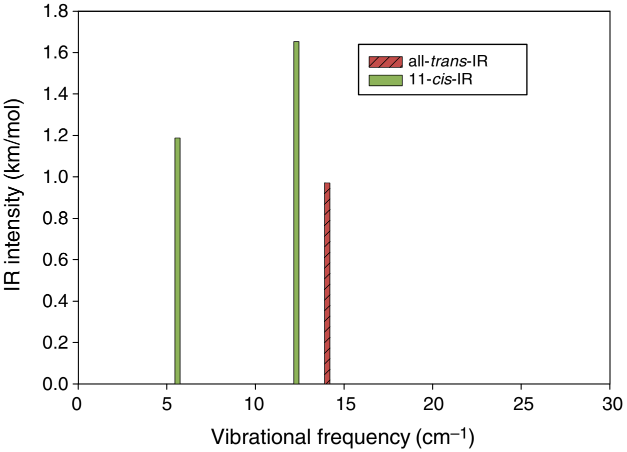


Figure 12 FIR spectra of 11-*cis*-retinal and all-*trans*-retinal ($<30\,\mathrm{cm}^{-1}$) [2]

measurements. This can be explained by the experimental difficulties in the range of $40\,\mathrm{cm}^{-1}$ and below.

This research effort also generated results for the potential energy barriers between the ground state and metastable-state conformations (which are important for specifying switching times) and the electronic excited-state energies (which are needed for specifying the necessary optical excitation).

Calculations of the potential energy curves (PEC) in the ground-state S_0 were carried out at both HF theory and density functional theory (DFT) levels with the split valence polarized 6-31G(d) basis set using the Gaussian 98 package [31]. The B3LYP density functional was used, which is Becke's three-parameter hybrid functional using the Lee, Yang, and Parr (LYP) correlation functional [38]. We calculated the energy of 9-*cis*-retinal, 11-*cis*-retinal, 13-*cis*-retinal and all-*trans*-retinal isomers and simulated the rotation of all-*trans*-retinal around C11=C12, C13=C14, and C9=C10 double

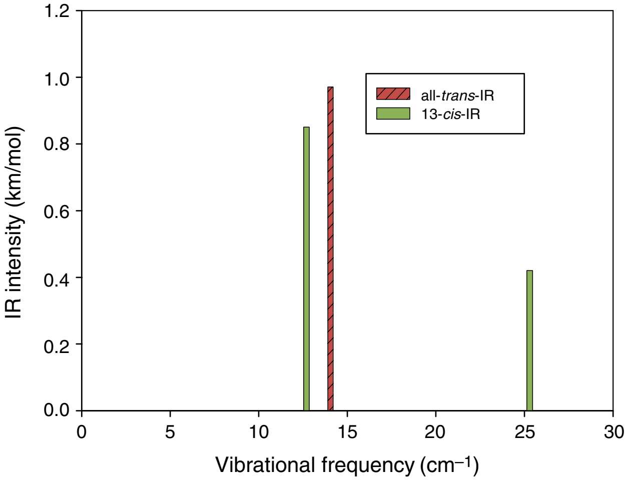


Figure 13 FIR spectra of 13-*cis*-retinal and all-*trans*-retinal (<30 cm^{-1}) [2]

Table 2 Vibrational frequencies of isolated retinal isomers in the range of FIR spectra (cm^{-1})

All-*trans*-retinal				9-*cis*-retinal		11-*cis*-retinal	13-*cis*-retinal	
Ref. [37] exp	Ref. [36] exp	Ref. [36] cal	This work	Ref. [37] exp	This work	This work	Ref. [37] exp	This work
		17	14.1		12.2	5.6		12.7
		29	30.0		22.0	12.3		25.3
		35	30.8		27.2	30.5		30.3
46.6	46	49	42.7	43.5		33.7	41.7	47.7
54.4	53	59	57.5	53.9	54.0	54.4	54.8	58.5
61.0					59.9	62.7	61.0	
66.2	64						64.8	
69.3	68	78	75.5		81.1		68.6	75.2
90.6	89	99	96.0		92.8	94.2		98.8
			100.0					

Table 3 Energy barrier in the PECs of isolated retinal in the ground state (eV)

	C9=C10		C11=C12		C13=C14	
	HF	B3LYP	HF	B3LYP	HF	B3LYP
Barrier between 0° and 90°	1.8360	1.4334	1.4878	1.1859	1.6728	1.2974
Barrier between 90° and 180°	1.8578	1.4634	1.7517	1.4035	1.6946	1.3274

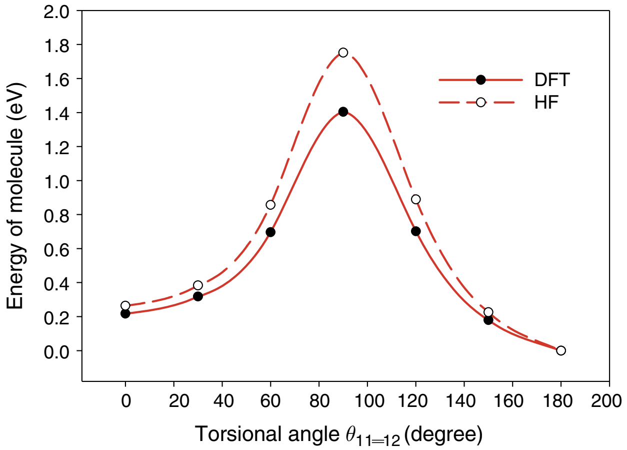


Figure 14 PECs of all-*trans*-retinal rotating around C11=C12 double bond calculated at HF and DFT level

bonds. We took the dihedral angles of the corresponding double bonds $\theta_{C11=C12}$, $\theta_{C13=C14}$ and $\theta_{C9=C10}$ as the reaction coordinates which were fixed during the optimizations, although all other parameters were free. The PECs of the ground state of rotation around C11=C12 double bond calculated at both HF and DFT levels are shown in Figure 14. According to our calculations, the results at the DFT level qualitatively agree with those at the HF level although they achieve better accuracies by including electron correlation. Also, both results predict significant barriers in the ground state that prevent the transition from one isomer to another without external excitation.

In addition, PECs of the ground state of rotation around C9=C10, C11=C12, and C13=C14 double bonds at DFT level are shown in Figure 15. It is clear that all-trans-retinal has the smallest energy among all retinal isomers and is the most stable conformation of retinal. 11-cis-retinal, 13-cis-retinal and 9-cis-retinal are local minima on the PECs which means they are metastable conformations. Rotation around C9=C10 double bond has the biggest barrier and around C13=C14 has the smallest barrier. The complete list of barriers of the three isomerizations in the ground states are illustrated in Table 3.

To investigate the external conditions (proper light frequency) needed to initiate the isomerization processes, we calculated the low-lying excited-states of all-*trans*-retinal and its isomers. The energies of the excited states as shown in Table 4 are found within time-dependent DFT (TDDFT) methods. For molecules with low-lying excited-states,

Table 4 Excitation energies of retinal isomers calculated with TDDFT method (eV) and OS represents oscillator strength

	9-*cis*-retinal		11-*cis*-retinal		13-*cis*-retinal		All-*trans*-retinal	
	E(eV)	OS	*E*(eV)	OS	*E*(eV)	OS	*E*(eV)	OS
S_1	3.0708	0.9110	3.0490	1.0821	3.0673	1.0771	3.0496	1.2225
S_2	3.0833	0.0006	3.0505	0	3.0991	0.0002	3.0646	0.0054
S_3	3.8894	0.5567	3.8612	0.5131	3.9118	0.5023	3.8893	0.5745

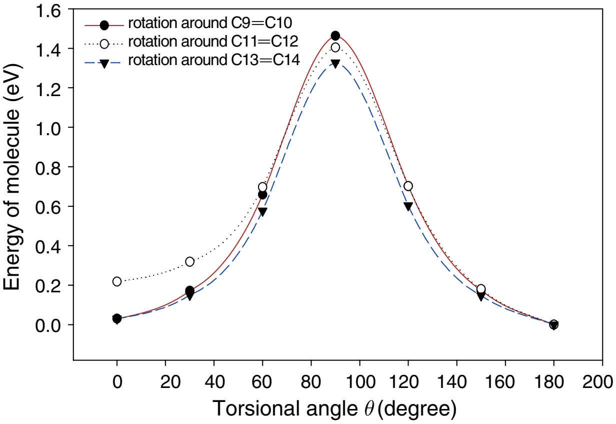


Figure 15 PECs of the ground state of all-*trans*-retinal rotating around C9=C10, C11=C12, and C13=C14 double bonds at DFT level

TDDFT makes a considerable improvement over HF-based methods (like CIS) [39]. The excitation energies are similar for the corresponding states of these four isomers and the lowest excitation energies 3.07 eV (ca. 403 nm) are in the visible region. The similar excitation energies show that our existing simulations are not sufficient to explain the isomerization mechanism around different double bonds. As mentioned before, in rhodopsin the isomerization from 11-*cis*-retinal to all-*trans*-retinal occurs but in bacteriorhodopsin the isomerization from all-*trans*-retinal to 13-*cis*-retinal takes place. So the proper environments, such as protein environment in this case, probably play the most important role in determining which double bond participates in the isomerization process.

3. Retinal nanostructure arrays

Monolayers of retinal immobilized on solid substrates are of considerable importance not only for applications as retinal-based sensors but also for investigations of the complex behavior of retinal molecules at interfaces. By modifying one end of retinal with a thiol linker, we could chemically graft the retinal to gold surfaces. Because two-dimensional gold nanostructure arrays can be produced on highly oriented pyrolytic graphite (HOPG) surfaces [40] and on semiconductor substrates [41], we can finally fabricate two-dimensional retinal nanostructure arrays. In addition, by controlling the two-dimensional spatial distribution of gold nanostructures including shapes and sizes, we can control retinal molecules quantitatively.

3.1. Two-dimensional nanopatterned retinal structure

Two-dimensional nanopatterned retinal structures can be created in the following three steps. First, connect retinal to cysteine (amino acid) which contains the thiol group by nucleophilic addition-elimination reaction between aldehyde (C=O) and primary amine ($-NH_2$) as described in most organic chemistry textbooks such as [42]. The reaction is shown in Figure 16 with the product *N*-retinylidene cysteine.

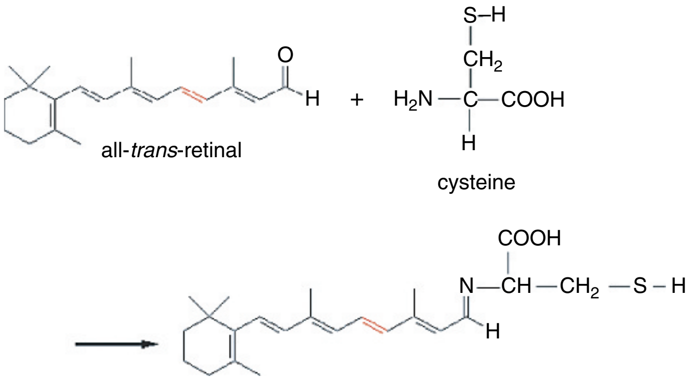


Figure 16 Nucleophilic addition-elimination reaction between all-*trans*-retinal and cysteine amino acid

Second, fabricate two-dimensional gold nanostructure arrays on pit-patterned HOPG surfaces [40] or semiconductor substrates by a dip-pen nanolithography (DPN)-based strategy [41]. A two-step experiment was performed in [40] to produce two-dimensional patterns of gold nanostructures on HOPG. First, nanometer-size etch pits were patterned as regular arrays on HOPG using highly focused 25 keV Ga^+ ion bombardment. The second step was to deposit gold onto the pit-patterned HOPG surfaces in vacuum. In [41], the investigators developed a method for fabricating arrays of Au nanostructures on a SiO_x/Si surface based on DPN and wet chemical etching.

Finally, due to the strength of the thiol groups at the end of cysteine amino acids, self-assembled monolayers of retinylidene cysteine molecules on gold nanopaticles can be formed. For example, in [43] they show how the etched Au nanopatterns can be used as templates to adsorb and assemble proteins from solution to form functionalized inorganic/biological nanostructures and in [44] self-assembled monolayers on Au(111) are formed by microcontact printing of dodecanethiol. The bonding of the sulfur head group to the gold substrate is in the form of a metal thiolate (RS^-) species [45]. This is an extremely strong surface bond and the resulting monolayers are quite stable. The formation of gold thiolate requires the loss of SH hydrogen but it has not been determined whether this proton is lost as H_2 or H_2O [46]. The presumed adsorption chemistry is shown in the equation of Figure 17 [47]. In this way, retinal derivative molecules are confined in a nano-size space and two-dimensional retinal nanostructure arrays are formed as shown in Figure 18.

3.2. Simulation and analysis of retinal derivatives

To approximately infer how the gold substrates will affect the properties of retinal molecules, simulations have been conducted to predict the properties of retinylidene

$$\mathrm{Au} + \mathrm{HS}\,(\mathrm{CH_2})_n\mathrm{X} \xrightarrow[\text{300 K, EtOH}]{-\mathrm{H_2}(+\mathrm{O_2}-\mathrm{H_2O}?)} \mathrm{Au{-}S^-}\,(\mathrm{CH_2})_n\mathrm{X}$$

Figure 17 Chemical reaction between molecules with thiol group and gold [47]

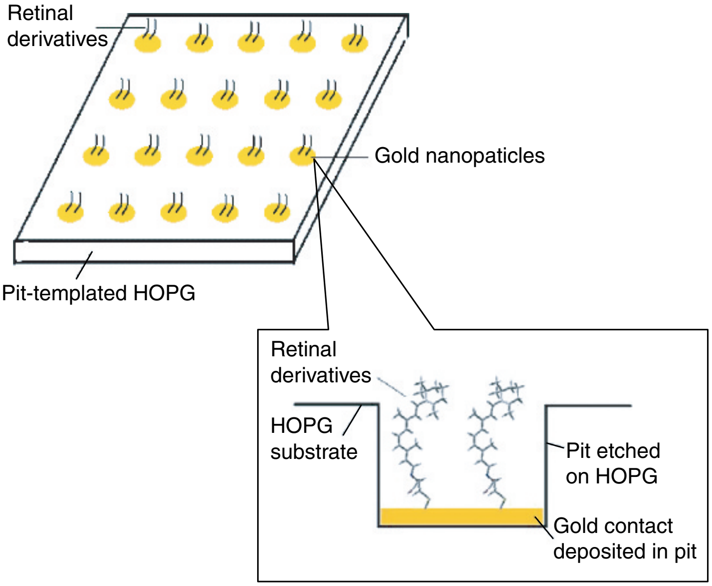


Figure 18 Two-dimensional nanopatterned *N*-retinylidene cysteine arrays

cysteine-gold molecules. While the long-range goal is to simulate retinylidene cysteine-gold monolayers on gold clusters, initial investigations have been focused on the case where one retinal isomeric molecule is connected to a gold atom via the link of a cysteine molecule (*N*-retinylidene cysteine-gold molecule) and one such *N*-all-*trans*-retinylidene cysteine-gold molecule is illustrated in Figure 19. Other retinylidene cysteine-gold isomers simulated in this chapter are constructed in the same manner.

The THz characteristics of various retinylidene cysteine-gold molecular conformations have been investigated at HF level using LANL2DZ basis set. The scaling factor 0.9 was applied to eliminate the systematic errors due to the method and basis set [48].

The very far infrared spectra ($<30\,\mathrm{cm}^{-1}$) of *N*-all-*trans*-retinylidene cysteine-gold and its isomers are shown in Figures 20–22. It is clear that there are many more vibrational modes in FIR spectra of retinal derivatives than those of isolated retinal isomers. The strongest absorption peaks of *N*-9-*cis*-retinylidene cysteine-gold, *N*-11-*cis*-retinylidene

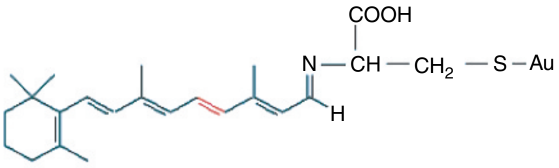


Figure 19 An all-*trans*-retinal molecule connected with a cysteine amino acid molecule and then a gold atom (*N*-all-*trans*-retinylidene cysteine-gold molecule)

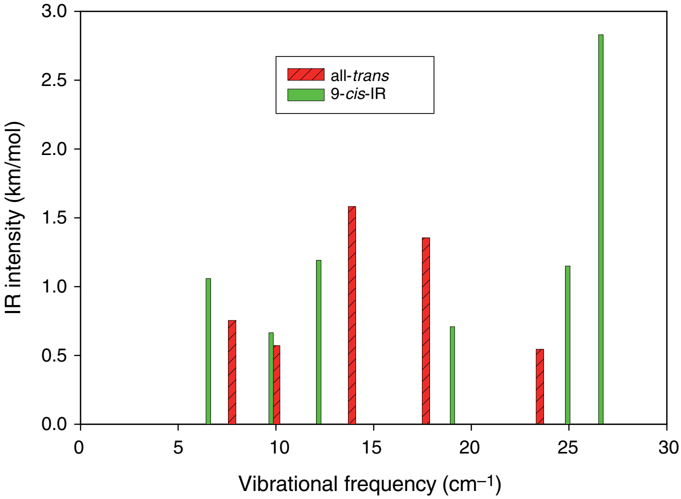


Figure 20 FIR spectra of *N*-9-*cis*-retinylidene cysteine-gold and *N*-all-*trans*-retinylidene cysteine-gold (<30 cm^{-1})

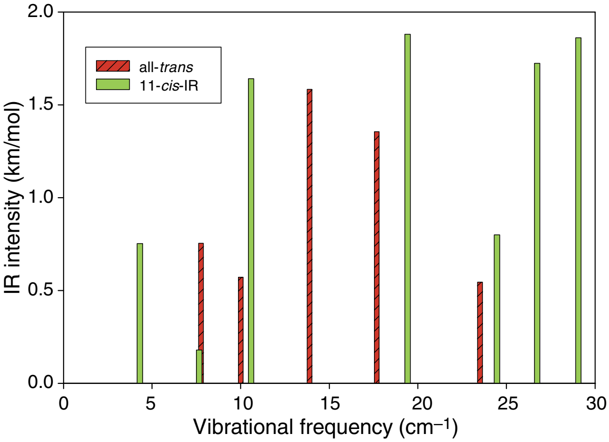


Figure 21 FIR spectra of *N*-11-*cis*-retinylidene cysteine-gold and *N*-all-*trans*-retinylidene cysteine-gold (<30 cm^{-1})

cysteine-gold and *N*-13-*cis*-retinylidene cysteine-gold molecule are quite distinguishable from that of *N*-all-*trans*-retinylidene cysteine-gold, which shows that the spectra fidelity of retinal isomers still exists in these retinal derivatives.

In addition, the PECs of the retinylidene cysteine-gold molecule were derived as were done earlier for the isolated retinal. Geometric optimizations and calculations of potential energies of molecules were performed at DFT using LANL2DZ basis set. B3PW91 is used by the DFT level, which uses the Becke-3 hybrid exchange functional and the Perdew–Wang 91 correlation functional. The PECs of rotation for *N*-all-*trans*-retinylidene cysteine-gold around C9=C10, C11=C12, and C13=C14 double bonds are

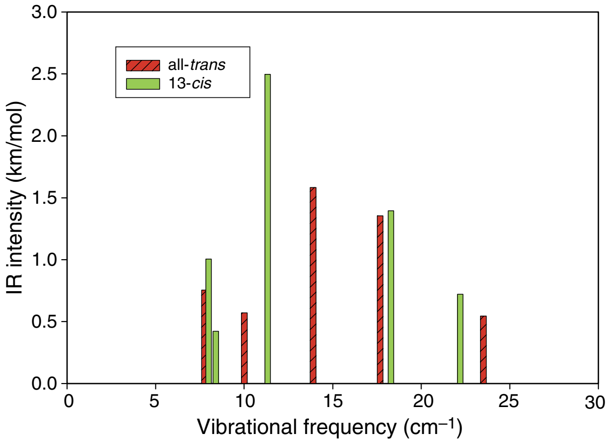


Figure 22 FIR spectra of *N*-13-*cis*-retinylidene cysteine-gold and *N*-all-*trans*-retinylidene cysteine-gold ($<30\,\mathrm{cm}^{-1}$)

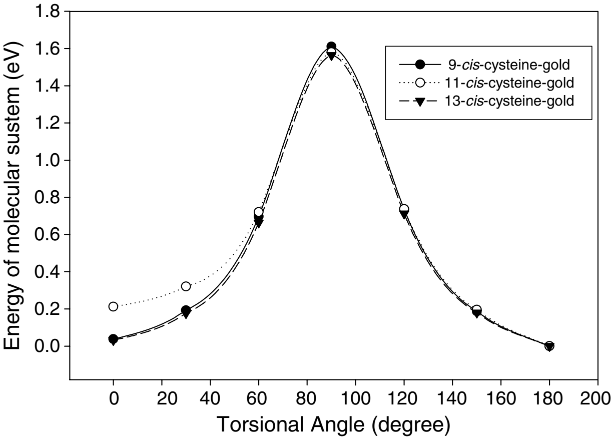


Figure 23 PECs of *N*-all-*trans*-retinylidene cysteine-gold molecule rotated around C9=C10, C11=C12, and C13=C14 double bonds

shown in Figure 23. Besides the spectra fidelity, other observations from the simulations of isolated retinal isomers can also be found from the simulations of retinal derivatives isomers. For example, significant barriers are observed in PECs that are even larger than those of isolated retinal. And another one is that the barrier for rotation around the C9=C10 double bond of retinylidene cysteine-gold is still the largest and around C13=C14 is still the smallest. The complete list of excitation energies of both isolated retinal and retinylidene cysteine-gold is shown in Table 5. Excited states were calculated with the time-dependent DFT (TDDFT) method. The excitation energies of the retinal derivatives are smaller than those of corresponding isolated retinal isomers but still relatively close to each other as isolated retinal isomers. According to the discussion

Table 5 Excitation energies calculated with TDDFT method (eV)

Isomers	S_1	S_2	S_3
All-trans-retinal (Isolated)	3.0496	3.0646	3.8893
N-all-trans-retinylidene cysteine-gold	1.6606	1.8640	2.7852
9-cis-retinal (Isolated)	3.0708	3.0833	3.8894
N-9-cis-retinylidene cysteine-gold	1.6900	1.8780	2.7732
11-cis-retinal (Isolated)	3.049	3.0505	3.0991
N-11-cis-retinylidene cysteine-gold	1.6654	1.8745	2.7825
13-cis-retinal (Isolated)	3.0673	3.0991	3.9118
N-13-cis-retinylidene cysteine-gold	1.6594	1.8656	2.7506

above, it is reasonable to conclude that many properties of isolated retinal isomers are preserved in retinal derivatives.

4. Directions for future work and conclusions

This chapter has presented a new bio-molecular electronic architecture that offers an enhanced capability for sensing THz-frequency bio-signatures. This chapter also presents modeling results that illustrate prototypical examples for the type of functional bio-molecules needed for implementing this bio-molecular system. In the future, this work will be used to guide new experimental investigations that focus on the measurement of very far infrared spectra of retinal isomers and retinal isomer derivatives. The studies presented in this chapter analyzed compound molecules of the type needed for realizing functional bio-molecular devices and derived their THz spectra and photo-induced methods for controlling the THz-frequency characteristics through changes to energy state and conformation. One possible fabrication approach for the construction of retinal-based two-dimensional nanostructures was also presented. Future theoretical work will consider such functional bio-molecules that are bonded to traditional nanoscale semiconductor systems. The ultimate goal of this research is to define functioning bio-molecular devices (e.g., light guiding structures) that can be used to achieve bio-signature sensing and data processing in ultra-small integrated platforms.

References

[1] D. L. Woolard, Y. Luo, B. L. Gelmont, T. Globus, and J. O. Jensen, "Bio-Molecular inspired electronic architectures for enhanced sensing of THz-frequency bio-signatures", *Int. J. High Speed Electron*, **16**, no. 2, 609–637 (2006).

[2] D. L. Woolard, Y. Luo, B. L. Gelmont, T. Globus, and J. O. Jensen, "A bio-molecular inspired electronic architecture: Bio-based device concepts for enhanced sensing", invited paper in *Proc. SPIE Secur. Defense Symp.*, **5790**, 180 (2005).

[3] D. L. Woolard, Y. Luo, B. L. Gelmont, T. Globus, J. O. Jensen, H-L Cui, and N. C. Seeman, "Patent title: Bio-molecular inspired electronic architectures for enhanced sensing of THz-frequency bio-signatures", Patent submission to the legal office of U.S.Army Reserch Office, March 2005.

[4] D. Woolard, T. Globus, E. Brown, L. Werbos, B. Gelmont, and A. Samuels, "Sensitivity Limits & Discrimination Capability of THz Transmission Spectroscopy as a Technique for Biological Agent Detection," in the proceedings to the 5th Joint Conference on Standoff Detection for Chemical and Biological Defense, Williamsburg, VA, 24–28 September (2001).

[5] D. L. Woolard, T. Koscica, D. L. Rhodes, H. L. Cui, R. A. Pastore, J. O. Jensen, J. L. Jensen, W. R. Loerop, R. H. Jacobsen, D. Mittleman, and M. C. Nuss, "Millimeter wave induced vibrational modes in DNA as a possible alternative to animal tests to probe for carcinogenic mutations," *J. Appl. Toxicol.*, **17**, 243–246, May (1997).

[6] Dwight L. Woolard, William R. Loerop and Michael S. Shur (eds), *Terahertz Sensing Technology, Volume 1: Electronic Devices and Advanced Systems Technology* (World Scientific, Singapore, 2003).

[7] Dwight L. Woolard, William R. Loerop and Michael S. Shur (eds), *Terahertz Sensing Technology, Volume 2: Emerging Scientific Applications & Novel Device Concepts* (World Scientific, Singapore, 2003).

[8] T. Globus, D. Woolard, M. Bykhovskaia, B. Gelmont, L. Werbos, and A. Samuels, "Millimeter and submillimeter wave spectroscopy of DNA and related materials (theory and experiment)," *Int. J. High Speed Electron. Syst. (IJHSES)*, **13**, no. 4, 903–936, December (2003).

[9] T. Globus, D. Theodorescu, H. Frirson, T. Kchromova, and D. Woolard, "Terahertz Spectroscopic Characterization of Cancer Cells," accepted SPIE, *Photonics West*, San Jose, California, January (2005).

[10] T. Globus, L. Dolmatova-Werbos, D. Woolard, A. Samuels, B. Gelmont, and M. Bykhovskaia, "Application of Neural Network Analysis to Submillimeter-wave Vibrational Spectroscopy of DNA Macromolecules," in the proceedings to the 2001 ISSSR, June 12–15, Quebec City, Canada (2001).

[11] M. Bykhovskaia, B. Gelmont, T. Globus, D. L. Woolard, A. C. Samuels, T. Ha-Duong, and K. Zakrzewska, "Prediction of DNA far IR absorption spectra basing on normal mode analysis," *Theor. Chem. Acc.*, **106**, 22–27 (2001).

[12] T. Globus, M. Bykhovskaia, D. Woolard, and B. Gelmont, "Sub-Millimeter Wave Absorption Spectra of Artificial RNA Molecules," published in the proceedings to the Advanced Workshop on Frontiers in Electronics (WOFE), St. Croix, Virgin Islands, 6–11 January (2002).

[13] T. Globus, D. L. Woolard, A. C. Samuels, T. Khromova, and J. O. Jensen, "Submillimeter Wave Fourier Transform Characterisation of Bacterial Spores, " in the Proceedings of the *Int. Symp. Spectral Sensing Res.*, Santa Barbara, California, 2–6 June 2003, file://F:\isssr2003\index.htm (2004).

[14] T. Globus, T. Khromova, D. Woolard, and B. Gelmont, "Terahertz Fourier transform characterization of biological materials in solid and liquid phases," presented at SPIE Photonic East, Program on Chemical and Biological Sensors, Providence, RI, October 2003, in *Proc. SPIE*, **5268–2**, 10–18 (2004).

[15] T. Globus, M. Bykhovskaia, D. Woolard, and Boris Gelmont, "Submillimeter-wave absorption spectra of artificial RNA molecules," invited to *J. Phys. D*, **36**, 1314–1322 (2003).

[16] T. R. Globus, D. L. Woolard, T. Khromova, T. W. Crowe, M. Bykhovskaia, B. L. Gelmont, J. Hesler and A. C. Samuels, "THz-Spectroscopy of biological molecules," *J. Biol. Phys.*, **29**, 89–100 (2003).

[17] Thomas W. Crowe, Tatiana Globus, Dwight L. Woolard, and Jeffery Hesler, "Terahertz sources and detectors and their application to biological sensing," invited to *Phil. Trans. R. Soc. Lond. A* **362**, 365–377 (2004).

[18] T. Globus, D. Woolard, T. Khromova, R. Partasarathy, A. Majewski, R. Abreu, J. Hesler, S.-K. Pan, and G. Ediss, "Terahertz signatures of biological-warfare-agent simulants," *Proc. SPIE Defense Secur. Symp.*, **5411–5**, Orlando, FL, April (2004).

[19] D. Woolard, T. R. Globus, B. L. Gelmont, M. Bykhovskaia, A. C. Samuels, D. Cookmeyer, J. L. Hesler, T. W. Crowe, J. O. Jensen, J. L. Jensen, and W. R. Loerop, "Submillimeter-Wave Phonon Modes in DNA Macromolecules," *Phys. Rev. E*, **65**, 1903–1914, May (2002).

[20] V. K. Saxena, B. H. Dorfman, and L. L. Van Zandt, *Phys. Rev.*, A **43**, 4510–4516 (1991).

[21] L. L. Van Zandt and V. K. Saxena, in *Structure & Functions, Volume 1: Nucleic Acids*, Eds R. H. Sarma and M. H. Sarma (Adenine Press, 1992) and references therein.

[22] T. Globus, D. L. Woolard, A. C. Samuels, B. L. Gelmont J. L. Hesler, T. W. Crowe, and M. Bykhovskaia, "Submillimeter-wave FTIR spectroscopy of DNA macromolecules and related materials," *J. Appl. Phys.*, **91**, 6106–6113 (2002).

[23] E. R. Brown, D. L. Woolard, A. C. Samuels, T. Globus, and B. Gelmont, "Remote detection of bioparticles in the THz region," *Proc. Int. Microw. Symp.*, Seattle, WA, June 2–6 (2002).

[24] D. L. Woolard, E. R. Brown, A. C. Samuels, T. Globus, B. Gelmont, and M. Wolski, "Terahertz-frequency spectroscopy as a technique for the remote detection of biological warfare agents," presented to the special session on terahertz electronics, *Proc. IEEE Int. Microw. Symp.*, 8–13 June, Philadelphia, PA (2003).

[25] M. K. Choi, K. Taylor, A. Bettermann, and D. W. van der Weide, "Spectroscopy with electronic terahertz techniques for chemical and biological sensing," pp. 35–48 in reference [7].

[26] D. W. van der Weide, M. K. Choi, K. Taylor, and A. Bettermann, "Sensing biomolecules with microwave and terahertz frequencies," invited to 16th International Zurich Symposium on Electromagnetic Compatibility, Zurich, 13–18 February (2005).

[27] G. Markelz and S. E. Whitmire, "Terahertz applications to biomolecular sensing," pp. 49–66 of reference [7].

[28] T. Globus, D. Woolard, M. Bykhovskaia, B. L. Gelmont, L. Werbos, and A. C. Samuels, "THz-frequency spectroscopic sensing of DNA and related materials, p. 1–34 of reference [7].

[29] E. R. Brown, J. E. Bjarnason, T. L. J. Chan, A. W. M Lee, and M. A. Celis, "Optical attenuation signatures of *Bacillus subtillis* in the THz region," *Appl. Phys. Lett.*, **84**, 3438–3440 (2004).

[30] B. L. Gelmont, "Computational modeling of light induced transformations in organic molecules," U.S. ARO Interim (10/1/02–12/31/02) Progress Report on Research Grant DAAD19-02-1-0439 (2003).

[31] M. J. Frisch, *et al.*, *Gaussian*, Inc., Pittsburgh, PA (2002).

[32] J. B. Foresman and Æleen Frisch, *Exploring Chemistry with Electronic Structure Methods*, 2nd edn Gaussian, Inc., Pittsburgh, PA (1996).

[33] D. C. McKean, M. W. Mackenzie, A. R. Morrisson, J. C. Lavalley, A Janin, V. Fawcett, and H. G. M. Edwards, *Spectrochimica Acta*, **41A**, 3, 435–450 (1985).

[34] S. Yamamoto, H. Wasada and T. Kakitani, and T. Yamato, *J. Molecul. Struct.*, **461–462**, 463–471 (1999).

[35] S. Yamamoto, H. Wasada, and T. Kakitani, *J. Molecul. Struct.*, **451**, 151–162 (1998).

[36] F. L. Gervasio, G. Cardini, P. R. Salvi, and V. Schettino, *Journal of Physical Chemistry A*, **102**, 2131–2136 (1998).

[37] M. Walther, B. Frischer, M. Schall, H. Helm, and P. Uhd Jepsen, *Chemi. Phys. Lett.*, **332**, 389–395 (2000).

[38] Æleen Frisch and M. J. Frisch, *Gaussian 98 User's Reference*, 2nd edn Gaussian, Inc., Pittsburgh, PA (1999).

[39] R. E. Stratmann, G. E. Scuseria, and M. J. Frisch, *J. Chem. Phys.*, **109**, 8218–8224 (1998).

[40] Y Zhu, A. Schnieders, J. D. Alexander, and T. P. Beebe, Jr., *Langmuir*, **18**, 5728–5733 (2002).

[41] H. Zhang, Z. Li, and C. A. Mirkin, *Adv. Mater.*, **14**, 1472 (2002); H. Zhang, S. W. Chung, and C.A. Mirkin, *Nano Lett.*, **3**, 43 (2003).

[42] F. A. Carey, *Organic Chemistry*, p. 664. (1987).

[43] H. Zhang, K. Lee, Z. Li, and C. A. Mirkin, "Biofunctionalized nanoarrays of inorganic structures prepared by dip-pen nanolithography, " *Nanotechnol.*, **14**, 10, 1113–1117 (2003).

[44] N. B. Larsen, H. Biebuyc, E. Delamarche, and B. Michel, "Order in microcontact printed self-assembled monolayers", *J. Am. Chem. Soc.*, **119**, 3017–3026 (1997).

[45] R. G. Nuzzo, B. R. Zegarski, and L. H. Dubois, *J. Am. Chem. Soc.*, **109**, 2358 (1987).

[46] Ulman, *An introduction to Ultrathin Organic Films: From Langmuir-Blodgett to Self-Assembly*, Academic Press, Inc. (1991).

[47] L H. Dubois and R. G. Nuzzo, "Synthesis, structure, and properties of model organic surfaces", *Annu. Rev. Phys. Chem.*, **43**, 437–463 (1992).

[48] http://srdata.nist.gov/cccbdb/vsf.asp.

Molecular and Nano Electronics: Analysis, Design and Simulation
J. M. Seminario (Editor)

Chapter 3

Charge delocalization in (*n*, 0) model carbon nanotubes

Peter Politzer, Jane S. Murray and Monica C. Concha

Department of Chemistry, University of New Orleans, New Orleans,
LA 70148, USA. ppolitze@uno.edu

1. Introduction

Carbon nanotubes are long cylindrical structures composed of hexagonal rings of carbon atoms. These rings can have various orientations with respect to the tube axis, as can be seen in the examples in Figure 1. The two extremes are when each ring has two sides parallel to the axis, Figure 1(a), and perpendicular to it, Figure 1(b). An intermediate case (one of many possibilities) is illustrated in Figure 1(c).

A system has been devised for labeling the different types of nanotubes in terms of two positive integers (n, m); it is explained in detail elsewhere [1–4]. Our present interest is primarily in tubes having two sides of the carbon rings parallel to the axis; these are in the category $(n, 0)$, with n being the number of rings around the circumference of the tube. Thus Figure 1(a) shows an (8, 0). Tubes having two ring sides perpendicular to the axis are labeled (n, n); all others are (n, m) with $n \neq m$.

The structures in Figure 1 are portions of "single-walled" nanotubes, meaning that each is a single cylinder. It is also possible to have several coaxial tubes; these are designated as "multi-walled." We shall focus solely upon the former.

Another feature of nanotubes, in practice, is that they are normally closed at both ends (capped) when synthesized. The caps can have various structures and shapes [2]. However, whereas the lateral sides of carbon nanotubes are composed entirely of six-membered rings (except for possible defects), each cap must contain six five-membered rings. This follows from Euler's theorem: Any hexagonal framework can achieve complete closure only through the introduction of exactly twelve pentagons [1, 2].

The unusual structural, electronic and mechanical properties of carbon nanotubes have aroused enormous interest and stimulated a great deal of work, as well as proposed applications in many different areas. For recent overviews, see White and Mintmire [3] and Politzer *et al.* [4].

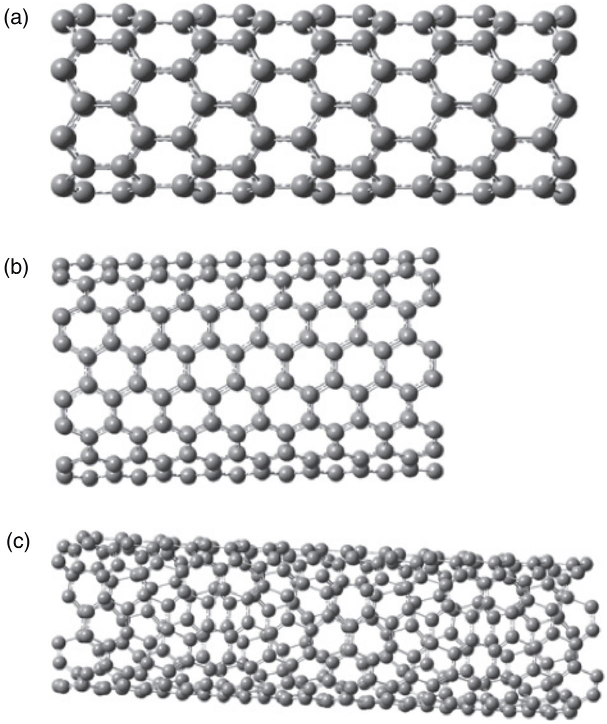


Figure 1 Portions of lateral surfaces of carbon nanotubes: (a) (8, 0); (b) (8, 8); (c) (8, 5)

A natural extension of all this activity has been to investigate the possibility of nanotubes with other elemental compositions. The boron/nitrogen combination seemed particularly promising because of B_xN_x being isoelectronic with C_{2x}, as well as other analogies between B/N and carbon compounds, such as the fact that solid boron nitride exists in both graphite-like and diamond-like forms [5, 6]. B_xN_x [7, 8] and $C_xB_yN_z$ [9, 10] nanotubes have indeed been prepared.

In this chapter, our emphasis will be upon the remarkable capacities of some (*n*, 0) model nanotubes for charge delocalization and transmission of electronic effects. We will limit the discussion to all-carbon systems, although some $C_xB_yN_z$ are also of interest in this respect. We became aware of these unique features in the course of computational analyses of the electrostatic potentials and local ionization energies on model nanotube surfaces [4, 11, 12]. We will therefore begin by providing some background relating to these two properties.

2. Electrostatic potential

The nuclei and electrons of any system produce an electrostatic potential $V(\mathbf{r})$ in the surrounding space; its value at any point $\mathbf{r}$ is given rigorously by Eq. (1):

$$\mathrm{V}(\mathbf{r}) = \sum_{\mathrm{A}} \frac{Z_{\mathrm{A}}}{|\mathbf{R}_{\mathrm{A}} - \mathbf{r}|} - \int \frac{\rho(\mathbf{r}')\mathrm{d}\mathbf{r}'}{|\mathbf{r}' - \mathbf{r}|} \tag{1}$$

Z_A is the charge on nucleus A, located at $\mathbf{R}_A$, and $\rho(\mathbf{r})$ is the electronic density. The sign of $V(\mathbf{r})$ in any region depends upon whether the positive contribution of the nuclei or the negative one of the electrons is dominant there.

The potential $V(\mathbf{r})$ is physically observable. It has been determined experimentally, by diffraction techniques [13–15], as well as computationally. In our studies of model nanotubes, we compute $V(\mathbf{r})$ on both their inner and outer surfaces; we take these to be defined by the 0.001 electrons/bohr^3 contours of their electronic densities, as suggested by Bader *et al.* for molecules [16].

The electrostatic potential on the surface of a molecule or other system, which we label $V_S(\mathbf{r})$, can be characterized in terms of several statistical quantities. Two obvious ones are the most positive and the most negative values of $V_S(\mathbf{r})$: $V_{S,\max}$ and $V_{S,\min}$. Others that will be relevant with respect to model nanotubes are the total variance, σ^2_{tot}, and its positive and negative components, σ^2_+ and σ^2_-:

$$\sigma^2_{\text{tot}} = \sigma^2_+ + \sigma^2_- = \frac{1}{m}\sum_{i=1}^{m}\left[V_S^+(\mathbf{r}_i) - \overline{V}_S^+\right]^2 + \frac{1}{n}\sum_{j=1}^{n}\left[V_S^-(\mathbf{r}_j) - \overline{V}_S^-\right]^2 \tag{2}$$

In Eq. (2), $\overline{V}_S^+$ and $\overline{V}_S^-$ are the average positive and negative values of $V_S(\mathbf{r})$:

$$\overline{V}_S^+ = \frac{1}{m}\sum_{i=1}^{m} V_S^+(\mathbf{r}_i) \tag{3}$$

$$\overline{V}_S^- = \frac{1}{n}\sum_{j=1}^{n} V_S^-(\mathbf{r}_j) \tag{4}$$

The first summation in Eq. (2) and that in Eq. (3) are over the points where $V_S(\mathbf{r})$ is positive, the others are over those where it is negative, on grids covering the inner and outer nanotube surfaces.

The quantities σ^2_{tot}, σ^2_+ and σ^2_- indicate how variable the total, positive and negative surface potentials are, i.e. how extensive their ranges are. Since the terms in Eq. (2) are squared, the extrema of $V_S(\mathbf{r})$ – the $V_{S,\max}$ and $V_{S,\min}$ – are particularly influential.

We have shown earlier that the quantities that we use to characterize $V_S(\mathbf{r})$ (which can include also its average deviation and an electrostatic balance parameter) provide an effective basis for correlating and predicting a variety of properties that depend upon noncovalent interactions: boiling points and critical constants, heats of phase transitions, solubilities and solvation energies, partition coefficients, diffusion constants, viscosities, surface tensions, etc. There have been several reviews of this work [17–20].

3. Average local ionization energy

The second property in terms of which we have analyzed model nanotube surfaces is the "average local ionization energy," $\bar{I}(\mathbf{r})$. We introduced this originally within the framework of Hartree–Fock (HF) theory [21], as:

$$\bar{I}(\mathbf{r}) = \sum_i \frac{\rho_i(\mathbf{r})\,|\varepsilon_i|}{\rho(\mathbf{r})} \tag{5}$$

in which $\rho_i(\mathbf{r})$ is the electronic density of the i^{th} occupied orbital of the system, having energy $|\varepsilon_i|$. Within Hartree–Fock formalism, the $|\varepsilon_i|$ are equal to the electrons' ionization energies, provided that their loss does not affect the remaining orbitals, an assumption for which Koopmans' theorem offers some support [22, 23]. Accordingly we interpret $\bar{\text{I}}(\mathbf{r})$ as the "average local ionization energy" that is required to remove an electron from the point $\mathbf{r}$ in the space of the system. The focus is upon the point in space, not a particular orbital. When we compute $\bar{\text{I}}(\mathbf{r})$ on $\rho(\mathbf{r}) = 0.001$ electrons/bohr^3 surfaces, as we normally do, then it is denoted by $\bar{\text{I}}_\text{S}(\mathbf{r})$.

Both $\bar{\text{I}}_\text{S}(\mathbf{r})$ and $\bar{\text{I}}(\mathbf{r})$ have been shown to be significant in a variety of areas, both fundamental and applied. $\bar{\text{I}}_\text{S}(\mathbf{r})$ correlates with atomic electronegativities [24] and shell structure [25], the latter reflecting its relationship to local temperature [26, 27]. $\bar{\text{I}}_\text{S}(\mathbf{r})$ is indicative (inversely) of local polarizability [28, 29], as well as C–C bond strain [30] and radical sites in polycyclic aromatic hydrocarbons [31].

With respect to chemical reactivity, it is the lowest values of $\bar{\text{I}}_\text{S}(\mathbf{r})$, the $\bar{\text{I}}_\text{S,min}$, that are of particular interest. These reveal the locations of the least tightly held electrons, those most reactive toward electrophiles. Thus, the $\bar{\text{I}}_\text{S,min}$ correctly predict the activating/deactivating and directing effects of benzene substituents [21, 32], and the sites most vulnerable to reaction with electrophiles (and radicals [31, 33]) in other organic molecules [33, 34]. As might be anticipated, the magnitudes of $\bar{\text{I}}_\text{S,min}$ correlate with proton affinities and pK$_\text{a}$ [35].

While the justification for our interpretation of $\bar{\text{I}}(\mathbf{r})$ comes from Hartree–Fock theory, $\bar{\text{I}}_\text{S,min}$ obtained by Kohn–Sham density functional methods have proven to be equally effective in predicting trends [32]. For a recent review and discussion of $\bar{\text{I}}(\mathbf{r})$ and $\bar{\text{I}}_\text{S}(\mathbf{r})$, see Politzer and Murray [27].

4. Computational procedure

The results that will be discussed were obtained by calculations at the Hartree–Fock (HF) STO-5G//STO-3G level. The use of minimum basis sets was dictated by the sizes of the systems; most of the model nanotubes that we have investigated in our earlier work had between 80 and 120 atoms, not including hydrogens. We believe, however, that the computational procedure was fully adequate for our purposes. The bond lengths produced by the HF/STO-3G optimizations are in good agreement with available experimental data [4, 36–38]. It has also been demonstrated that HF/STO-5G $V(\mathbf{r})$ [39, 40] and $\bar{\text{I}}_\text{S}(\mathbf{r})$ [40] are quite satisfactory on a relative basis, for showing trends, which has been our objective in these studies.

It was already mentioned that when nanotubes are prepared, they are generally capped at both ends. To facilitate various applications, these caps are often removed, by one of several methods [41–43]. Computationally, we have treated both capped and uncapped (open) model nanotubes. For the latter, in order to avoid having unfulfilled valencies at the ends, we follow the common practice of terminating the tubes with hydrogens [44–49]. In doing so, we must keep in mind that their presence may have some perturbing effect, hopefully small, upon the results.

5. Electrostatic potentials of model carbon nanotubes

We have computed $V_S(\mathbf{r})$ on the inner and outer surfaces of more than 30 model nanotubes of various compositions–carbon, B_xN_x and $C_xB_yN_z$ – and structural types [4, 11, 12, 36–38]; some of the carbon systems bore substituent groups. The number of non-hydrogen atoms was generally between 80 and 120. Before discussing these results, we would like to present some data that will help to put them into proper perspective.

In Table 1 are given, for a representative group of molecules, some of the computed quantities in terms of which we characterize molecular surface electrostatic potentials: $V_{S,\max}$, $V_{S,\min}$, $\overline{V}_S^+$, $\overline{V}_S^-$, σ_+^2, σ_-^2 and σ_{tot}^2. Our objective is simply to point out some general features:

1. Negative regions tend to be stronger than positive ones. Thus, usually $|V_{S,\min}| > V_{S,\max}$, $|\overline{V}_S^-| > \overline{V}_S^+$ and $\sigma_-^2 > \sigma_+^2$. The most positive $V_{S,\max}$ are in the neighborhood of 30 kcal/mole, and are associated with hydroxyl or sometimes amino hydrogens; in contrast, the $V_{S,\min}$ due to oxygen and nitrogen lone pairs are often -40 to -50 kcal/mole, or even more negative.
2. Exceptions to the preceding generalizations can arise when the molecule has several strongly-electron-withdrawing groups [51]. In Table 1, for example, 2,4,6-trinitrotoluene has $V_{S,\max} > |V_{S,\min}|$, $\overline{V}_S^+ > |\overline{V}_S^-|$ and $\sigma_+^2 > \sigma_-^2$. $V_{S,\max}$ is above the ring, which has lost much of its π electronic charge to the NO_2 groups.
3. Hydrocarbons usually have weak $V_S(\mathbf{r})$, with both $\overline{V}_S^+$ and $|\overline{V}_S^-|$ being less than 5 kcal/mole, and they show little variability, $\sigma_{\text{tot}}^2 < 20\,(\text{kcal/mole})^2$.

Table 1 Computed surface quantities (HF/STO-5G*//STO-3G*) for some representative molecules[a,b]

Molecule	$V_{S,\max}$	$V_{S,\min}$	$\overline{V}_S^+$	$\overline{V}_S^-$	σ_+^2	σ_-^2	σ_{tot}^2
benzene	9.5	−10.2	4.63	−4.87	7.3	8.5	15.8
butylbenzene	8.5	−12.4	3.74	−4.51	4.9	13.0	18.0
dimethylsulfide	11.2	−17.0	4.07	−5.69	7.2	25.7	33.0
1-hexanol	12.6	−36.7	5.00	−9.17	9.9	132.5	142.4
cyclohexanol	26.7	−43.5	5.61	−15.94	18.6	207.5	226.1
methanol	29.9	−41.6	10.31	−18.44	48.9	181.9	230.8
ammonia	18.1	−29.0	9.16	−12.78	27.6	73.7	101.3
methylamine	21.4	−53.2	9.32	−20.20	34.6	263.0	297.6
piperazine	22.9	−47.8	9.32	−17.40	26.3	207.2	233.5
phenol	34.9	−36.6	8.62	−8.54	64.0	73.1	137.1
acetamide	32.0	−40.6	12.73	−20.61	68.0	150.8	218.9
para-cresol	34.0	−35.9	7.32	−8.11	53.6	68.9	122.5
para-dichlorobenzene	15.2	−11.1	6.17	−7.01	18.0	10.2	28.1
2,4,6-trinitrotoluene	37.5	−28.8	20.63	−14.69	104.3	53.2	157.5

[a] Units: $V_{S,\max}$, $V_{S,\min}$, $\overline{V}_S^+$ and $\overline{V}_S^-$ are in kcal/mole; σ_+^2, σ_-^2 and σ_{tot}^2 are in $(\text{kcal/mole})^2$.
[b] Data taken from [50].

Another point of interest, although not evident in Table 1, is that the positive regions on molecular surfaces, while weaker than the negative ones, are typically larger in area, often considerably so [50]. This originates in the fact that $V(\mathbf{r})$ is positive everywhere for ground-state, spherically averaged atoms [52, 53], the contribution of the nucleus dominating those of the dispersed electrons. When atoms combine to form a molecule, some negative regions normally develop, e.g., near localized lone pairs or the π electrons of unsaturated hydrocarbons, but most of the surface remains positive.

In proceeding now to model nanotubes, we will focus first upon the important role that their curvatures play in determining their electrostatic potentials. A curvature has two general consequences, which are common to all of the systems that we have studied, regardless of composition or structure (n, m):

1. The inner surfaces are always somewhat more positive (or less negative) than the outer. This is because each inner point is in closer proximity to more nuclei than is the corresponding point on the outer surface. This becomes more significant as the diameter of the tube decreases, and it is particularly important inside the caps of closed tubes, since these tend to have high levels of curvature.
2. A second factor, which makes the outer surfaces more negative (or less positive), is believed to arise from the effects of curvature upon the covalent bonds between the atoms [54, 55]. It is argued that curvature forces the atoms to deviate from the sp^2 configurations of a planar hexagonal framework and to acquire some sp^3 character, thereby introducing electronic localization in the fourth unfulfilled valency, as well as some degree of strain. An alternate explanation focuses upon curvature diminishing $2p\pi$–$2p\pi$ overlaps, with similar results.

What we have observed, therefore, is that $V_S(\mathbf{r})$ is normally more positive on the inner surface of a tube, especially inside the caps, and more negative on the outside, again more so on the caps. The differences become less significant as curvature decreases; thus a planar sheet of carbon hexagons, which is called graphene, has the same $V_S(\mathbf{r})$ on both sides [56].

Table 2 lists a selection of the model carbon nanotubes that we have studied. For each one are presented the same computed surface quantities, all related to $V_S(\mathbf{r})$, as are in Table 1. For most of the systems in Table 2, color plots of their surface electrostatic potentials can be found in the respective references.

The $V_S(\mathbf{r})$ of the two closed tubes, (6, 0) C_{96} and (5, 5) C_{120}, are perhaps the most realistic, since any perturbation by the terminal hydrogens is avoided. The surfaces of these two tubes are primarily weakly positive, with the $V_{S,\max}$ inside the caps. (The relatively high $V_{S,\max}$ of the (6, 0) C_{96} is due to the greater curvature of its caps [38].) Most of the outer surfaces have $V_S(\mathbf{r}) < 10\,\text{kcal/mole}$; the inner are somewhat more positive, but overall $\overline{V}_S^+ < 4\,\text{kcal/mole}$. The only negative regions are on the outsides of the caps, but the $V_{S,\min}$ are only -5.9 and $-1.8\,\text{kcal/mole}$.

The terminal hydrogens on the open tubes do affect $V_S(\mathbf{r})$ to some extent, as anticipated. They donate some charge to the carbons, so that the lateral outer surfaces become very slightly negative, mainly between 0 and $-5\,\text{kcal/mole}$. The $V_{S,\max}$ are now associated with the hydrogens, which also cause the $\overline{V}_S^+$ to be somewhat higher.

Overall, the surface electrostatic potentials of these model carbon nanotubes, whether closed or open, are very weak and rather bland, with little variability; note

Table 2 Computed surface quantities (HF/STO-5G//STO-3G) for some model carbon nanotubes[a]

Composition	(n, m)	$V_{S,max}$	$V_{S,min}$	$\overline{V}_S^+$	$\overline{V}_S^-$	σ_+^2	σ_-^2	σ_{tot}^2	Ref.
Closed (capped)									
C_{96}	(6, 0)	24.5	−5.9	3.1	−1.8	13.0	3.2	16.2	b
C_{120}	(5, 5)	10.3	−1.8	2.6	−0.5	4.9	0.2	5.0	c
Open (uncapped)									
$C_{62}H_{16}$	(7, 1)	13.7	−7.3	7.0	−3.0	17.7	1.7	19.4	c
$C_{68}H_{14}$	(6, 1)	14.8	−7.4	6.3	−2.7	19.9	1.2	21.1	c
$C_{68}H_{18}$	(8, 1)	13.6	−7.3	7.2	−3.1	16.5	1.6	18.1	c
$C_{72}H_{12}$	(6, 0)	15.2	−4.4	7.3	−2.0	21.4	0.8	22.2	d
$C_{80}H_{20}$	(5, 5)	14.7	−8.0	7.0	−3.9	17.2	2.1	19.3	c
Open, substituted									
$C_{68}H_{13}OH$	(6, 1)	35.8	−34.0	7.0	−3.2	40.1	15.6	55.6	e
$C_{68}H_{13}NH_2$	(6, 1)	29.1	−35.9	7.5	−3.8	43.9	10.4	54.3	d
$C_{80}H_{19}OH$	(5, 5)	37.2	−31.5	8.1	−4.3	38.9	7.1	46.0	e
$C_{72}H_{11}OH$	(6, 0)	74.7	−41.8	22.2	−18.1	297.9	121.2	419.1	e
$C_{72}H_{11}NH_2$	(6, 0)	73.8	−45.5	24.7	−19.1	409.3	139.4	548.7	e
$C_{72}H_{11}NO_2$	(6, 0)	58.5	−85.1	19.2	−24.0	264.3	669.0	933.3	e
$H_2NC_{72}H_{10}NO_2$	(6, 0)	75.8	−87.2	21.8	−21.6	410.4	625.7	1036.1	e

[a] Units: $V_{S,max}$, $V_{S,min}$, $\overline{V}_S^+$ and $\overline{V}_S^-$ are in kcal/mole; σ_+^2, σ_-^2 and σ_{tot}^2 are in (kcal/mole)^2.
[b] Reference [37]. [c] Reference [36]. [d] Reference [4]. [e] Reference [11].

that $\sigma_{tot}^2 < 25\,(\text{kcal/mole})^2$. In these respects, they are similar to the hydrocarbons in Table 1, but contrast markedly with most of the other molecules in that table.

What happens if one of the hydrogens on an open tube is replaced by a substituent group, e.g. OH or NH_2? As long as the tube is not of the type (n, 0), the effects are quite localized. In the vicinity of the substituent, $V_S(\mathbf{r})$ reflects its presence; elsewhere, however, $V_S(\mathbf{r})$ is essentially the same as for the original open tube, as described above. Table 2 includes three examples of this: (6, 1) $C_{68}H_{13}OH$; (6, 1) $C_{68}H_{13}NH_2$; and (5, 5) $C_{80}H_{19}OH$. The $V_{S,max}$ and $V_{S,min}$ are seen to be in the ranges found for hydroxyl and amino groups in typical organic molecules (Table 1), while the positive and negative averages over the whole surfaces are similar to the unsubstituted tubes in Table 2.

The results are dramatically different when substitution takes place on an (n, 0) tube. It develops a striking gradient in $V_S(\mathbf{r})$ along its entire length, from very positive at one end to very negative at the other. This can be seen in Figure 2, and is also brought out clearly by the data in Table 2. Both $V_{S,max}$ and $V_{S,min}$ are much larger in magnitude for the four substituted (6, 0) tubes than for any of the others in Table 2, as well as most of the molecules in Table 1. The same is true of $\bar{V}_S^+$ and $\bar{V}_S^-$, and particularly for the σ_{tot}^2, which are much higher than those in Table 1. The latter point reflects the remarkable variation in $V_S(\mathbf{r})$ from one end of the tube to the other.

For the (6, 0) tubes bearing OH or NH_2, $V_{S,max}$ is near the site of the substituent, and $V_{S,min}$ is at the other end. Since these groups are well known to be resonance donors of electronic charge (see Table 3), it appears that they are acting in that capacity here,

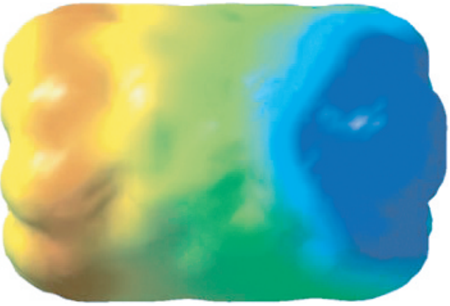
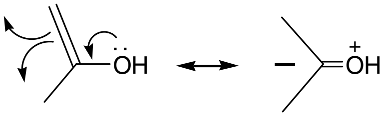


Figure 2 Computed (HF/STO-5G//STO-3G) electrostatic potential $V_S(\mathbf{r})$ on the outer surface of the (8, 0) $C_{80}H_{15}NH_2$ model carbon nanotube. The NH_2 group is at the right end. The variation in $V_S(\mathbf{r})$ is between −26.1 kcal/mole at the left end and +26.1 kcal/mole at the right

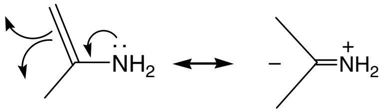
Table 3 Resonance (σ_R°) and inductive (σ_I) substituent constants [57]

Substituent	σ_R°	σ_I
OH	−0.40	0.25
NH_2	−0.48	0.17
NO_2	0.15	0.67
F	−0.34	0.54
CN	0.13	0.57
CH_3	−0.11	−0.01

and that the charge that they provide is delocalized along the full length of the tube, producing an increasingly negative $V_S(\mathbf{r})$:



1



2

This interpretation is supported by the fact that the computed C–OH and C–NH_2 distances in these (6, 0) tubes are shorter than usual, and that the C–NH_2 portion is planar [4, 11, 12], as predicted by structures **1** and **2**.

When the substituent is NO_2, which is strongly electron-withdrawing by induction (Table 3), then the electronic delocalization is toward its end, which becomes the site of

$V_{S,min}$. When both NH_2 and NO_2 are introduced, at opposite ends, their effects reinforce each other and the resulting σ^2_{tot}, 1036 (kcal/mole)^2 (Table 2), is one of the highest that we have ever encountered.

We have also investigated the effects of other substituents on (6, 0) systems, and found analogous tube-long polarization. For F and CN, the direction is the same as for NO_2, as expected on the basis of their relative inductive and resonance tendencies (Table 3); for CH_3, however, which is weakly donating through both resonance and induction, the direction is as for OH and NH_2.

As Figure 2 shows, it is not only substituted (6, 0) tubes that exhibit this dramatic polarization of $V_S(\mathbf{r})$; other (n, 0) do as well. It is notable that a single substituent can so affect the entire surface of the tube.

6. Average local ionization energies on model carbon nanotube surfaces

The average local ionization energies on the outer surfaces of carbon tubes have local minima, $\bar{I}_{S,min}$, above the carbons [4, 11, 12, 38], just as was found for graphene [56], but they are smaller in magnitude, mainly 13–14 eV, compared to 14.8–14.9 eV in the interior of graphene. The difference is due to curvature, which produces more negative $V_S(\mathbf{r})$ on the outsides of the tubes (as was already discussed) from which it is accordingly easier to remove an electron. For comparison, the HF/STO-5G//STO-3G $\bar{I}_{S,min}$ above the carbons of benzene are 14.3 eV [56]. ($\bar{I}_{S,min}$ are always somewhat higher than the $|\varepsilon_i|$ of the highest occupied orbitals because, as seen in Eq. (5), the former reflect some probability of inner, more-tightly-held electrons being at the point in question.)

Since lower $\bar{I}_{S,min}$ implies more energetic and easily removed electrons, higher levels of curvature, e.g. at caps, should be associated with greater chemical reactivity toward electrophiles. This has indeed been observed [1]. It has even been suggested that enhanced reactivity could be induced in carbon nanotubes by prior mechanical twisting or kinking [55, 58]; when the desired reaction was completed, the tube could be straightened.

When a substituent is introduced on a model tube other than (n, 0), the effect upon $\bar{I}_S(\mathbf{r})$ is largely localized to its vicinity, just as for $V_S(\mathbf{r})$ [4, 11, 12]. For (n, 0) systems, however, a tube-long gradation in $\bar{I}_S(\mathbf{r})$ is now superposed upon the pattern of $\bar{I}_{S,min}$ above the carbons. Overall, $\bar{I}_S(\mathbf{r})$ changes from low to high in the same direction as $V_S(\mathbf{r})$ goes from strongly negative to strongly positive. Thus, the polarization observed in $V_S(\mathbf{r})$ is reflected in $\bar{I}_S(\mathbf{r})$.

7. Polarization of unsubstituted model carbon nanotubes

We have attributed the striking polarization of the $V_S(\mathbf{r})$ of (n, 0) carbon tubes to the presence of substituents at the open ends. To our surprise, we have found that a similar effect, although weaker, can be produced even by minor structural asymmetry. This will be demonstrated for some relatively small systems, but qualitatively the same has been observed for larger ones.

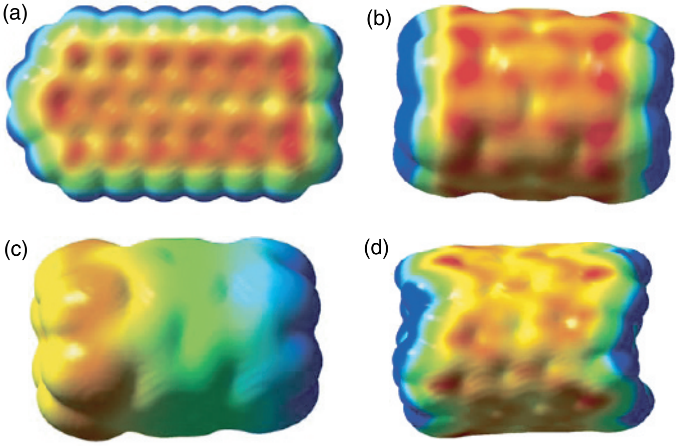


Figure 3 Computed (HF/STO-5G//STO-3G) surface electrostatic potentials $V_S(\mathbf{r})$. (a) Planar $C_{54}H_{20}$. (b) Symmetric (6, 0) $C_{48}H_{12}$ tube. (c) Asymmetric (6, 0) $C_{48}H_{12}$ tube. (d) (8, 5) $C_{104}H_{26}$ tube. In (a), (b) and (d), the range of values is less than 2 kcal/mole, from red (most negative) to blue (most positive). In (c), it is −3.9 (red) to +3.9 (blue) kcal/mole

Figure 3(a) shows $V_S(\mathbf{r})$ for a planar sheet of 18 carbon hexagons, arranged in 3 rows of 6 apiece, with hydrogens around the periphery. As in our previous study of model graphenes [56], $V_S(\mathbf{r})$ is weakly negative everywhere above the carbons, due to the hydrogens. There is no indication of any overall polarization.

This sheet of carbon hexagons can be rolled into a short (6, 0) tube. When the C–H distances are all constrained to be exactly the same, in this instance 1.070 Å, the resulting $V_S(\mathbf{r})$, in Figure 3(b), is typical of open model carbon nanotubes, as described earlier in this chapter. There is no polarization.

Figure 3(c) shows $V_S(\mathbf{r})$ for the same tube, but with the C–H bond lengths at the two ends differing slightly: 1.0870 Å and 1.0805 Å. This very small asymmetry is sufficient to cause a weak but clearly evident tube-long polarization of $V_S(\mathbf{r})$; the end with the shorter C–H bonds is negative, $V_{S,\text{min}} = -3.9\,\text{kcal/mole}$, and the other is positive, $V_{S,\text{max}} = 3.9\,\text{kcal/mole}$.

Analogous results have been obtained for other (*n*, 0) tubes, e.g. (8, 0) and (9, 0), although the polarization diminishes as the tube diameter increases and the curvature decreases. Thus for the (9, 0) of the same length, $V_{S,\text{min}} = -2.3\,\text{kcal/mole}$, $V_{S,\text{max}} = 2.3\,\text{kcal/mole}$. On the other hand, there is no polarization of the $V_S(\mathbf{r})$ of (8, 8) and (8, 5) tubes; see Figure 3(d). It appears to be only the (*n*, 0) that have this capability.

8. Why?

Why does asymmetry, whether due to substituents or to structure, lead to such distinctive polarization in (*n*, 0), and *only* (*n*, 0), carbon tubes? We will suggest some possible factors and offer some speculation.

A unique feature of $(n, 0)$ tubes, as can be seen in Figure 1, is that alternate C–C bonds are parallel to the tube axis. No other type of tube, (n, n) or (n, m), has any bonds at all with this property. Only for such bonds is 2pπ–2pπ overlap unaffected by curvature, so their presence might mean that $(n, 0)$ systems have an enhanced capacity for charge delocalization parallel to the axis. They do in fact have very high (computed) axial polarizabilities [47, 59–61], but this is also true of the (5, 5) [60, 61].

In any case, having axially oriented bonds cannot be enough to produce the polarization that is observed, since it does not occur in the planar analogues of $(n, 0)$ tubes, such as that in Figure 3(a). Curvature is evidently essential; indeed the polarization becomes weaker as the tube diameter increases. However, curvature is also not sufficient, since it is characteristic of all of the types of tubes, to varying degrees.

The key point may be the specific symmetry that open $(n, 0)$ systems have with respect to the tube axis. It is not only the longitudinal C–C bonds that are parallel to the axis, but also the C–H bonds and those to end substituents, if any. Another important element of axial symmetry is the "zig-zag" pattern of the transverse C–C bonds. As a result of these framework properties, which are specific to $(n, 0)$ systems, the electronic effects of even a minor perturbation at one end of the tube are transmitted around its entire circumference and along its full length. We continue to try to gain further insight into these phenomena.

9. A possible application: Nonlinear optics

The response of a molecule's dipole moment (or bulk matter's polarization) to an external electric field can be represented by Eq. (6) [62–66]:

$$\mu_i(\boldsymbol{\varepsilon}) = \mu_i(0) + \sum_j \boldsymbol{\alpha}_{ij}\boldsymbol{\varepsilon}_j + \frac{1}{2}\sum_{jk} \boldsymbol{\beta}_{ijk}\boldsymbol{\varepsilon}_j\boldsymbol{\varepsilon}_k + \frac{1}{6}\sum_{jkm} \boldsymbol{\gamma}_{ijkm}\boldsymbol{\varepsilon}_j\boldsymbol{\varepsilon}_k\boldsymbol{\varepsilon}_m + \cdots \qquad (6)$$

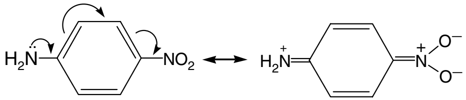
$\mu_i(\boldsymbol{\varepsilon})$ is the component of the dipole moment along an axis i in the presence of the electric field ε. The tensors $\boldsymbol{\alpha}, \boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are the polarizability and the first and second hyperpolarizabilities, respectively (or the first-, second- and third-order susceptibilities in the case of bulk polarization).

Nonlinear optical activity reflects an enhanced higher-order (i.e. nonlinear) response to the electric field associated with electromagnetic radiation; thus it depends upon $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$. A high level of nonlinear optical activity can have important applications in optoelectronics, e.g. in frequency converters, signal processors, switches, modulators, etc. [62, 63, 65, 67]. There have in fact been a number of investigations of the possible use of carbon nanotubes in optoelectronic devices [61, 68–75].

Molecules having electron donor and acceptor groups separated by a conjugated bridge, e.g. **3** and **4**, are known to be good candidates for showing enhanced second-order nonlinearity [62–66, 76]. The remarkably strong charge delocalization that we have found in substituted $(n, 0)$ model tubes suggests that these may be particularly suitable for such purposes. As an initial test, we estimated computationally, at the local density functional SVWN/6-31G* level, the second-order responses (i.e., the first hyperpolarizabilities) of *para*-nitroaniline (**4**) and our NH_2/NO_2-substituted (6, 0) model tube (Table 2) [11]. The value was nine times larger for the latter! It seems clear that appropriately substituted (*n*, 0) systems merit particular consideration in optoelectronics.

$$(N{\equiv}C)_2C{=}CH{-}C{\equiv}C{-}C_6H_4{-}N(CH_3)_2$$

3



4

10. Concluding remarks

While this chapter has dealt almost entirely with open (*n*, 0) model carbon nanotubes, both substituted and unsubstituted, we have also found very notable charge delocalization and transmission of electronic effects in capped (*n*, 0) $C_xB_yN_z$ tubes; see Politzer *et al.* [4, 12, 38]. The common element is the (*n*, 0) framework and the introduction of some perturbation of the symmetry of the system.

At present, we are aware only of our computational evidence for the very unusual features of (*n*, 0) systems, although it may be relevant to mention Zipper *et al.*'s observation of enhanced electrical conduction by (*n*, 0) carbon nanotubes [77, 78]. Given the rapid advances in technology and synthesis techniques, however, experimental confirmation may soon be forthcoming.

References

[1] P. M. Ajayan, *Chem. Rev.* 99 (1999) 1787–1800.

[2] P. J. F. Harris, *Carbon Nanotubes and Related Structures*, Cambridge University Press, Cambridge, UK (1999).

[3] C. T. White, J. W. Mintmire, *J. Phys. Chem. B* 109 (2005) 52–65.

[4] P. Politzer, J. S. Murray, P. Lane, M. C. Concha, in A. A. Balandin, W. L. Wang (eds), *Handbook of Semiconductor Nanostructures and Devices*, American Scientific Publishers, Los Angeles, 2006.

[5] M. Windholz (ed.), *The Merck Index*, 10th edn, Merck, Rahway, NJ (1983).

[6] J. K. Burdett, *Chemical Bonding in Solids*, Oxford University Press, New York (1995).

[7] E. Bengu, L. D. Marks, *Phys. Rev. Lett.* 86 (2001) 2385–2387.

[8] S. Y. Bae, H. W. Seo, J. Park, Y. S. Choi, J. C. Park, S. Y. Lee, *Chem. Phys. Lett.* 374 (2003) 534–541, and references cited.

[9] K. Suenaga, C. Colliex, N. Demoncy, A. Loiseau, H. Pascard, F. Willaime, *Science* 278 (1997) 653–655.

[10] D. Golberg, Y. Bando, M. Mitome, K. Kurashima, N. Grobert, M. Reyes-Reyes, H. Terrones, M. Terrones, *Chem. Phys. Lett.* 360 (2002) 1–7.

[11] P. Politzer, J. S. Murray, P. Lane, M. C. Concha, P. Jin, Z. Peralta-Inga, *J. Mol. Model.*, 11 (2005) 258–264; *J. Mol. Model.* 12 (2006), 528 (Erratum).

[12] P. Politzer, J. S. Murray, P. Lane, M. C. Concha, in W. A. Sokalski (ed.), *Molecular Materials with Specific Interactions: Modeling & Design*, Springer, London, 2006, ch. 5.

[13] R. F. Stewart, *J. Chem. Phys.*, 57 (1972) 1664–1668.
[14] P. Politzer, D. G. Truhlar (eds), *Chemical Applications of Atomic and Molecular Electrostatic Potentials*, Plenum Press, New York (1981).
[15] G. Naray-Szabo, G. G. Ferenczy, *Chem. Rev.* 95 (1995) 829–847.
[16] R. F. W. Bader, M. T. Carroll, J. R. Cheeseman, C. Chang, *J. Am. Chem. Soc.* 109 (1987) 7968–7979.
[17] J. S. Murray, P. Politzer, in J. S. Murray, P. Politzer (eds), *Quantitative Treatments of Solute/Solvent Interactions*, Elsevier, Amsterdam (1994) ch. 8.
[18] J. S. Murray, P. Politzer, *J. Mol. Struct. (Theochem)* 425 (1998) 107–114.
[19] P. Politzer, J. S. Murray, *Trends Chem. Phys.* 7 (1999) 157.
[20] P. Politzer, J. S. Murray, *Fluid Phase Equilib.* 185 (2001) 129–137.
[21] P. Sjoberg, J. S. Murray, T. Brinck, P. Politzer, *Can. J. Chem.* 68 (1990) 1440–1443.
[22] T. A. Koopmans, *Physica* 1 (1933) 104–113.
[23] R. K. Nesbet, *Adv. Chem. Phys.* 9 (1965) 321.
[24] P. Politzer, M. E. Grice, J. S. Murray, *Coll. Czech. Chem. Comm.* 70 (2005) 550–558.
[25] P. Politzer, J. S. Murray, M. E. Grice, T. Brinck, S. Ranganathan, *J. Chem. Phys.* 95 (1991) 6699–6704.
[26] Nagy, R. G. Parr, S. Liu, *Phys. Rev. A* 53 (1996) 3117–3121.
[27] P. Politzer, J. S. Murray, in A. Toro-Labbé (ed.), *Theoretical Approaches to Chemical Reactivity*, Elsevier, Amsterdam, 2006, in press.
[28] P. Jin, T. Brinck, J. S. Murray, P. Politzer, *Int. J. Quantum Chem.* 95 (2003) 632–637.
[29] P. Jin, J. S. Murray, P. Politzer, *Int. J. Quantum Chem.* 96 (2004) 394–401.
[30] J. S. Murray, J. M. Seminario, P. Politzer, P. Sjoberg, *Int. J. Quantum Chem., Quantum Chem. Symp.* 24 (1990) 645–653.
[31] J. S. Murray, F. Abu-Awwad, P. Politzer, *J. Mol. Struct. (Theochem)* 501 (2000) 241–250.
[32] P. Politzer, F. Abu-Awwad, J. S. Murray, *Int. J. Quantum Chem.* 69 (1998) 607–613.
[33] J. S. Murray, Z. Peralta-Inga, P. Politzer, K. Ekanayake, P. LeBreton, *Int. J. Quantum Chem.* 83 (2001) 245–254.
[34] P. Politzer, J. S. Murray, M. C. Concha, *Int. J. Quantum Chem.* 88 (2002) 19–27.
[35] T. Brinck, J. S. Murray, P. Politzer, *Int. J. Quantum Chem.* 48 (1993) 73–88.
[36] Z. Peralta-Inga, P. Lane, J. S. Murray, S. Boyd, M. E. Grice, C. J. O'Connor, P. Politzer, *Nano Lett.* 3 (2003) 21–28.
[37] P. Politzer, P. Lane, J. S. Murray, M. C. Concha, *J. Mol. Model.* 11 (2005) 1–7.
[38] P. Politzer, P. Lane, M. C. Concha, J. S. Murray, *Microelectr. Eng.* 81 (2005) 485–493.
[39] P. Politzer, J. S. Murray, in K. B. Lipkowitz and D. B. Boyd (eds), *Reviews in Computational Chemistry*, vol. 2, VCH Publishers, New York (1991), ch. 7, and references cited.
[40] J. S. Murray, T. Brinck, P. Politzer, *J. Mol. Struct.* (*Theochem*) 255 (1992) 271–281.
[41] C. Dillon, K. M. Jones, T. A. Bekkedahl, C. H. Kiang, D. S. Bethune, M. J. Heben, *Nature* 386 (1997) 377–378.
[42] Kuznetsova, J. T. Yates, Jr., J. Liu, R. E. Smalley, *J. Chem. Phys.* 112 (2000) 9590–9598.
[43] Fujiwara, K. Ishii, H. Suematsu, H. Kataura, Y. Maniwa, S. Suzuki, Y. Achiba, *Chem. Phys. Lett.* 336 (2001) 205–211.
[44] M. S. C. Mazzoni, H. Chacham, P. Ordejon, D. Sanchez-Portal, J. M. Soler, E. Artacho, *Phys. Rev. B* 60 (1999) R2208–R2211.
[45] W. Bauschlicher, Jr., *Nano Lett.* 1 (2001) 223–226.
[46] S. Irle, A. Mews, K. Morokuma, *J. Phys. Chem. A* 106 (2002) 11973–11980.
[47] M. D. Halls, H. B. Schlegel, *J. Phys. Chem. B* 106 (2002) 1921–1925.
[48] R. L. Jaffe, *J. Phys. Chem. B* 107 (2003) 10378–10388.
[49] Z. Zhou, M. Steigerwald, M. Hybertsen, L. Brus, R. A. Friesner, *J. Am. Chem. Soc.* 126 (2004) 3597–3607.
[50] P. Politzer, J. S. Murray, Z. Peralta-Inga, *Int. J. Quantum Chem.* 85 (2001) 676–684.

[51] J. S. Murray, P. Lane, P. Politzer, *Mol. Phys*. 93 (1998) 187–194.
[52] H. Weinstein, P. Politzer, S. Srebrenik, *Theor. Chim. Acta* 38 (1975) 159.
[53] K. D. Sen, P. Politzer, *J. Chem. Phys*. 90 (1989) 4370.
[54] J. Chen, R. C. Haddon, S. Fang, A. M. Rao, W. H. Lee, E. C. Dickey, E. A. Grulke, J. C. Pendergrass, A. Chavan, B. E. Haley, R. E. Smalley, *J. Mater. Res*. 13 (1998) 2423.
[55] Srivastava, D. W. Brenner, J. D. Schall, K. D. Ausman, M.-F. Yu, R. S. Ruoff, *J. Phys. Chem. B* 103 (1999) 4330–4337.
[56] Z. Peralta-Inga, J. S. Murray, M. E. Grice, S. Boyd, C. J. O'Connor, P. Politzer, *J. Mol. Struct.* (*Theochem*) 549 (2001) 147–158.
[57] O. Exner, *Correlation Analysis of Chemical Data*, Plenum Press, New York (1988).
[58] Dekker, *Phys. Today* 52(5) (1999) 22.
[59] L. X. Benedict, S. G. Louie, M. L. Cohen, *Phys. Rev. B* 52 (1995) 8541–8549.
[60] L. Jensen, O. H. Schmidt, K. V. Mikkelsen, P.-O. Åstrand, *J. Phys. Chem. B* 104 (2000) 10462–10466.
[61] L. Jensen, P.-O. Åstrand, K. V. Mikkelsen, *J. Phys. Chem. A* 108 (2004) 8795–8800.
[62] S. R. Marder, J. E. Sohn, G. D. Stucky (eds), New Materials for Nonlinear Optics, ACS Symposium Series 455, American Chemical Society, Washington, DC, 1991.
[63] R. Kanis, M. A. Ratner, T. J. Marks, *Chem. Rev*. 94 (1994) 195–242.
[64] N. Matsuzawa, D. A. Dixon, *J. Phys. Chem*. 98 (1994) 2545–2554.
[65] S. P. Karna, *J. Phys. Chem. A* 104 (2000) 4671–4673.
[66] R. Andreu, M. J. Blesa, L. Carrasquer, J. Garin, J. Orduna, B. Villacampa, R. Alcala, J. Casado, M. C. R. Delgado, J. T. L. Navarrete, M. Allain, *J. Am. Chem. Soc*. 127 (2005) 8835–8845 and references cited.
[67] W. L. Wilson, in J. H. Moore and N. D. Spencer (eds), *Encyclopedia of Chemical Physics and Physical Chemistry*, vol. III, Institute of Physics, London, 2001, C2.15.
[68] X. Wan, J. Dong, D. Y. Xing, *Phys. Rev. B* 58 (1998) 6756–6759.
[69] X. Liu, J. Si, B. Chang, G. Xu, Q. Yang, Z. Pan, S. Xie, P. Ye, J. Fan, M. Wan, *Appl. Phys. Lett*. 74 (1999) 164–166.
[70] J. Jiang, J. Dong, D. Y. Xing, *Phys. Rev. B* 59 (1999) 9838–9841.
[71] Ya Slepyan, S. A. Maksimento, V. P. Kalosha, J. Herrmann, E. E. B. Campbell, I. V. Hertel, *Phys. Rev. A* 60 (1999) R777–R780.
[72] P. Chen, X. Wu, X. Sun, J. Lin, W. Ji, K. L. Tan, *Phys. Rev. Lett*. 82 (1999) 2548–2551.
[73] S. R. Mishra, H. S. Rawat, S. C. Mehendale, K. C. Rustagi, A. K. Sood, R. Bandyopadhyay, A. Govindaraj, C. N. R. Rao, *Chem. Phys. Lett*. 317 (2000) 510–514.
[74] S. Barazzouk, S. Hotchandani, K. Vinodgopal, P. V. Kamat, *J. Phys. Chem. B* 108 (2004) 17015–17018.
[75] D. A. Stewart, F. Léonard, *Nano Lett*. 5 (2005) 219–222.
[76] S. R. Marder, D. N. Beratan, L.-T. Cheng, *Science* 252 (1991) 103.
[77] M. Szopa, M. Marganska, E. Zipper, M. Lisowski, *Phys. Rev. B* 70 (2004) 75406–75412.
[78] E. Zipper, M. Szopa, M. Marganska, M. Lisowski, Nano and Giga Challenges Conference, Jagellonian University, Krakow, Poland, September 2004.

Molecular and Nano Electronics: Analysis, Design and Simulation
J. M. Seminario (Editor)

Chapter 4

Analysis of programmable molecular electronic systems

Yuefei Ma and Jorge M. Seminario

Department of Chemical Engineering, and Department of Electrical and Computer Engineering, 3122 TAMU, Texas A&M College Station TX77843, USA.
seminario@tamu.edu

1. Introduction

The continuing scaling down in size of microelectronics devices has motivated the development of molecular electronics, often called moletronics, which uses molecules to function as electronic devices [1–14]. One of the goals of moletronics is the construction of programmable molecular arrays [15, 16]. We will focus this chapter on one of the scenarios for these devices, a quasi-ordered array of metallic islands interlinked by molecules that are addressed by a small number of input/output leads located on the periphery of this programmable molecular array, which is also called nanoCell [15], and which is populated with different molecules interconnecting the metallic islands. Experimentally and theoretically, the electrical characteristics of the nanoCells suggest that the metallic islands contribute to some nonlinear features of the current–voltage characteristics, such as negative differential resistance and memory phenomena [16–19]. The importance of high speed electronics also compels us to carefully understand the limitations faced by microelectronics due to device fabrication and physical limits. Therefore, challenging new techniques are described that may complement conventional microelectronics and permit the attainment of higher computing speeds.

1.1. Importance and current status of high speed electronics

In 1965, only seven years after Jack Kilby invented the integrated circuit, Gordon Moore made his famous observation popularly known as "Moore's Law" [20, 21]. Moore's Law predicted that the number of transistors per integrated circuit would double every 18 months, leading to an exponential growth in the complexity of devices. Amazingly, this empirical rule is still holding true. The number of transistors in the 4004 processor was 2,250 in the year 1971. The number of transistors increased to 275,000 for the Intel386™ processor in 1985. By 2003, it grew to 410,000,000 in the Intel® Itanium® 2 processor. Then, on 25 January 2006, Intel demonstrated a fully functional 45 nm SRAM chip with more than a billion transistors.

The fundamental driving force behind Moore's Law is the constant craving for ultrafast computing. It is always said, "Time is money"; in fact time is everything. As the technology grows, the need for faster computation is demanded in many disciplines, such as mathematics, chemistry, physics, meteorology, etc. Faster computation facilitates state-of-the-art research and makes it more productive and efficient. It would be better to finish a computation in a shorter time, because it will shorten the time needed for feedback and thus for completion of a project. Also, faster computers means faster communication, which will bring people closer together, all around the world. In the meantime, the information being processed is growing exponentially, which requires a larger number of memory devices and faster speed to process and deliver all the information. All these computational and memory functions require more and more transistors in a smaller space. Thus, the size of transistors in a single chip needs to shrink tremendously.

1.1.1. Limitations in device fabrication

The continuous growth resulting from following Moore's law creates obstacles in device fabrication in the following two aspects.

1. Lithography Since the resolution of lithography depends on the wavelength of the light source, a light source with shorter wavelength is needed to create smaller images. New techniques such as phase-shift masks and optical proximity correction make it possible to print smaller patterns than those normally expected from the wavelength of the light source. For example, currently, 193 nm light is used to produce a minimum feature size of 45 nm. However, the goal of feature sizes of 30 nm or less is still pushing us to seek new light sources with shorter wavelength. As researchers search for new light sources, technologies for improved photo resist, mask, and aligner systems need to be developed simultaneously. Some alternatives have been found, such as E-beam and X-ray lithography, which can yield nano-sized features, but these alternatives are not suitable for mass production [22]; fabrication with feature sizes smaller than 45 nm in large-scale production is extremely difficult.
2. Doping Semiconductor materials are made conductive by adding certain impurities into the bulk. This process of doping is done by diffusion or ion implantation. Both processes face a limitation of solid solubility, meaning that the maximum number of impurity atoms that can exist in a solid is limited. Thus, as the size of the device is decreased, it might be possible in the future that only one or a few doping atoms will exist in the source/drain area, resulting in fuzzy boundaries and therefore bad performance.

1.1.2. Limitations in device operation

The building blocks of current electronic circuits are Complementary Metal Oxide Semiconductor (CMOS) Field Effect Transistors (FET), in which both NMOSs and

PMOSs are used to implement the logic functions. In NMOS, electrons are majority carriers in the channel region; while in PMOS, holes are majority carriers. According to the International Technology Roadmap for Semiconductors (ITRS), CMOS technology will still be around until 2015 [23]. Even if some of the difficulties in device fabrication are overcome, there will still be a number of intractable issues in solid-state physics as follows.

- Velocity saturation in Metal Oxide Semiconductor Field Effect Transistors (MOSFET) The speed of carriers in a MOSFET is related to the mobility of the majority carriers in the channel, i.e., the electron/hole mobility. The mobility is inversely related to the lateral electric field across the channel. As the device is scaled down, the electric field increases, which results in decreasing mobility. Therefore it restricts the MOSFET current and the speed of the device.
- Punch through During normal operation, the current in a MOSFET is carried through the channel, which is induced by the voltage applied to the gate. However, if the device is scaled down so that the channel is extremely short, the depletion region of the drain expands, because of the application of the drain voltage. Eventually, it will touch the source depletion region. Thus, the electrons will "punch through" from the source to the drain, resulting in leakage current.
- Parasitic capacitance As the device is scaled down, the surface-to-bulk ratio increases. This may increase the parasitic capacitance of the Gate-Drain capacitor, the Gate-Source capacitor, the Drain-Bulk capacitor, the Source-Bulk capacitor, etc. The general problem is that large capacitors decrease the speed of the device.
- Gate leakage Scaling of the MOSFET results in an extremely thin gate oxide and due to the "uncertainty principle" between the velocity and the momentum of the electron, it is impossible to find an electron in an exact position [24]. As is predicted by quantum mechanics, electrons tunnel through the thin gate resulting in a reduced gate field. To maintain the same channel depth as well as the same channel current, larger voltages must be applied to the gate, which increase the power consumption.

1.2. Future perspective of electronics

We study the electrical conductance through the metallic islands and their conformational changes under bias. Furthermore, a scenario is proposed to use molecular vibronics and electrostatic potentials to transport and process signals inside the programmable molecular array [16, 19, 25–36].

All these limitations urge us to search for new solutions for further expanding Moore's Law. Amazingly, three years before the publication of Moore's Law, Richard P. Feynman gave a talk, "There is plenty of room at the bottom." He spoke about the problem of manipulating and controlling things on a small scale. This seminal talk laid the foundation for the development of nanotechnology and molecular electronics; thus, a number of possible solutions have been found for the end of the Silicon era:

- Single molecular transistor The fundamental idea behind the single molecular transistor is to use a single molecule to function as an electronic device. This will decrease the size of a device tremendously. Since the seminal paper describing a molecular rectifier [37] was published in 1974, the field of molecular electronics has grown rapidly. However, fine lithography remains as a limitation for its realization as single molecules need to be addressed precisely.

- Programmable molecular array A programmable molecular array is a two-dimensional structure made of chemically arranged molecules. Microsized metallic input and output leads are located on the periphery of the structure. Thus, there are tens or hundreds of single molecular devices connected in series–parallel between input and output. This intriguing feature further shrinks the dimensions of devices. However, programmable molecular array has to be programmed by applying a sequence of voltage pulses on the metallic leads in order to perform specific post-fabrication functions.
- Use of molecular vibrational modes and molecular potential to transmit and process information When an atom in a molecule is displaced, the displacement signal will be transmitted through the molecule by the molecular vibrational modes. It is similar to the case of a mass-and-spring system. Furthermore, the displacement introduces a change in the distribution of the molecular electrostatic potential (MEP). By combining these two methods, the signals can be transmitted and processed without any electron current.
- Spintronics Spintronics is a technology that uses electron spin (and sometimes the nuclei spin), instead of using electron charge to store and transfer information [38, 39]. The spin can be detected as a weak magnetic energy state characterized as "spin up" or "spin-down". Spintronics has been successfully applied to a device called a spin valve, which utilizes a layered structure made of thin films of magnetic materials to change the electrical resistance which depends on the direction of magnetic field being applied. Currently, researchers are developing new magnetic semiconductors based on room-temperature ferromagnetism.
- Cross bar approach In a cross bar structure, two layers of regularly arranged nanowires (or nanotubes) are crossing each other with electrically switchable molecules. By applying a sequence of voltage impulses to the nanowires and using switches of opposite polarities, the device can perform specific logic functions. In addition, it can restore a logic level in a circuit to its ideal voltage value, allowing a designer to chain many simple gates together to perform an arbitrary computation.

In addition to extending the unique scenario based on the charge-current to process and encode information in integrated circuits to the molecular electronic systems, a dual scenario is proposed to use molecular vibronics and electrostatic potentials to transport and process signals inside programmable molecular arrays [16, 19, 25–36].

2. Programmable molecular arrays

In this section, first the programmable molecular array will be introduced. Next, related research work in other research groups is reviewed. Then the fabrication process of the programmable molecular array as well as the process of forming self-assembled monolayers will be explained. And then the electrical measurement set-up is described. Last, the electrical characterization of memory and switching phenomena, start-up transitional behavior, effect of different molecular depositions, and programmability of the device are investigated.

2.1. Introduction to programmable molecular cells

As conventional silicon-based microelectronics encounters more and more challenges when it scales down to smaller sizes, several companies and research groups are trying to incorporate single or small groups of molecules into electronic circuits. These molecules are expected to implement complex computations, for example, logic gates, memories, etc. By doing so, the size of a rather complicated circuit can be greatly reduced, and at the meantime it will bring down the fabrication cost. However, to address a single molecule is another challenge for conventional fabrication techniques, especially lithography. This is the reason why molecular programmability has been proposed [40–43].

A programmable molecular array can also be called a molecular random access memory cell or nanoCell. It is a two-dimensional structure constructed on an insulator, e.g., SiO_2. Inside a nanoCell, there are chemically organized molecules and metallic nanoclusters, nanowires and/or nanotubes, as shown in Figure 1.

The nanowires or nanotubes serve as anchors to attach the molecules by self-assembled monolayer techniques. They can also conduct electrical current. Microsized metal leads are located on the periphery of the two-dimensional structure to allow the interconnection between the molecules and external circuit. It has been shown theoretically that some molecules can function as electronic device [3, 8, 44, 45]. Therefore, the nanoCell can be viewed as hundreds of devices connecting any of the two contacts in series-parallel. After fabrication, the nanoCell will be programmed by a computer program to function as the purpose of desired device.

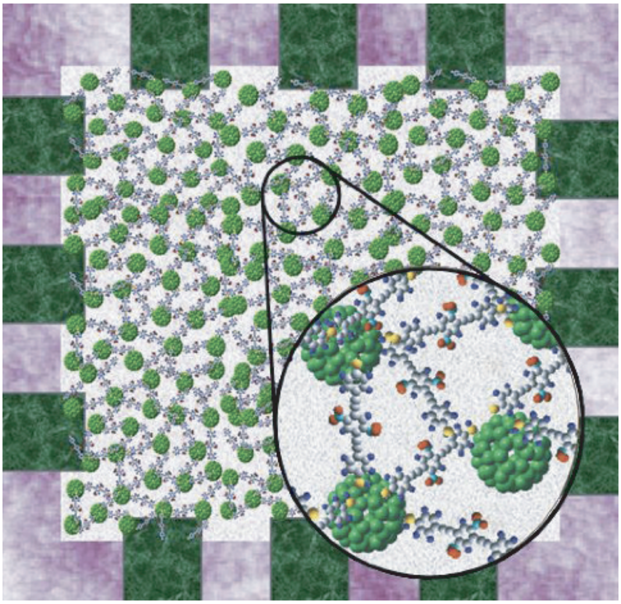


Figure 1 Schematic of a nanoCell composed of gold clusters (green) and interlinking molecules (grey) sitting on a substrate of SiO_2 (white). The electrodes (green) are located around the nanoCell [15]

The advantage of a nanoCell is to relieve the painful fabrication process, especially for photolithography, which currently cannot fabricate features at molecular size. The smallest feature in a nanoCell that needs to be defined by photolithography is the metal lead. Except for this, the components of a nanoCell are chemically located. Since the structure of a nanoCell is mainly made of molecules, it also has the advantage of reducing power consumption and heat dissipation. In addition, it allows the device to be advanced in the future to conquer the physical and technical limitations of silicon-based electronics.

The programmable molecular array was originally proposed by a group of researchers from University of South Carolina (USC), Yale University, and Penn State University (PSU) [40]. The preliminary idea is to chemically place the molecules inside a box and address them by using an external electrical field. The simulation conducted by Seminario *et al.* shows that the nanoCell can function as a logic device and can be programmable after fabrication [44]. The first two-dimensional nanoCell was fabricated by Tour and Franzon. [43, 46]. Since it is almost impossible for the molecules to lie parallel to the substrate surface and to interconnect each other, a discontinuous gold film is vapor deposited inside the box [46]. Molecules are then self-assembled onto the gold islands through which the molecules are interconnected. To enhance conductance, molecules covered by nanowires are assembled onto the surface after being deposited [43]. The first molecule being deposited onto the nanoCell was mononitro OligoPhenylene Ethynylene (OPE). It exhibits Negative Differential Resistance (NDR) between Au probes both experimentally and theoretically [8, 47, 48]. The NDR is believed to be essential to obtain switching effects [42, 49]. The current–voltage characteristics of OPE deposited nanoCell were studied and reported. Switching effects and memory effects have been demonstrated in the device.

At the same time, Husband *et al.* addressed the issue relating to programming the nanoCell to perform specific functions [42, 49]. In a computer simulation, they were able to program a nanoCell into a half adder. However, successfully programming a nanoCell in real experiments has never been reported.

2.2. Sample fabrication

The fabrication process of the nanoCell device can be divided into two parts based on the components: the substrate and the molecules.

2.2.1. Substrate fabrication

The first nanoCell substrate (Figure 2) was fabricated by researchers at North Carolina State University (NCSU). As shown in Figure 2, instead of metallic nanoclusters, a discontinuous gold film is deposited onto the p-type Si substrate uniformly covered with thermally grown wet SiO_2 [46]. The thickness of the SiO_2 film is 1000 Å. The patterning of the discontinuous gold film is performed using a lift-off technique. In this technique, a layer of photoresist is deposited onto the substrate and a lithography process is used to define the pattern. Afterward, a discontinuous gold film is vapor deposited in high vacuum ($\sim 10^{-10}$ torr). During the deposition, the substrate temperature is maintained at 100 °C in order to increase the stability of the film. Standard silicon process techniques are utilized to fabricate the contact leads and contact pads.

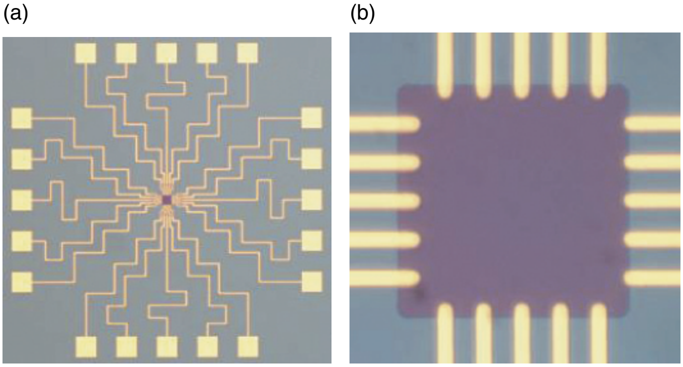


Figure 2 Optical microscopic images of an experimentally constructed nanoCell. (a) Optical microscopic image of a nanoCell. The yellow squares are contact pads. The purple square in the middle is a discontinuous gold film. The meander lines are contact leads. (b) Enlarged view of a discontinuous gold film

According to the scanning atomic force microscope (AFM, Digital Instrument) image (Figure 3a), the diameter of the gold islands ranges from 5 to 60 nm and the separation between the islands is around 5 nm, which is sufficiently small to fit two molecules. Atomic force microscope (AFM, Digital Instruments) shows that the thickness of the film ranges from 3 to 4 nm in Figure 3 However, because the AFM tip might be larger than or about the same size as the gap between the gold islands, the tip might not touch the substrate when it scans through the device. So, the thickness of the gold film might be larger than 4 nm.

2.2.2. Formation of self-assembled monolayers on nanoCell

Two kinds of molecules were chosen separately to deposit on different chips. Molecule **1** (Figure 4) is 4, 4′-(diethynylphenyl)-2′-nitro-1-benzenethioacetyl, which belongs to the mononitro oligo (phenylene ethynylene) (OPE) group. The molecule is a three-benzene-ring oligomer with two nitro group (NO_2) substituents in the central ring. Sulfur atoms acting as “alligator clips” join one end of the molecular devices with the Au clusters. Electrical properties of the molecule have been studied both experimentally and theoretically. The study shows that the molecule yields NDR which makes it a potential molecular electronic device [47, 48]. This is the reason why this molecule has been chosen for this research work. The molecule is synthesized by Tour’s group in Rice [47]. The acetyl group is used to protect the molecule against reaction with oxygen during transportation and it is removed in situ during the self-assembly process under acid conditions ($CH_2Cl_2/MeOH/H_2SO_4$), which is reported to yield better results than the traditional NH_4OH/THF mixture [51]. Once the acetyl group is removed, the S^- ion will attach to the Au atoms, resulting in a self-assembled monolayer (SAM).

Molecule **2** (Figure 4) is octyltrichlorosilane, which self-assembles on SiO_2 [52]. The adsorption of the molecule onto the SiO_2 substrate takes place through the hydrolysis of the Si-Cl bonds to form Si-OH groups which interact with OH groups on the SiO_2 surface and form Si-O-Si bonds.

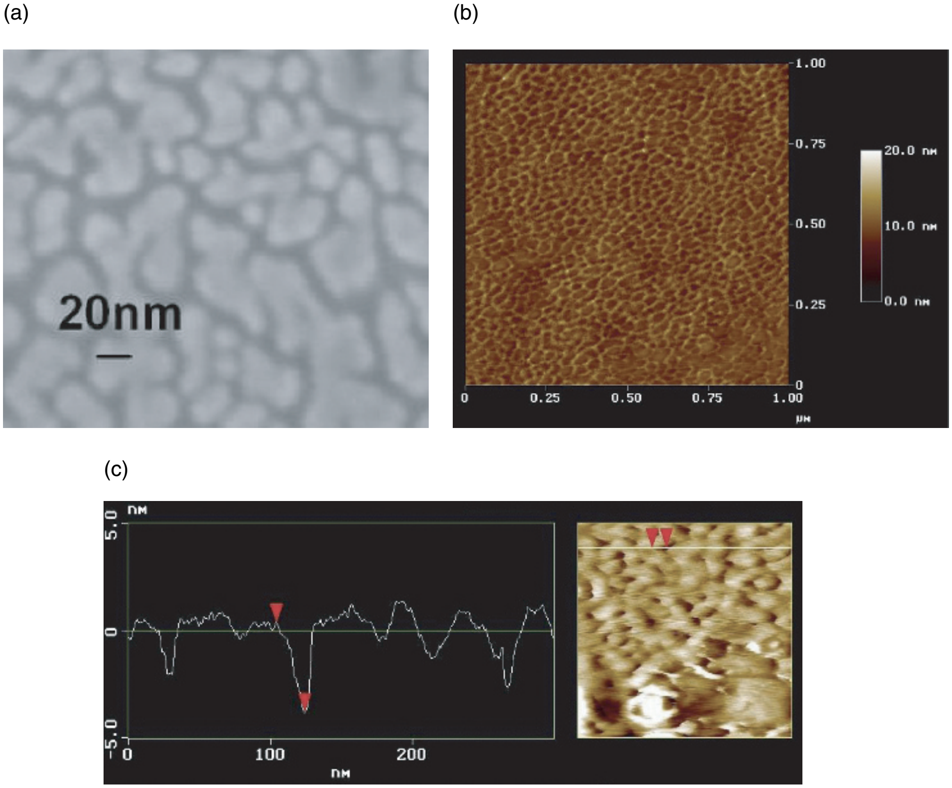
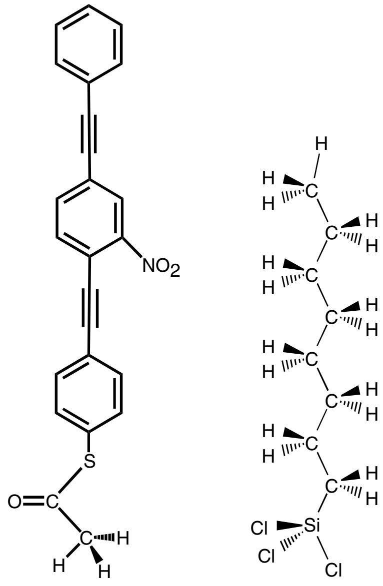


Figure 3 (a) Scanning electron microscope image of a discontinuous gold film [50]. (b) Atomic force microscope image of a discontinuous gold film. (c) Cross-section analysis of the gold film

Two different self-assembly procedures are performed at room temperature. Before the self-assembly, the nanoCell chip is cleaned using methanol and acetone, then rinsed with distilled water and dried by using nitrogen gas. For self-assembling molecule **1**, 1 mM of molecule **1** is added to a solvent composed of methylene dichloride and methanol (2:1). The nanoCell chip is soaked for approximately 12 hours in the solution. During the reaction, the acetyl group at the end of the molecule **1** is cleaved and then reacts with the gold surface, forming SAMs. Finally, it is rinsed by methanol to stop the reaction and dried by nitrogen gas. For self-assembling **2**, the nanoCell chip is soaked for 12 hours in 1% toluene solution. During the reaction, the three chlorine atoms are displaced by surface hydroxyls on the SiO_2 substrate. Thus the molecules are self-assembled on the SiO_2 surface. It is then rinsed by methanol and dried by nitrogen gas.

2.3. Measurement set-up

The nanoCell devices are measured in a probe station (Lakeshore Cryogenic probe station). Since the system being tested involves nanoscale features, such as molecules and gold islands, any free particles in the ambient environment might interfere with the



1 **2**

Figure 4 Molecules used for self assembling. **1** yields the 4, 4′-(diethynylphenyl)-2′-nitro-1-benzenethiolate during self-assembly wherein the acetyl group ($COCH_3$) is cleaved and the sulfur attaches to the gold islands. **2** is the octyltrichlorosilane. During the self-assembly process, the three chlorine atoms are displaced by surface hydroxyls on the SiO_2 substrate [50]

normal operation of the system. In order to exclude this environmental influence, the sample is placed inside a high vacuum (about 10^{-7} torr) chamber (Figure 5) during the measurement.

The temperature of the sample stage inside the vacuum chamber can be reduced by a continuous flow of liquid nitrogen, which is pumped out from a nitrogen dewar by a flow of nitrogen gas. A temperature sensor is attached to the sample holder inside the vacuum chamber. The other end of the temperature sensor is connected to a temperature controller. Users could input the set point of the desired temperature and the temperature controller calculated the output power of a heater, which is located at the chamber wall, based on the difference between the desired and the current temperature. Thus, the temperature gradually approaches the set point. In this chapter, the nanoCell devices are analyzed at room temperature (297 K).

The electrical measurement is performed by a HP 4145 semiconductor parameter analyzer remotely controlled by a computer. The applied voltage is swept in a staircase manner from start voltage V_1 to stop voltage V_2 in a voltage step of ΔV. The current is measured at the end of each voltage step.

2.4. Electrical characteristics of nanoCells

Current–voltage measurement is carried out on the simple structure with only two opposing electrodes. There are only OPE molecules self-assembled in the nanoCell.

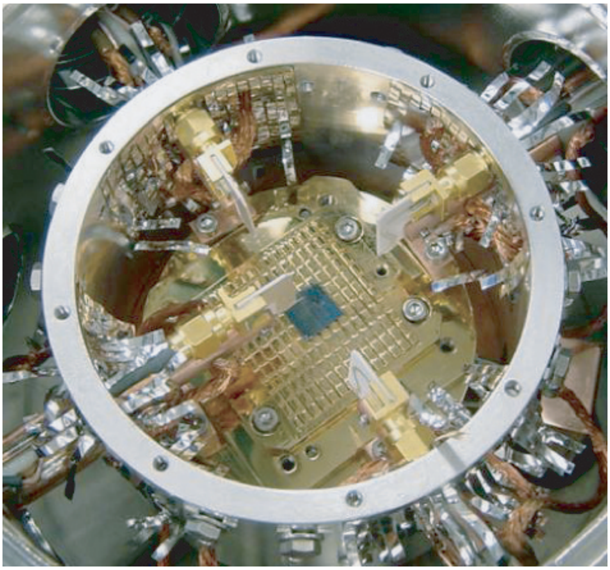


Figure 5 Probe station (Lakeshore Cryogenic) used to measure the nanoCell device. The probe station provides a high vacuum environment ($\sim 10^{-7}$ torr) to eliminate any free particles in the neighborhood of the sample

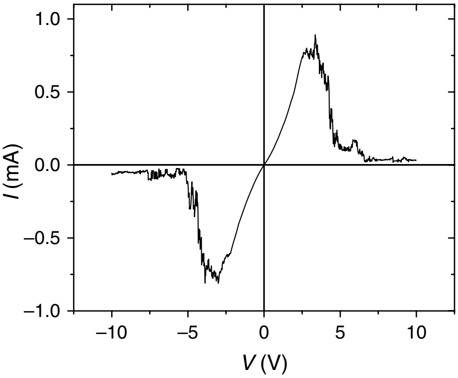


Figure 6 Typical NDR-like behavior shown in nanoCell device

Most of the devices tested exhibit NDR-like behavior, as shown in Figure 6, i.e., the current changes inversely according to the applied voltage. Thus, NDR-like behavior has two conductive states at the same current. The current peak value and position, and even the number of peaks may vary for each device. In addition, the devices show similar current–voltage characteristics under both polarities of biased voltage, although the peak value and the position may not be exactly the same.

2.4.1. Transition states before NDR

Initially, no NDR behavior is found in the nanoCells; instead, transition states are found before NDR appeared. Based on the tested nanoCells, 50% of them exhibit initial transitional behavior as shown in Figure 7. During the first voltage sweeps from 0 to 5 V, the nanoCell exhibits repeatable linear I-Vs (Figure 7a). When the voltage sweeps from 0 to 10 V, the current drops sharply at a certain voltage $V_{\text{th1}} = \sim 6.8$ V (Figure 7b). When a voltage less than 6 V is applied, the current is relatively low, between 9 and 10 nA, as shown in Figure 7c. If the voltage is above 6 V, the current increases about 3 orders of the magnitude, as shown in Figure 7c. Therefore, voltage of 6 V is defined as the threshold voltage, which is denoted as V_{th2}. Repeatable NDR-like characteristics appear in the subsequent voltage sweeps (Figure 7d). This sequence of events is referred to as the first observed initial transitional behavior, which is composed of a high conductance ohmic behavior (Figure 7a), a breakdown behavior (Figure 7b) and a transitional behavior (Figure 7c).

The initial transition states of the other 50% of the nanoCells do not contain the high conductance ohmic behavior and the breakdown behavior. Instead, they contain only the transitional I-V. It is referred to as the second observed initial transitional behavior. For

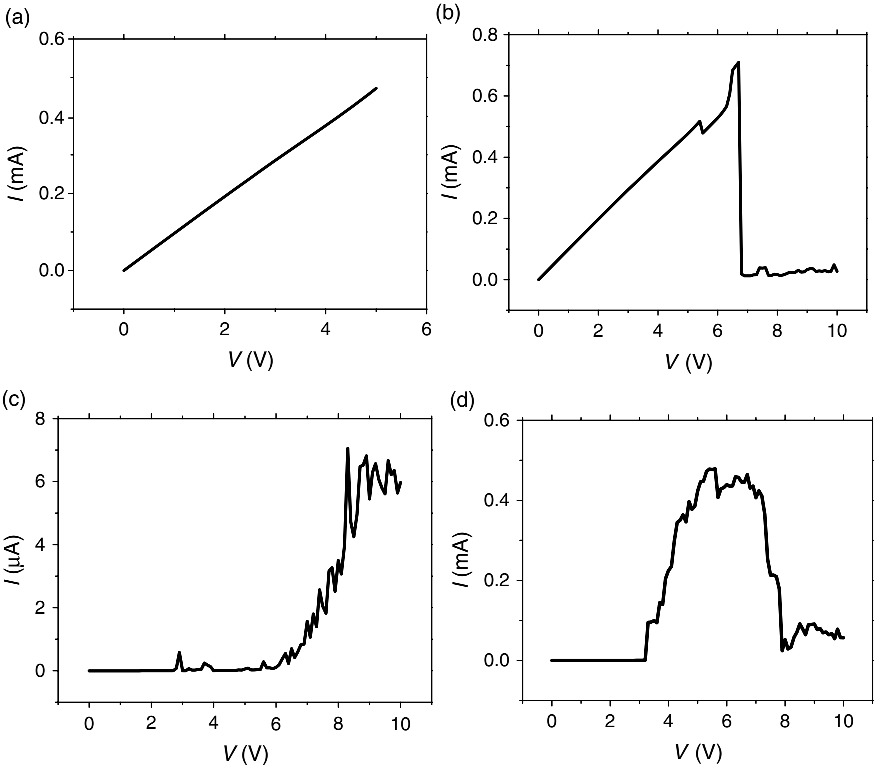


Figure 7 First observed current–voltage transition behavior of a nanoCell device. (a) high conductance ohmic I-V from 0 to 5 V; (b) breakdown once the bias voltage exceeds ~6.8 V; (c) transitional I-V, low current increases sharply at about 6 V; (d) NDR-like behavior [50]

the last 10% of the nanoCells, the transitional I-V does not occur before NDR appears. It is referred to as the third observed initial transitional behavior.

An interesting feature of all these three transitional behaviors is that it is not reversible, i.e., the nanoCell cannot be switched back to the original state by the application of a biased voltage. For instance, for the first observed one, once the NDR appears, the ultra-high conductance ohmic behavior cannot be observed. In addition, on unbiased nanoCell devices, negative voltage sweeps also induce similar sequences of I-Vs and finally reach the NDR-like behavior. However, on biased nanoCells, after the NDR has appeared in the forward biased range, it also shows up in the negative biased range without the initial and transitional sequence of I-Vs, and vice versa.

2.4.2. Memory phenomenon in nanoCells

Memory is another phenomenon that has been observed in a nanoCell. Based on the forward biased I-V characteristics which indicate NDR characteristics, as shown in Figure 7d, the operating voltage range of the nanoCell can be obviously divided into two regions separated by a threshold voltage V_T. When the applied voltage is confined below V_T, the I-V curve is relatively smooth and follows a predictable track. When the voltage goes beyond V_T, the I-V curve becomes less predictable and includes one or several negative resistance regions. The interesting feature of nanoCell is that the conductance of the first region can be changed by applying a voltage beyond the second region.

As shown in Figure 8, if a voltage sweep with stop value higher than V_T is applied (curve 1 in Figure 8a), the next voltage sweep with stop value lower than V_T yields a low conductance of around $5.5 \times 10^{-8}\,\Omega^{-1}$ (curve 1 in Figure 8b). This low conductance can be switched to high conductance by applying another voltage sweep with stop value higher than V_T (curve 2 in Figure 8a). The resulting conductance is around $1.7 \times 10^{-5}\,\Omega^{-1}$ (curve 2 in Figure 8b). Thus, the process of applying a voltage that is higher than V_T is "write", while the process of applying a voltage that is lower than V_T is "read". If we assign the high conducive state "1" or "on" and low "0" or "off", the nanoCell can be switched between "1" and "0" by applying the writing process.

The assignment of "1" and "0" to different conductive states is arbitrary since the conductance of the read I-V is dependent on the final current value of the writing process. For example, the conductance curves 1 and 2 in Figure 8d are both low if we compare them to curve 1 in Figure 8b. However, it is obvious that they are different since their corresponding conductance are $5.5 \times 10^{-8}\,\Omega^{-1}$ and $1.0 \times 10^{-9}\,\Omega^{-1}$ for curves 1 and 2, respectively. Certainly, in order to insure a high on–off ratio, these conductive states will not be used to differentiate "1" and "0".

Both conductive states are repeatable, i.e., any subsequent read voltage sweep generates similar I-V curve as in the previous one. Besides, switching between the two conductive states is repeatable, i.e., any induced conductive state can be switched to the other one by applying a write voltage. In addition, as shown in Figure 9, the read voltage is reversible, i.e., the I-V characteristics generated from reverse read voltage sweep also follows the same pattern.

Switching could also be carried out by applying a voltage pulse. Similarly, reading can be done by a constant voltage. For example, for the nanoCell that already shows a "0" state, if a voltage pulse of 6 V for 0.05 second is applied, a higher current is obtained at a read voltage of 2 V (Figure 10).

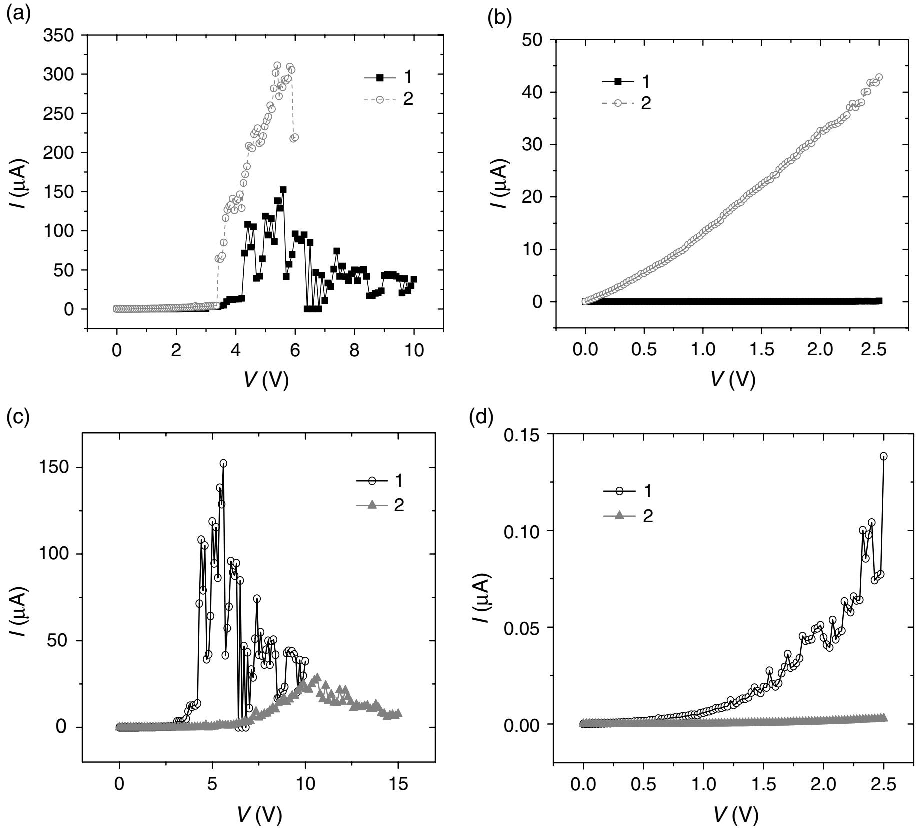


Figure 8 Repeatable memory effect in a nanoCell with $V_T = 3\,V$. (a) Writing process in which the final currents are 38 and 219 μA for curves 1 and 2, respectively; (b) reading process in which the conductance are $5.5 \times 10^{-8}\,\Omega^{-1}$ and $1.7 \times 10^{-5}\,\Omega^{-1}$ for curves 1 and 2, respectively; (c) writing process in which the final currents are 38 and 8 μA for curves 1 and 2, respectively; (d) reading process in which the conductance are $5.5 \times 10^{-8}\,\Omega^{-1}$ and $1.0 \times 10^{-9}\,\Omega^{-1}$ for curve 1 and 2, respectively

One interesting experiment is to prepare and test four different ensembles (Figure 11) by depositing different combinations of molecules **1** and **2**. The first ensemble has only the discontinuous gold film; the second one has molecule **1** self-assembled on the gold islands of the discontinuous gold film; the third ensemble has molecule **2** self-assembled on the silicon oxide of the substrate; and the fourth one has **1** and **2** self-assembled on the gold islands and SiO_2, respectively.

2.5. Influence of molecules on electrical behavior

It is expected that the first and the third ensemble should have extremely low conductance, since alkane and vacuum are both good insulators; the fourth ensemble should have a conductance somewhat in-between. However, the four ensembles showed similar

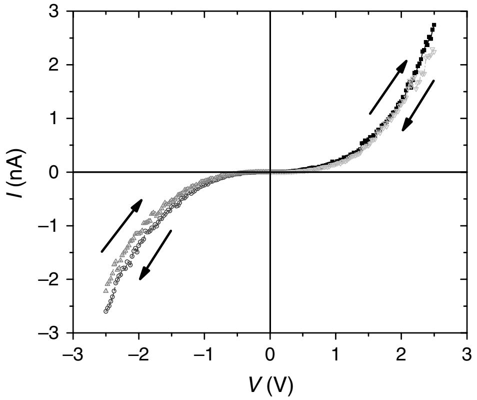


Figure 9 Read conductance in forward and reverse biases. The arrows indicate the direction of applied voltage

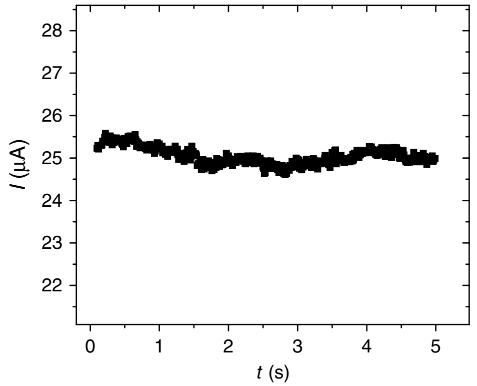


Figure 10 A read current at 2 V vs time after a write voltage of 6 V for 0.05 second

I-V characteristics with peak current values in the same order of magnitude. Thus, we reach the conclusion that the above observed behavior is due to the electron migration through discontinuous gold film.

In order to further investigate the influence of molecules on the nanoCells' electrical behavior, one of the chips with nanoCells featuring NDR is cut into two pieces. Molecule OPE is deposited on one-half of the chip and alkane molecule on the other half of the chip.

The nanoCells with OPE molecule (Figure 11b) exhibit similar switching characteristics like the nanoCells with only the discontinuous gold film (Figure 3a), including a similar threshold voltage and NDR peak value. This further demonstrates that only the formed gold filaments are responsible for the NDR behavior or at least that the contribution of the OPE to the electron conductance is negligible compared to the contribution of the filaments.

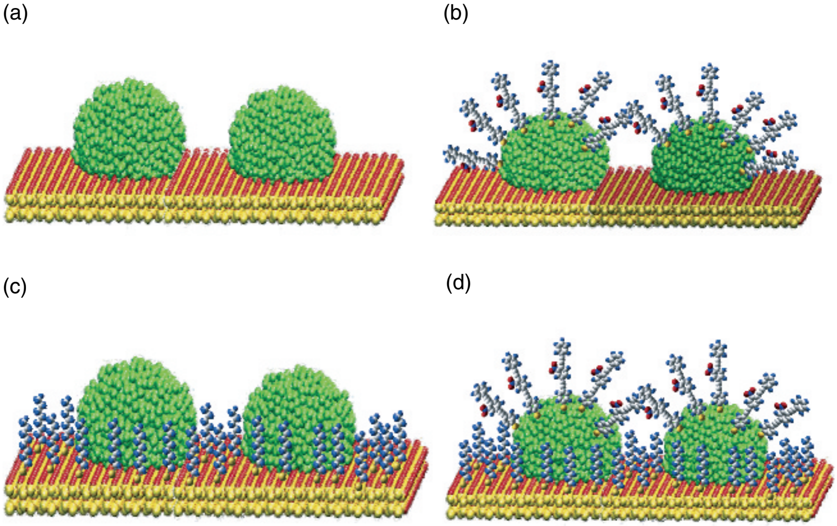


Figure 11 Schematic drawings of four ensembles of the nanoCells with (a) only the discontinuous gold film on SiO_2; (b) molecule **1** deposited on the discontinuous gold film islands; (c) molecule **2** deposited on the silicon oxide surface; (d) molecules **1** and **2** deposited on gold and SiO_2, respectively

The nanoCells with alkane molecule (Figure 11c) self-assembled on SiO_2 show increases in threshold voltage V_{th2} ($\sim$10 V higher). This means that the insulating alkane molecule creates a higher barrier for the electron transfer through gold islands. After reaching V_{th2}, the barrier is overcome and we obtain an NDR characteristic similar to those found without any molecule or with only OPE. We carry out deposition of alkane on the chips already containing molecule OPE (Figure 11d). This fourth ensemble (Figure 3d) exhibits behavior similar to the third ensemble (Figure 11c), which is consistent with our previous conclusion that the I-Vs of the nanoCell with and without OPE are very similar.

To further investigate the influence of the OPE, we deposit the OPE on a chip where no current is found. After deposition, there is still no current. This, again, proves that the OPE has no significant influence on the I-Vs of the nanoCells.

2.6. *Programming of nanoCell*

The nature of randomness on addressing molecules inside the nanoCell makes this device dependent on computer programming.

In a multi-leads nanoCell, each pair of the leads shows similar switching and memory phenomenon as in the two-lead nanoCell. Once a conductive state has been written to one of the lead pairs, the conductance of the same nanoCell between the other pairs changes accordingly. During the measurement, only the two leads of interests are connected to the semiconductor parameter analyzer, while others are dangling. Take the nanoCell shown in Figure 12 as an example. When the conductive state between K and

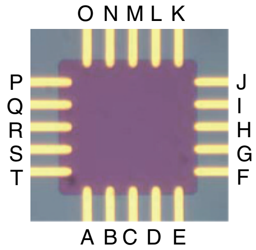
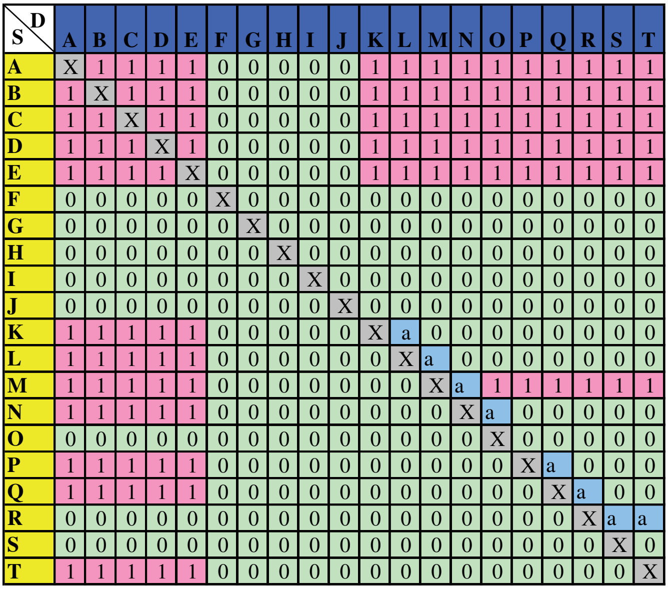


Figure 12 Multiple-leads nanoCell with alphabetic notation on each lead



S \ D	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
A	X	1	1	1	1	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
B	1	X	1	1	1	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
C	1	1	X	1	1	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
D	1	1	1	X	1	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
E	1	1	1	1	X	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
F	0	0	0	0	0	X	0	0	0	0	0	0	0	0	0	0	0	0	0	0
G	0	0	0	0	0	0	X	0	0	0	0	0	0	0	0	0	0	0	0	0
H	0	0	0	0	0	0	0	X	0	0	0	0	0	0	0	0	0	0	0	0
I	0	0	0	0	0	0	0	0	X	0	0	0	0	0	0	0	0	0	0	0
J	0	0	0	0	0	0	0	0	0	X	0	0	0	0	0	0	0	0	0	0
K	1	1	1	1	1	0	0	0	0	0	X	a	0	0	0	0	0	0	0	0
L	1	1	1	1	1	0	0	0	0	0	0	X	a	0	0	0	0	0	0	0
M	1	1	1	1	1	0	0	0	0	0	0	0	X	a	1	1	1	1	1	1
N	1	1	1	1	1	0	0	0	0	0	0	0	0	X	a	0	0	0	0	0
O	0	0	0	0	0	0	0	0	0	0	0	0	0	0	X	0	0	0	0	0
P	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	X	a	0	0	0
Q	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	X	a	0	0
R	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	X	a	a
S	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	X	0
T	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	X

Figure 13 Truth table of a nanoCell as K–E is set to “1”. The truth value in any entry is measured between the drain-terminal (D) and the source-terminal (S) when a voltage sweep is applied to the drain

E is set to “1”, the conductive states between other pairs of leads are summarized in Figure 13.

Three kinds of conductive states exhibit on this nanoCell. The “1” and “0” correspond to the high (higher than $10^{-6}\,\Omega^{-1}$) and low conductance (lower than $10^{-6}\,\Omega^{-1}$), respectively. These are expected to be observed. However, there is a third state with high conductance and ohmic behavior which is similar to the initial transitional behavior in the two-leads nanoCell. In the programming of nanoCell, this kind of conductance can be ignored.

The programming of nanoCell can be carried out using a multiple probe testing board and computer controlled oscilloscope. The fundamental idea is to first collect all the switching information between each leads. For example, the switching information between KE when KE is “1” has been obtained. The switching information between

KE when KE is "0" also needs to be obtained. Similarly, the information between AB, CD, JO, etc. needs to be obtained. Once a database of switching information is formed, the computer calculates the specific pair of leads that a voltage pulse can write into to perform desired function.

3. Electrical conductance of discontinuous metallic film

In this section, first the theoretical models proposed by several researchers to explain the electrical conductance through discontinuous metallic films are outlined. Then the electrical conductance is investigated both experimentally and theoretically in two categories: below the threshold voltage and beyond the threshold voltage.

3.1. Theoretical models in discontinuous metallic film

Since nanoCells with only discontinuous gold film have large current ($>1\,\mu A$) and exhibit memory and switching phenomena, the understanding of the electrical conductance of discontinuous metallic film becomes critical for the analysis of programmable molecular array. It will not only benefit the research work in discontinuous-metal-film-based nanoCell, but also provide in-depth knowledge of electron transport through any two-dimensional molecular array. Since the molecules that will be used in the programmable molecular array are semiconducting, ultimately the programmable molecular array can be viewed as a two-dimensional array of electron transport junctions, as the discontinuous metallic film does.

As early as 1960s, electrical conductance through discontinuous thin metal films has been studied by a number of researchers. Non-ohmic behavior of the electrical conductance have been observed in several types of discontinuous metal film [53–56]. Neugebauer *et al.* proposed that when island sizes and separations between them are both small, activated tunneling of electrons is the dominant mechanism of electron transfer [57]. The barrier to tunneling is the energy difference between a Fermi-level of a particle and the lower edge of the conduction band of the substrate. Hill *et al.* developed a two-island model to account for the transport. However, Uozumi *et al.* argued that the non-ohmic conductance arises from non-ohmicity of the tunneling current at high electric field, and the expression for the current density is given in [58]. They also found that the non-ohmicity occurs when the electric field between the two metal islands exceeds the critical field E_c, which is defined in [58]:

$$E_c = kT/ed \tag{1}$$

where k is the Boltzman constant, T is the temperature, e is the electron charge, and d is the island separation.

Uozumi *et al.* extended the two-island model to one-dimensional series-connected islands with identical sizes but different gap lengths [59]. Their computational result showed that the voltage between two islands can be about a hundred times higher than the average voltage calculated from the one between the two electrodes and that the logarithm of the conductance depends almost linearly on the square root of the voltage between the two electrodes above the threshold voltage [59]. Their experimental results further confirmed their suggestion [60].

A tremendous breakthrough was accomplished by Shin *et al.*, who proposed a ring-shaped model to account for the transport properties [61, 62]. In this model, the ring-shaped array of small islands is located between the two electrodes. Thus, unlike the one-dimensional arrays proposed by others [58–60, 63–68], there are two branches or paths for electron transfer. Electrons may get trapped in one of the islands [69]. Thus, multiple Coulomb blockade gaps may appear in current–voltage characteristics [61] as current peaks followed by NDR behavior. This is the distinct feature of this model in contrast to the linear one-dimensional array. In addition, trapped electrons block the electrical conduction through the array, so no current flows. Using Monte Carlo simulation, Shin *et al.* were able to calculate the current for constant voltage between the two electrodes.

Interestingly, NDR and memory phenomena have been observed and reported in thin insulating films made of SiO [70] and in organic light emitting diodes [71–82]. Tang *et al.* found that for those devices, nanosized metallic islands exist inside the insulating film. Thus, they explained the NDR and memory phenomena based on the ring-shaped model of four islands [83]. The state that the trapped electrons block the conductance path is the “off” state; otherwise it is “high”. The thermal fluctuations results in the transition between “on” and “off”, thus the NDR appears [83]. When the temperature increases, the current peak broadens. In addition, The memory effect is due to the “charging” and “discharging” of the electrons in the system [83].

3.2. Electron transport through discontinuous metallic film below activation energy

As it is mentioned in the last chapter, the current–voltage characteristics of nanoCell can be divided into two regions: below V_T, the current increases with increasing voltage and the I-V curve follows a predicted behavior; beyond V_T, the I-V curve changes violently and include one or several local current maximum.

When the applied voltage is below V_T, the I-V characteristic is relatively smooth for both low conductance and high conductance. In order to find out the electron conduction mechanism, we performed a temperature variant current–voltage measurement on the nanoCell sample that carries room temperature I-V curves shown in Figure 14. There are no molecules deposited on the discontinuous film. Thus, the gaps between the gold islands are considered as vacuum.

Interestingly, although discontinuous gold film is a two-dimensional array of junctions, the electrical characteristics resembles the one that is predicted by Uozumi *et al.* Figure 15 shows the electric field dependence of the high conductance G of a discontinuous gold film at various temperatures. The electric field is calculated directly from the applied voltage by dividing by 2×10^{-4} cm, which is the distance between the two electrodes for this particular sample. The interesting feature of these curves is that a remarkable ohmic conductance can be observed beyond 220 K at low electric fields, i.e., around 20 $(\mathrm{V/cm})^{-1/2}$.

Thus the critical applied field strength from 220 K to room temperature is 400 V/cm. A second interesting point is for all temperatures, there is another critical applied field

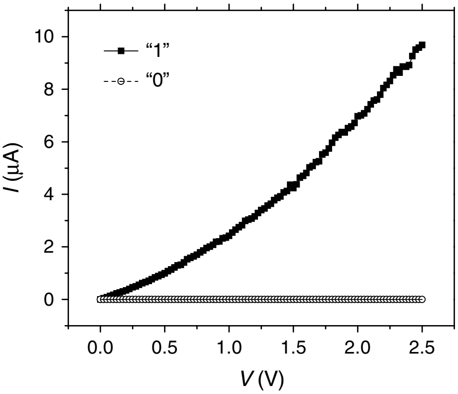


Figure 14 Room temperature I-V curves of the sample whose field dependences of conductance are shown in Figures 15 and 16

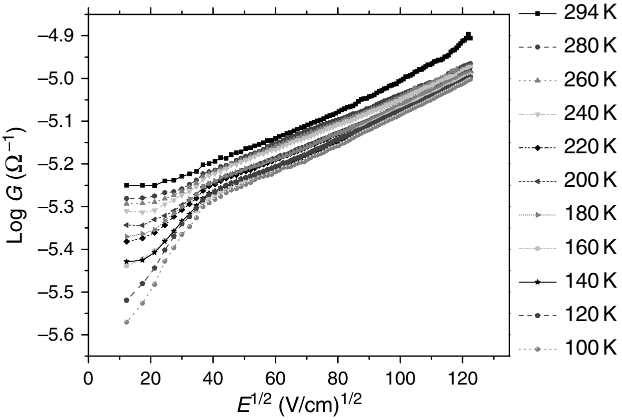


Figure 15 Electrical field dependence of the high conductance of a discontinuous film

above which the logarithm of the conductance G depends almost linearly on the electric field in a form of

$$\log G = NE^{1/2} + C \tag{2}$$

where N is the slope and C is the interception.

In contrast, the electric field dependence of the low conductance G at various temperatures is shown in Figure 16. Since the current is close to zero, the measurement will inevitably generate some artifact current value that is lower than zero. It should be noted that although the current is extremely small, there is a slightly larger variation of the current in low field. Since the largest variation happens in 260 K and 220 K, it seems that this variation does not depend on temperature.

While several mechanisms have been proposed to explain the NDR behavior, modern molecular simulation technique enables us to visualize the atomic movement in an

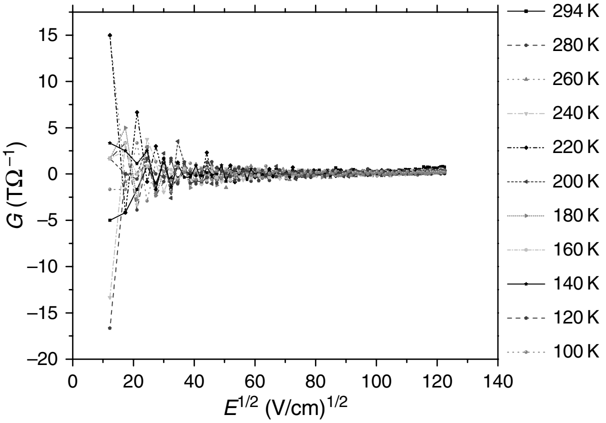


Figure 16 Electrical field dependence of the low conductance of a discontinuous film

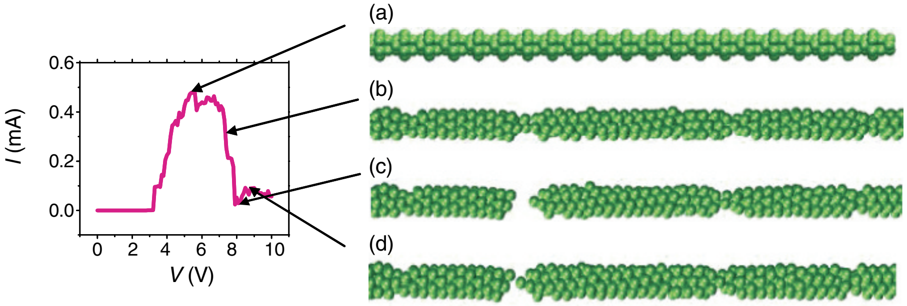


Figure 17 Comparison of molecular simulation results and experimentally obtained I-V [50]

electrical field. Molecular dynamics (MD) simulations have been performed on a one-dimensional nanowire with hundreds of gold atoms. The simulation starts from an ideal state where all the atoms are closely packed and the whole nanowire is in a regular shape (Figure 17a). Since this shape has relatively high conductance, we could say the peak current is associated with it.

3.3. NDR region beyond threshold voltage

When an electric field is acting upon the nanowire, the local temperature of the nanowire increases, the atoms vibrate persistently, yielding the irregular shape shown in Figure 17b, which also accounts for the decreasing current. The atoms continue to vibrate, and the shape of the nanowire is constantly changing, which results in the rising and falling of the current. At some point the atoms vibrate so quickly that the nanowire breaks, yielding nano-clusters, then current rapidly drops to zero (Figure 17c). After the nanowire breaks, the local temperature decreases. Then under the high voltage electron migration induced the moving of the gold atoms, thus bridging the nanoclusters

(Figure 17c). Hence the current increases again. This could also explain the rising of the current when the biased voltage is between V_T and V_{max}. Thus, by the assistance of MD simulation, the I-V characteristic beyond V_T can be explained and is mainly due to the combination effect of thermal motion and electro migration.

3.4. Effects of morphology of discontinuous gold film

The gold islands in the nanoCells are removed when a chip with discontinuous gold film is immersed in piranha solution (a mixture of H_2SO_4 and H_2O_2, in a ratio of 3:1) for 30 minutes. The temperature in this exothermic reaction rises to 130 °C. However, we find that only the gold filaments and clusters can be removed without significantly damaging the discontinuous Au film if the chip is immersed for a much shorter time.

An experiment is carried out in which one chip with only the discontinuous gold film is immersed in piranha solution for 45 seconds. Thus, the gaps between the gold islands become wider after the piranha bath. On this chip, prior to the piranha bath treatment, some of the nanoCells have yielded NDR behavior. After the piranha bath, we find no NDR behavior in these nanoCells. Instead, the two types of observed behaviors (first and second) are found again; therefore, the nanoCell can be reset to its original conduction state through the bathing process. Moreover, the nanoCell does not have memory of its previous behavior; in other words, a nanoCell, which initially exhibits the first observed behavior, may exhibit the second observed behavior after the piranha bathing, (i.e., second behavior, piranha bath, and first behavior) and vice versa. Then, the same nanoCell chip is immersed in piranha for another 2 minutes and then 90% of the nanoCells do not show any current in excess of the noise (~0.1 pA) even when they are biased up to 100 V and observations under an optical microscope indicate that the discontinuous gold film (gold islands) is not removed.

4. Static and transient current–voltage characteristics at the nanoscale

In this section, the standard transient response of conventional systems is first briefly reviewed. Then, the similarities and differences between the nanosized systems and the conventional system are introduced. A situation is proposed in which the number of electrons may be in the order of magnitude as the number of nuclei. The effect of this situation is investigated based on the static current–voltage measurement of discontinuous gold film.

4.1. Standard transient response of conventional systems

For many physical systems or devices, there is a transient status between two stationary stages. The nature of the transient and stationary responses is determined by the individual characteristics of the system or device. What is most interesting is that most realistic systems have uniquely determined responses by the same equations, implying that both static and transient responses can be studied using similar techniques through several fields of science and engineering. In particular, systems with equivalent characteristics constitute the topic of several if not all of present engineering.

For the electrical RLC circuit, the rising and falling of the voltage is determined by the movement and rearrangement of electrons in the RLC circuit, thus time scales associated are usually from fractions of seconds to fractions of nanoseconds; however, the chemical plant has a transient behavior depending on the movement and arrangement of big, macroscopic masses involving large portions of fluids and large mechanical systems with transient times from minutes to several hours.

The interesting analogy between chemical and electrical systems of totally different size, no matter human-fabricated or natural, has usually been taken for granted and we never ask the question “why?” Why this strong similarity between a macroscopic and microscopic system is not observed at the nanoscopic scale? The answer is based on the atomistic nature of the system. At microscopic level we can observe the robustness of the nuclei compared to the fluid, the electrons. For instance, a typical current of 10^{-4} A in an n-channel silicon-based MOSFET with channel length $L = 1\,\mu\text{m}$, width $W = 10\,\mu\text{m}$ and depth $d = 100\,\text{nm}$ represents a total of 6.25×10^5 electrons flowing through the channel in one nanosecond, which is the typical frequency ($\sim$GHz) of today’s microelectronic device. We know the density of silicon is 5×10^{22} atoms per cm^3 [84]. Thus, in such a volume holding the little transistor, one single electron causes a small perturbation to an average of 10^5 atoms. As a result, the nuclei in a crystal are not strongly affected by the dynamics of the electrons. Chemically speaking, when the nuclei are kept together by equally strong chemical bonds in the three dimensions, the strong dependence of transient and stationary responses is valid. This also gives rise to specific and sharply defined transient and stationary responses. A similar situation takes place at macroscopic level, for instance, in the chemical plant where the fluid is not a small perturbation to the materials forming the plant. Since the weight of the fluids in chemical plant is in the same order of magnitude as the materials making the plant, the nature of the flows may cause much more fluctuations in plants than electronics, which also explains why chemical plants require much more maintenance than the transistors in integrated circuits.

As the electronic device scales down dimensions whereby the number of electrons and the number of atoms become comparable, then a similar problem encountered in a chemical plant also is faced by the electrical circuits. The plant components should be robust enough to hold the fluids; however, in microelectronics, the materials under process are not the electrons but information encoded on their flow. In electronics, the amount of information that can be processed per unit of matter has grown exponentially; however, in classical engineering, the amount of material that can be processed by a unit of plant material has undergone only a linear growth and it is practically constant with time. Thus the exponential growth of the former is going to strongly affect the electronic devices at nano-dimensions if we want to continue with such an exponential growth.

Thus, as devices and systems approach the nanoscale driven by the present trends in nanotechnology [85], the strong relationship and sharp separation between transient and stationary responses is broken and blurred as the electron–nuclei interaction takes a major contribution. Thus from a system point of view, transient times in nanosystems are longer and may differ from one similar system to another as construction and distribution of atoms at the atomistic level need to be considered individually and not averaged over a large number of nuclei.

4.2. Set-up of transient response measurement

To test the above assumptions we have performed several electrical experiments of non-chemically robust metallic structures. The structure we test is a piece of discontinuous gold film deposited on a silicon dioxide substrate. Scanning electron microscope (SEM) image shows that the gap between the gold islands is around 5 nm. There are two probe tips made of palladium and titanium at each side of the film to allow the application of electric fields. The separation between the tips is 3 μm.

For the static current–voltage measurements, we combine the time-domain measurement function of the HP4145A and our custom designed Labview control program. In the time-domain measurement, a fixed voltage is applied while the current is measured in every interval $\Delta t'$. Once the current is considered stable or stationary, the voltage goes down to zero and then the next voltage is applied.

4.3. Atomic scale response of discontinuous gold films

Before the static voltage sweep is applied, we performed a transient current–voltage measurement.

The voltage sweeps from 0 to 5 V with a step size of 0.05 V. Figure 18 shows a transient current–voltage characteristic obtained from the device constructed by discontinuous gold film. The delay time is 0.1 second. In the low voltage region, the current voltage curve is slightly smoother than the high voltage region.

The static voltage sweeps from 0 to 5 V with a step size of 0.1 V. The current is recorded in intervals of 10 seconds. Figure 19 shows the measured current vs time when applied voltage is varied from 0 to 5 V and then back to 0 V. It is clear from the plot that for a fixed voltage, the current oscillates around a relatively stable value.

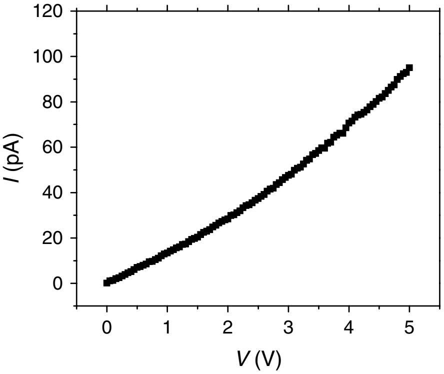


Figure 18 Transient current–voltage characteristic of a discontinuous gold film device

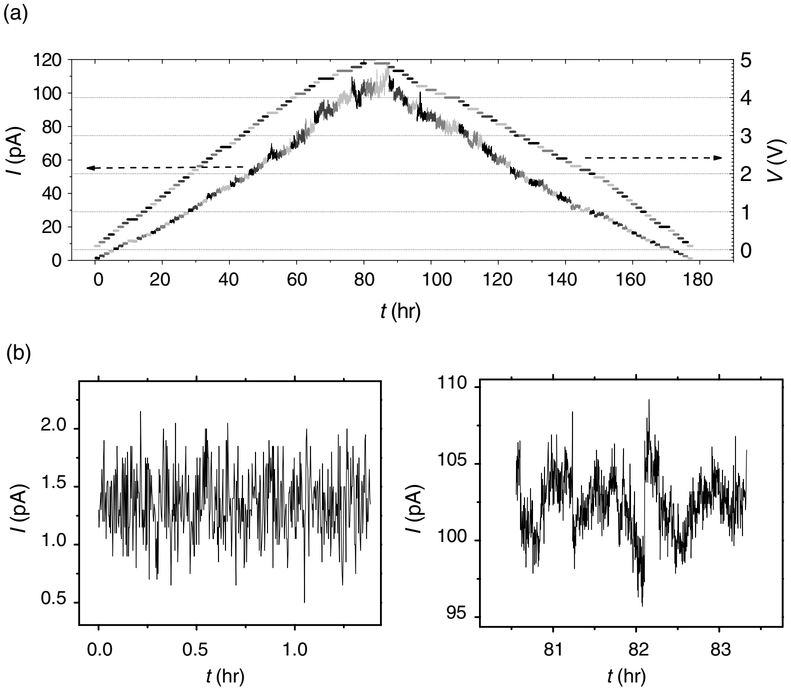


Figure 19 Time-domain current (axis on left side) vs applied voltage (axis on the right) from 0 to 5 V and then back to 0 V. (b) Current vs time plot when applied voltage V = 0.1 V; (c) current vs time plot when V = 5.0 V

This current oscillation is the "transient" behavior of the device, analog to macroscopic electrical devices, for example a resistor, which yields a fixed current when a fixed voltage is applied. Consider now the discontinuous gold film as the structure under test, which is a two-dimensional array of gold clusters with nanosized separations. Since there is no other substance inside the system, the conduction of current should depend on the conformations of the gold atoms. As the voltage is applied, the gold atoms diffuse by electron migration and form quantum nanosized filaments [30, 86, 87] (which are different in nature from widely studied classical microfilaments). As the current flows through the nanofilaments, the local temperature inside the filaments also increases, which results in the breaking of nanofilaments. As a result, the current increases by forming the nanofilaments and decreases by breaking the nanofilaments forming a variety of nanocluster structures. We find that the higher the applied voltage, the more the current oscillates. This phenomenon is also clearly shown by observing the standard deviation of the current vs. voltage plot, as in Figure 20. Since the local temperature increases with the applied voltage, and the self-diffusion coefficient of the gold clusters increases accordingly [30], the gold atoms vibrate more vigorously. This explains the strong oscillation of current at high voltages.

In addition, the level of uncertainty of the current also increases as the voltage increases. For instance, when applied voltage is 4.8 V, the current ranges from 80 to 100 pA. At the next step, i.e., V = 4.9 V, the current ranges from 97 to 108 pA. This

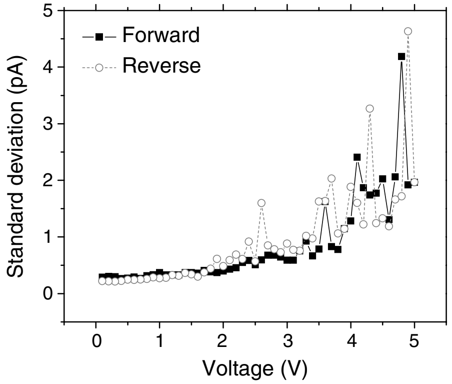


Figure 20 Standard deviation of the current flowing through the nanoCell at each applied voltage

uncertainty have a strong impact on the transient current–voltage measurements yielding results in the choppy region of current–voltage curve in the high voltage area, as shown in Figure 18.

On the other hand, the current uncertainty also washes out or creates artificial NDR characteristic. If, during a measurement, the current is 100 and 98 pA at 4.8 and 4.9 V, respectively, an NDR behavior is observed. However, if the current is 95 pA at 4.8 V and 100 pA at 4.9 V, then the NDR is not observed. In fact, NDR has been reported in several experimental papers [30, 47, 48] where only non-"static" current–voltage measurements were performed.

In addition, it takes longer time for the device to reach a stationary state at higher voltage. As seen from Figure 19(b), when the applied voltage is 0.1 V, the system immediately enters into a quasi-static stage, i.e., although the current oscillates, the mean value of the current remains almost constant. However, at higher voltage, $V = 5.0\,\mathrm{V}$ in Figure 19(c) for instance, the current oscillates for more than 4 hours and still does not reach to a static or quasi-static state.

4.4. Time-dependent NDR and hysterisis

A large amount of work on single molecules has been reported indicating switching and NDR; it would be informative to re-analyze those results to find out whether they are really electrical in nature as it takes place in standard microelectronics or they are due to nuclear dynamic effects due to the strong interaction with the electron currents as the experiments reported in this work.

We actually provide a new perspective to study the electrical characteristic of nano systems at the atomistic level. As the dimensions of today's electronic devices scale down to the nano level, the relative amount of electrons flowing through the electronic system increases when compared to the microelectronic systems. This relatively increasing amount may result in the displacements of atoms, which in turn causes a more complex transient behavior due to the creation of conformational states of gold clusters. The design of devices needs to consider this new conformational behavior and might utilize these inherent characteristics for some specific functions of nano circuits.

5. Vibronics and molecular potential

In this section, first the limitation of charge–current transport as a method for signal transmission is reviewed. Then the concept of vibronics is introduced and related research works are reviewed. Next the fundamental theory of MD simulation is explained. And then the DSP techniques for encoding and decoding information are introduced. The results of the simulation are reported. Additionally, the molecular electrostatic potential method is explained and simulation results are provided. Finally, the feasibility of combining the two methods into programmable molecular array is investigated.

5.1. Limitations of charge–current transport as a method for signal transmission

One of the biggest problems faced in microelectronics development is the amount of power consumed by the various components in the circuit. This leads to extremely high amount of heat dissipated by the chips containing the integrated circuits. Although smaller transistors consume less power, as transistor density and speed rise, the overall chip consumes more power and generates more heat.

Presently, a charge–current approach is widely used to process and transfer information in any integrated circuit. In this approach, the data is stored by induced charge separation in a memory device and data signals are transmitted and processed using electron current variations. Based on these methods, detection of signals and driving of circuits require minimum signal-to-noise ratio having strong effects in energy dissipation, which is the major limiting factor for charge–current approaches. Nevertheless, for practical but not optimal reasons, this approach is still being used in molecular electronics [16].

The programmable molecular circuit, which is studied in the previous two chapters, is proposed to complement conventional electronic devices in order to reach molecular size devices. However, it may not be the right solution if it still uses the charge–current approach, i.e., the data is transmitted through the device using electron current variations, which may still, and certainly will, lead to the problem of energy dissipation.

5.2. Previous research work related to vibronics

Due to the limitations of charge transport, molecular vibronics combined with MEP is proposed as a substitute method to transport information [88].

“Vibronics” stands for “vibrational electronics”. When a signal is injected into a molecule, the vibrational modes of the atoms around the injection point change. Since the atoms in a molecule are bonded to each other, the changes in the vibrational states of the molecule triggers the movement of atoms, which transfers to their neighbor atoms by means of bond bending, bond stretching, Van der Waal interactions, coulombic interaction between charges, etc. [89]. Thus using vibronics, the signals are transferred through the molecule [16, 27, 90].

Our MD simulation demonstrates that a signal can be encoded and can modulate the vibrational state at one site of the molecule, and the propagation of the vibrational movement allows the vibrational signal be detected at the other site of the molecule. Thus the molecule can be used as a single signal processing unit. Digital signal processing techniques are used to encode and decode the signals from molecular vibrational states. The typical molecular vibrations are in the range of terahertz (THz) thus starting a new era for terahertz signal transmission techniques in molecular electronics. All this work is triggered by preliminary MD simulations performed in a "nanoCell" where the modulation of a signal was performed [31, 91]. It should be noted that to implement vibronics in real electrical circuits, transducers are necessary to convert the electrical signals to vibrational signals and vibrational to electrical.

5.3. Molecular dynamics simulation

Atoms in molecules vibrate around their equilibrium position. Usually, the vibrational modes are assigned to certain bond length stretching or bond angle bending, etc.; however, bond length stretching or bond angle bending can cause the vibrational movement of other atoms that are directly or indirectly bonded to this vibrating bond. Therefore a vibrational mode is actually the vibrational movement of a few atoms in a molecule. Although the propagation of vibrational movements are beyond the ability of direct observation by modern experimental techniques, modern computer simulation techniques allow us to trace the trajectory of all the atoms, and therefore, to view the vibrational movements of individual atoms in a molecule.

Molecules are modeled as a collection of mass centers, which may be a single atom or a group of atoms (united atom) that are part of the molecule. The mass centers, which will generally be called atoms in this context, bear electric charges. The atoms interact with each other via bonding forces, non-bonding forces, and electrostatic forces.

During the simulation, two atoms, F_1 and F_2, from each end of the molecules are fixed. The vibrational signal is injected into the molecule following every step of updating coordinates. In this extra step, an atom X bonded to one of the fixed atoms, F_1 for example, oscillates along the direction X-F_1 at the distance defined by the modulated signal. The step time is set to 1 fs, and 200,000 equilibration steps are run before the production runs to relax the system. Then, 1,000,000 production steps are run corresponding to 1 ns, and the coordinates of all the atoms at each step are recorded in a trajectory file.

5.4. DSP techniques for encoding and decoding signals

From the trajectory file, time series of bond lengths, which contain the vibrational information, is calculated at each time step, and digital signal processing (DSP) techniques are applied to analyze the signal transmission along the molecule.

5.4.1. Modulation techniques

Before the signal is transmitted in a molecule, it has to be modulated with a carrier signal. This carrier signal is usually a sinusoidal signal. The main reason for modulation is that it makes the signal properties physically compatible with the propagation medium, i.e., the molecule.

Amplitude modulation (AM) and frequency modulation (FM) are two methods used frequently in signal modulation. In AM, amplitude of a carrier wave is varied in direct proportion to that of a modulating signal; while in FM, frequency of the carrier is varied. In our simulation, the frequency of the carrier wave is selected as the intrinsic vibrational mode of the backbone of the molecule. For example, for polypeptide, the frequency f_c is 23.81 THz.

For the sake of identifying, the modulating signal, i.e., the actual information signal) is designed to consist of a series of squares and triangles.

The frequency modulating (FM) wave, $x_F(t_i)$, is slightly different from the AM signal that it consists a series of trapezoids and triangles.

5.4.2. Decoding information

Digital signal processing (DSP) techniques are used to analyze bond length vibration signals in the molecular wires [92, 93]. At a sampling period of 1.0 fs, which is identical to the time step for the MD simulations, a discrete time series $D_{\alpha\beta}(i)\,[i = 0,\ 1, \ldots,\ N-1]$ is calculated from $\mathbf{R}_\alpha$ and $\mathbf{R}_\beta$, the i-th step coordinates of atoms α and β, respectively, recorded in trajectory file.

The time series $D_{\alpha\beta}$ contains both the direct (DC) and alternating (AC) components. The AC series, $\tilde{D}_{\alpha\beta}(i)\,[i = 0, 1, \ldots, N-1]$, is calculated from the initial bond length series $D_{\alpha\beta}(i)\,[i = 0, 1, \ldots, N-1]$ by subtracting its series average $D_{\alpha\beta,0}$, which is the DC component.

The frequency spectrum of the time-domain signal $F_{\alpha\beta}(k)\,[k = 0, 1, \ldots, N-1]$ is obtained by a fast Fourier transform (FFT) of the AC time series $\tilde{D}_{\alpha\beta}(i)\,[i = 0, 1, \ldots, N-1]$ [92, 93]:

$$F_{\alpha\beta}(k)\,[k = 0, 1, \ldots, N-1] = \sum_{i=0}^{N-1} \tilde{D}_{\alpha\beta}(i)\, e^{-j2\pi i k/N} \tag{3}$$

$$F_{\alpha\beta}(k) = A_{\alpha\beta}(k)\, e^{j\varphi_{\alpha\beta}(k)} \tag{4}$$

where N is the total number of samples in the series, $A_{\alpha\beta}(k)$ and $\varphi_{\alpha\beta}(k)$ are the amplitude and phase of the frequency-domain signal, respectively.

5.4.3. Recovering amplitude-modulated signal

To recover signal from an AM wave, the time-domain AC series is filtered using a Bessel bandpass filter centered at a carrier frequency and with a bandwidth of around 1 THz. The Bessel filter is selected because its output closely resembles the original waveform adding small and gentle edge rounding [92]. This bandpass filter filters out all the low frequency and high frequency vibration signals and only allows vibrational signal centered at the carrier frequency f_c with a narrow band to pass. Then, the bandpass filtered signal is rectified using a rectifier and the signal is recovered by passing the rectified signal through a Bessel lowpass filter with a small cutoff frequency; these last two steps can also be done using a peak detector. Figure 21 shows the front panel of the Labview program for amplitude modulated signal recovering.

As any computer program, the Labview program consists of input section and output sections. The user needs specify the trajectory data, i.e., the time series of bond lengths;

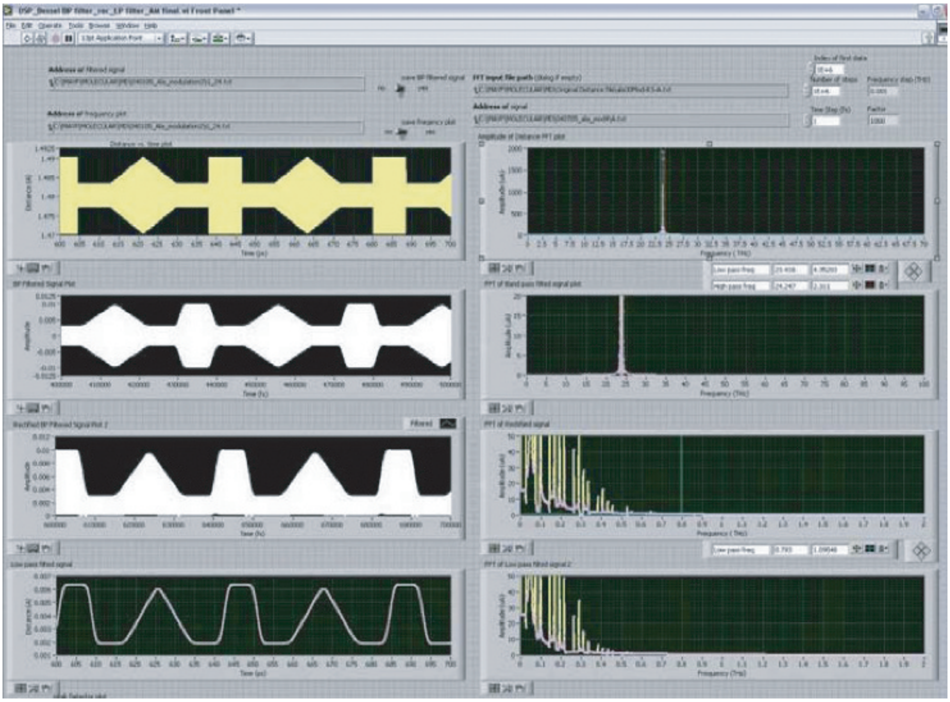


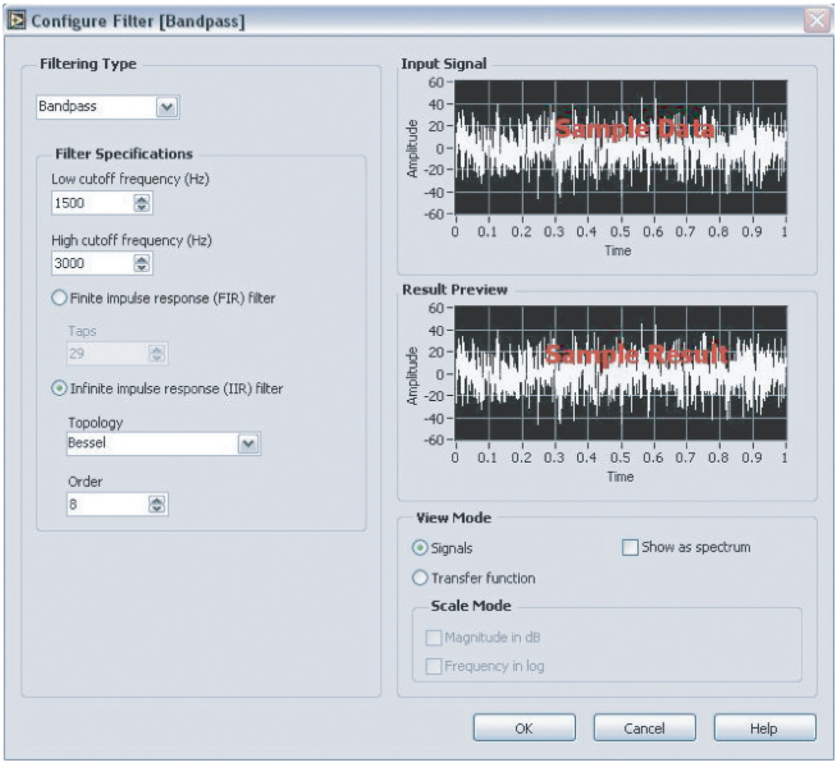
Figure 21 Front panel of Labview program to recover AM signal. The input fields are located on the top of the panel; the output time-domain signals (lower left) and their corresponding frequency-domain signals (lower right) are shown

the sampling interval, i.e., the time steps that has been used to record the trajectory data; the cut-off frequencies for the bandpass filter and lowpass filter. Because the fast Fourier transform of the trajectory data is needed for the selection of cutoff frequencies, the last two inputs are specified when the program is running, instead of before the program starts. The output signals are visualized in the front panel as well.

The program is constructed using a sequence structure in which each diagram, or frame, executes sequentially. This structure is chosen because of its ease of data manipulation. In the first diagram which primarily prepares the signal data for analysis, the program reads the data from the selected file and transforms it into a one-dimensional array. It is preferable to skip the initial part of the signal since it may contain excessive noise due to large vibration. The DC series of the data, i.e., the mean value, is extracted from the data so that only the AC signal is remained. Frequency spacing Δf is also calculated in this diagram:

$$\Delta f = \frac{1}{N \cdot \Delta t} \tag{5}$$

where Δf is the frequency resolution, N is the number of data and Δt is the sampling interval. The input data is plot in a waveform graph shown as the first one in the left column of Figure 21. To perform the bandpass filtering of the input AC data, the user needs to manually select the cutoff frequencies of the bandpass filter by



Configure Filter [Bandpass]

Filtering Type

Bandpass

Filter Specifications

Low cutoff frequency (Hz): 1500

High cutoff frequency (Hz): 3000

○ Finite impulse response (FIR) filter

Taps: 29

● Infinite impulse response (IIR) filter

Topology: Bessel

Order: 8

View Mode

● Signals

☐ Show as spectrum

○ Transfer function

Scale Mode

☐ Magnitude in dB

☐ Frequency in log

Figure 22 Front panel of the bandpass sub-program

dragging the vertical cursors in the interactive graph of the frequency-domain signal (shown as the first one in the right column of Figure 21). The values of the cutoff frequencies are inputted into the Bandpass sub-program, whose front panel is shown in Figure 22.

The infinite impulse response (IIR) Bessel filter is the right choice for bandpass filter since it has maximally flat response in both magnitude and phase and nearly linear-phase response in the passband [94]. The rectification is realized by simply using a Labview sub-program which returns the absolute value of each data point. The lowpass filtering of the rectified signal is similar to the bandpass filter in the previous step. The topology of the lowpass filter is also the Bessel IIR filter. The difference is that the user only needs to select the high frequency limit for the filter.

The lowpass filtering can be replaced by a peak detector which finds the location and amplitude of the input signal. The resulting signal is similar to the lowpass filtered signal.

5.4.4. Recovering frequency-modulated signal

The demodulation of the FM signal is carried out by means of slope detector. First, the time-domain AC series, which is obtained as described in the AM signal recovery, is filtered using a Bessel bandpass filter centered at carrier frequency. Then the bandpass filtered signal is clipped to constraint the amplitude variation and thus eliminating the thermal noise due to the finite temperature. The clipped signal is demodulated by a slope demodulator, a special Bessel bandpass filter, which has a center frequency slightly larger than the carrier frequency, and the entire frequency variation of the FM signal falls on the linear part of the left slope of the frequency response curve of the filter [92]. Figure 23 shows the frequency response of the slope detector used to demodulate signal transmitted through molecule GLY58. Since the carrier frequency is 23.81 THz and the frequency variation of the bandpass filtered signal is ±0.2 THz, the center and bandwidth of the slope demodulator are set to 24.9 and 2.6 THz, respectively. Finally, the signal is recovered by passing the slope-detected wave through a full wave rectifier and a lowpass filter.

The Labview program to recover the FM signal is essentially similar to the amplitude modulated one with two additional steps. The first one is a clipper after the bandpass filter to eliminate excessive noise. The clipper is realized by a sub-program, which clips the elements to within the boundary defined by upper and lower limits. The limits are set by the user in interactive waveform graph. If x_i represents the original data, a and b represents the upper and lower limits, respectively, the clipped data y_i is

$$y_i = \begin{cases} a & x_i > a \\ x_i & b < x_i < a \\ b & x_i < b \end{cases} \tag{6}$$

After the clipper, there is a demodulator to demodulate the frequency-modulated signal. The demodulator is basically a Bessel bandpass filter. The difference between this filter and the filters in the previous section is that it has different center and bandwidth. Frequency response of the filter is necessary to find the correct bandwidth and center. Figure 24 shows the Labview program for obtaining frequency response. By varying the

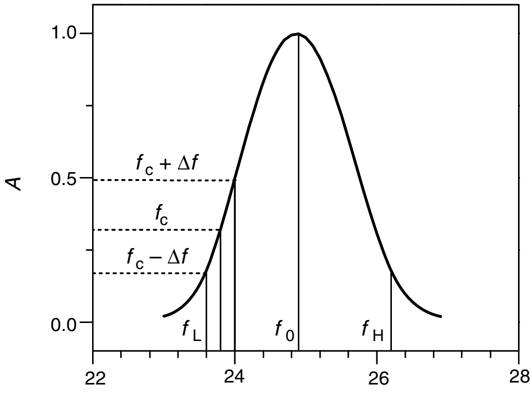
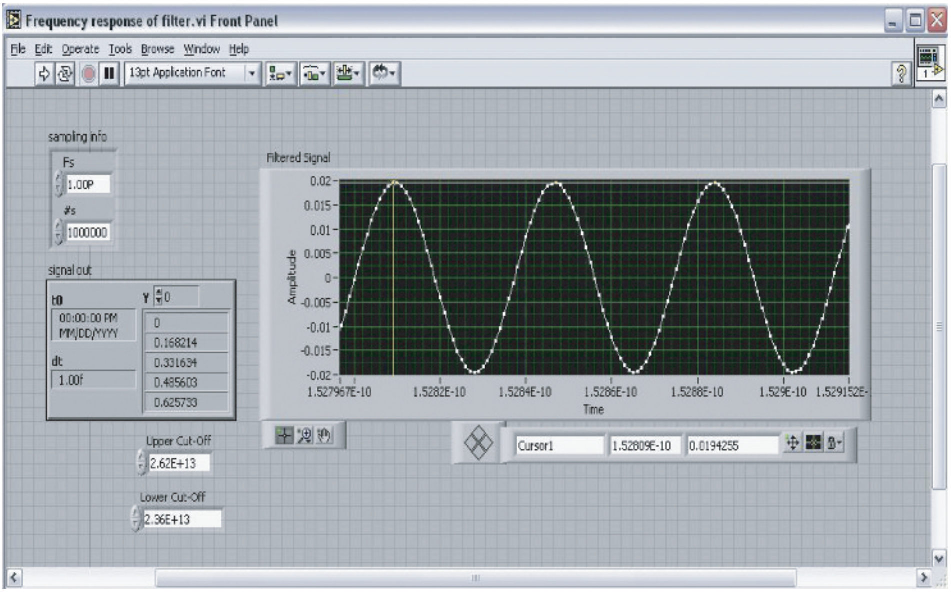


Figure 23 Frequency response of the slope demodulator, a special Bessel bandpass filter, used for FM signal demodulation; the center frequency f_0 is 24.9 THz and the bandwidth is 2.6 THz. The lower and upper cutoff frequencies of the slope demodulator, f_L and f_H, are 23.6 and 26.2 THz, respectively. The carrier signal has a center frequency f_c of 23.81 THz and frequency variation Δf of 0.2 THz



Fs 1.00P

#s 1000000

t0 00:00:00 PM MM/DD/YYYY

dt 1.00f

Y 0

0
0.168214
0.331634
0.485603
0.625733

Upper Cut-Off 2.62E+13

Lower Cut-Off 2.36E+13

Cursor1 1.52809E-10 0.0194255

Figure 24 Front panel of Labview program for obtaining frequency response of a bandpass filter

frequency of the input sinusoidal signal, the amplitude of the filtered signal is recorded and plotted versus the input frequency.

5.5. Simulation results of molecular vibrational signal transmission

Molecular dynamics simulation of molecular vibrational signal transmission is carried out for molecule GLY58 (Figure 25) at room temperature (298.1 K) for 1 ns. The sampling time, i.e., the time step is 1 fs and total number of steps is 1,000,000. The total length of the molecule is around 218 Å. The intrinsic vibrational mode of the backbone of this molecule is 23.81 THz.

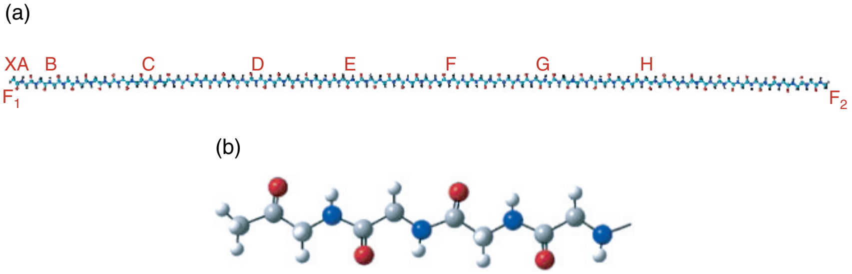


Figure 25 (a) Molecular structure of GLY58; (b) detailed structure of GLY58 at one end of the molecule. The grey atoms are Carbon (C); the blue Nitrogen (N); the red Oxygen (O); the white Hydrogen (H) [95]

During MD simulation, two carbon atoms, F_1 and F_2, from each cap are held fixed. Input signals are coupled to the carbon atom X that is bonded to F1 at left. The signals are detected at different sites A, B, C, D, E, F, G and H in the backbone of GLY58 as shown in Figure 25. Sites A, B, C, D, E, F, G and H are located at 1.4, 8.3, 34, 67, 90, 116, 142, 168 Å from the input site, respectively.

5.5.1. Amplitude-modulated signal transmission

Amplitude-modulated signal is injected into the molecule GLY58 by coupling the input signal into the movement of carbon atom X. Figures 26a and b show the time-domain and frequency-domain signals detected from site A (1.4 Å from the input site X in Figure 25a) and site E (90 Å from the input site X), respectively. In the frequency-domain signal, the carrier frequency is clearly shown as a peak in 23.81 THz. Apparently, the frequency-domain signal carries not only the carriers signal, shown as a peak at 23.81 THz in Figure 26d, but also other molecular vibrational signal due to thermal noise, solvent effect, etc. The modulating signal is recovered as described in the previous section. The Bessel bandpass filter is centered at 23.81 THz with a bandwidth of 0.7 THz. The Bessel lowpass filter has the cutoff frequency of 0.5 THz. The recovered signal is shown in Figure 26c. The similarity between the original modulating signal and the recovered signal clearly tells us that the recovered signal is the original modulating signal that propagates from the site X.

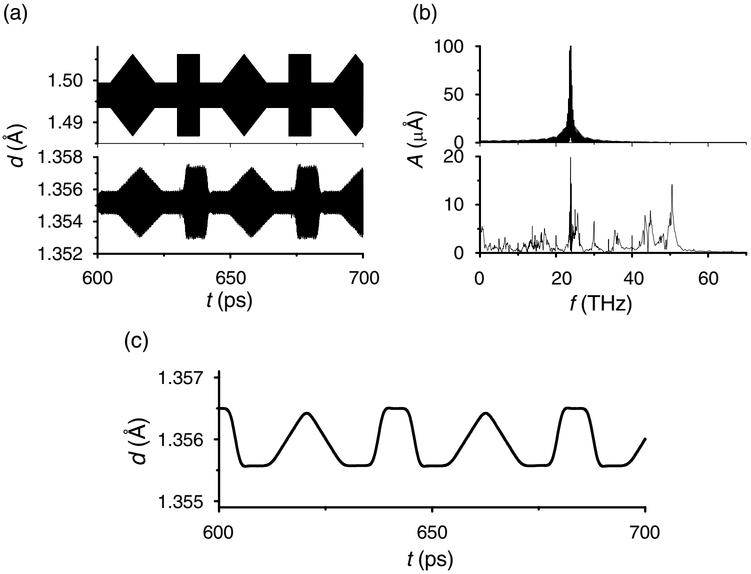


Figure 26 Signal transmission along GLY58 using amplitude modulation by a carrier at 23.81 THz. (a) Time-domain vibrational signal detected at a remote site A (upper plot) and E (lower plot); (b) frequency-domain vibrational signal at site A (upper) and E (lower); (c) signal recovered using DSP techniques [95]

5.5.2. Time delay of molecular vibrational signal transmission using AM

Figure 27 shows the recovered AM AC signals at different sites in molecule GLY58 as indicated in Figure 25a. The average velocity of the vibronic signal can be determined following the time delays as signal propagates along the molecule. From Figure 27, the peak of the first triangle-shaped signal is located at 617.705 ps at site A and 622.775 ps at site H. The distance between site A and H is 166.6 Å. This yields a speed of 3286 m/s. When this type of information is collected at all sites, an average speed of 3279 m/s will be obtained (Figure 28). Although this speed is at least one order of magnitude smaller than the speed of electrons in a bulk semiconductor, the advantage of our vibronic approach is that the energy needed to transfer one bit of information is a fraction of eV as compared to the few ten-thousands of eV needed under the present electron current approaches.

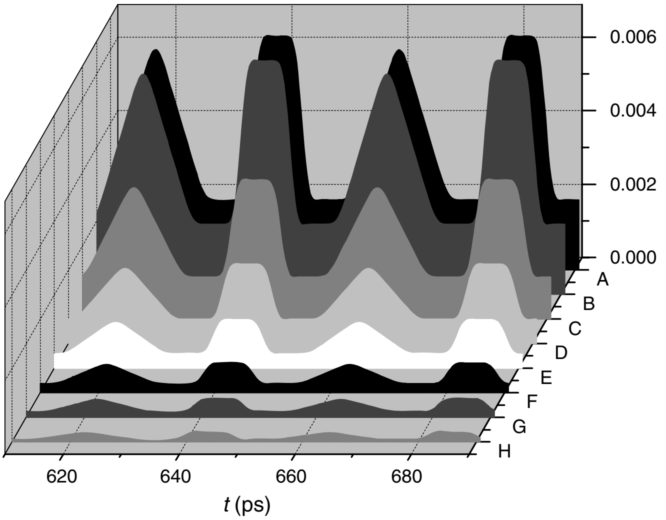


Figure 27 Detected AM AC signals in molecule GLY58 at different sites as indicated in Figure 25a [95]

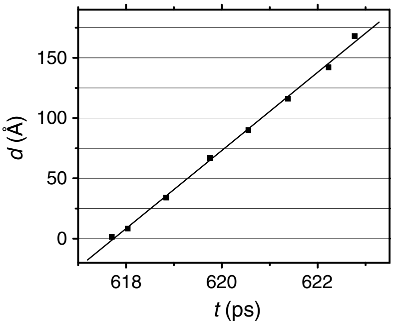


Figure 28 Distance vs time-delays of the first triangle apex from AM signals in Figure 27 [95]

5.5.3. *Noise and attenuation of molecular vibrational signal transmission using AM*

Signal-to-noise ratio is a determining factor in any data processing systems. As signal travels, its amplitude decreases to the point that is washed out by the noise (usually thermal noise). Amplitude modulation is prone to this as noise simply adds to the signal. The vibronic approach is suited to interconnect molecular and nano devices that cannot be interconnected by standard lithographic techniques, thus in the range of nanometers. Figure 27 shows that the signal keeps its original shape after it travels to site G at 14.2 nm, which is a reasonable distance for a nanosized circuit. However, the amplitude of the signal at the carrier frequency, which is 23.81 THz, attenuates (Figure 29). The distance dependence of the amplitude shows that the amplitude attenuates exponentially with a factor of 0.0195 Å^{-1} (Figure 30). Thus, for a distance of 1.0 nm, the signal amplitude is still 82.3% of its original amplitude.

5.5.4. *Effects of different carrier frequencies on AM signal transmission*

When signals are coupled to atom X (forced movement of X), the atom that is directly bonded to X will also be driven to vibrate at the applied frequency; and this atom will again drive another bonded atom. As this process progresses, the vibrational signal propagates. As is discussed before, the carrier frequency of the signal should be one of the intrinsic vibrational frequencies of the backbone in the molecule.

We compare the response to two signals with different carrier frequencies: one at 23.81 THz corresponding to one of the intrinsic vibrational mode; and the other at 71.43 THz corresponding to a frequency range with no intrinsic vibrational modes. Figure 31 compares the evolution of these two signals. Notice that the later signal at

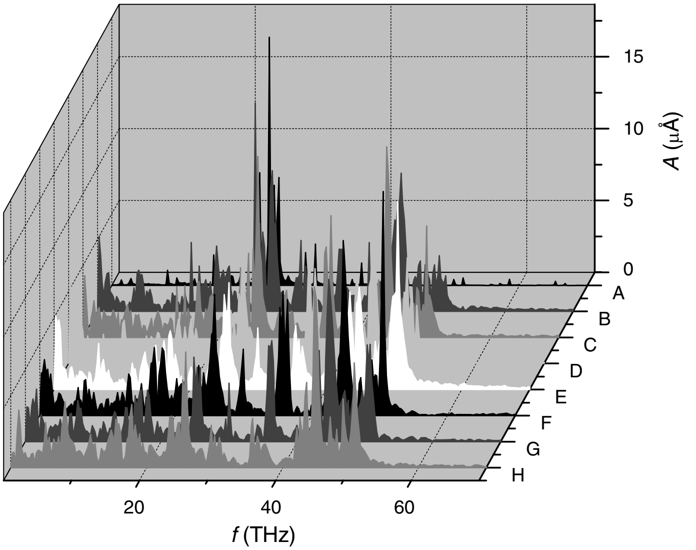


Figure 29 Detected frequency-domain AM signals in molecule GLY58 at different sites as indicated in Figure 25a [95]

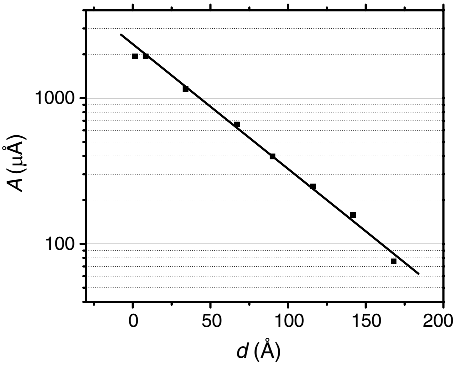


Figure 30 Attenuation of the AM signal along molecule GLY58 [95]

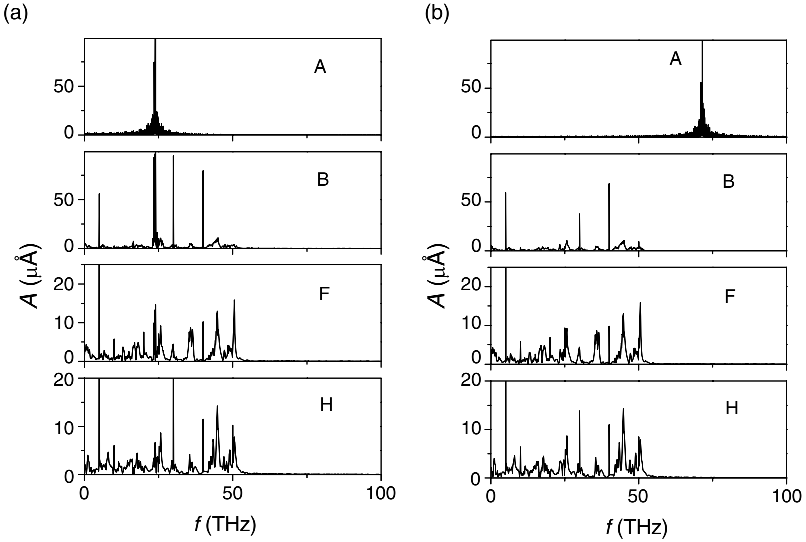


Figure 31 Frequency response of GLY58 of different AM carrier frequencies. (a) Frequency spectrums at sites A, B, F and H under excitation frequency of 23.81 THz; (b) frequency spectrums at sites A, B, F and H under excitation frequency of 71.43 THZ [95]

71.43 THz is much faster than the former that corresponds to an intrinsic mode of the polypeptide; actually, it almost dissolved immediately after it propagates through site A.

5.5.5. Molecular vibrational signal transmission using frequency modulation

Frequency-modulated signal is injected into the molecule GLY58 by coupling the input signal into the movement of carbon atom X as indicated in Figure 25. The time-domain signal at site A, 1.4 Å from the input site X, is shown in the upper plot of Figure 32a. Since the signal is only varied by frequency, it is difficult to distinguish the change in the time domain. In the frequency-domain signal (upper plot of Figure 32b), the carrier

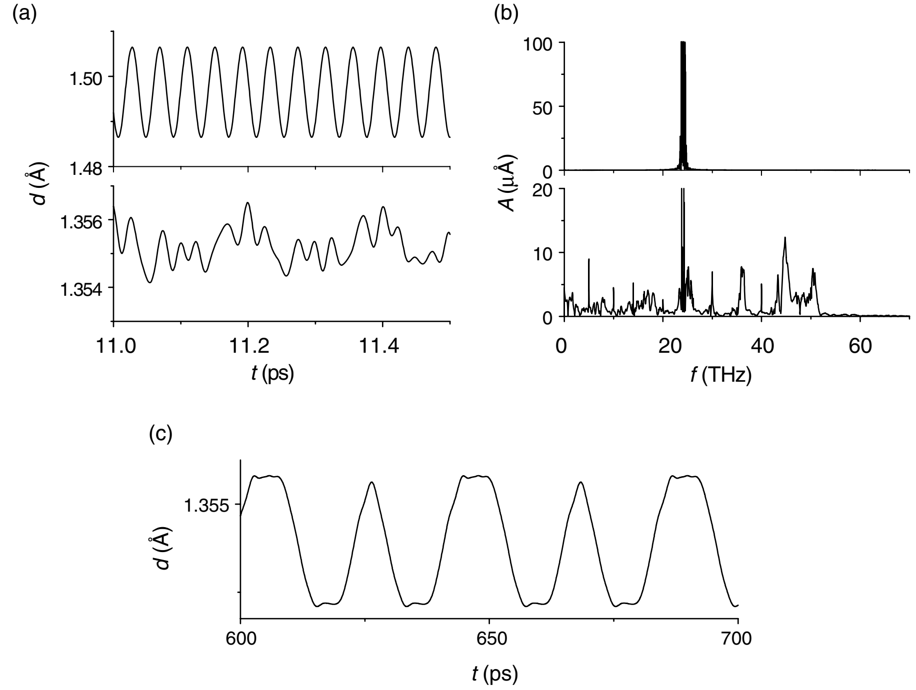


Figure 32 Signal transmission along GLY58 using frequency modulation by a carrier at 23.81 THz. (a) Time-domain vibrational signal detected at a remote site A (upper plot) and G (lower plot); (b) frequency-domain vibrational signal at site A (upper) and G (lower); (c) signal recovered using DSP techniques [95]

frequency is clearly shown as a peak in 23.81 THz. Figure 32a and 32b also show the time- and frequency-domain signals detected at site G, which is 142 Å from the input site.

The modulating signal is recovered as described in Section 5.4.2. The first Bessel bandpass filter is centered at 23.81 THz with a bandwidth of 1.0 THz. The bandpassed AC signal is clipped to constrain the amplitude variation to ±200 μÅ. The slope demodulator is centered at 24.9 THz with a bandwidth of 2.6 THz and the Bessel lowpass filter has a cutoff frequency of 0.5 THz. The recovered signal is shown in Figure 32c.

5.6. Using molecular potential to process information

For any molecular system, there is an electrostatic potential ϕ. The electrostatic potential ϕ at a point P is defined as the reversible work per unit charge needed to move an infinitesimal test charge Q_t from infinity to P [96]. The molecule can be viewed as a collection of point-charge nuclei and electronic charge smeared out into a continuous distribution. The probability of finding a molecular electron in a volume $\mathrm{d}V = \mathrm{d}x\mathrm{d}y\mathrm{d}z$ is [96]

$$\rho(x, y, z) = n \sum_{allm_x} \int \cdots \int |\psi(x, y, z, x_2, \ldots, z_n, m_{s1}, \ldots, m_{sn}|^2 \mathrm{d}x_2 \ldots \mathrm{d}z_n \qquad (7)$$

where ψ is the electronic wave function of the system comprising n electrons, and $|\psi(x_1, \ldots, z_n, m_{s1}, \ldots, m_{sn}|^2 \mathrm{d}x_1\mathrm{d}y_1\mathrm{d}z_1 \ldots \mathrm{d}x_n\mathrm{d}y_n\mathrm{d}z_z$ is the probability of simultaneously finding electron 1 with spin m_{s1} in the volume $\mathrm{d}x_1\mathrm{d}y_1\mathrm{d}z_1$ at (x_1, y_1, z_1), electron 2 with spin m_{s2} in the volume $\mathrm{d}x_2\mathrm{d}y_2\mathrm{d}z_2$ at (x_2, y_2, z_2), and so on [96]. The wave function ψ can be obtained by solving Schrödinger equation numerically using Gaussian 03 program.

Knowing the probability, we know the amount of electronic charge in $\mathrm{d}V$ is $-e\rho\mathrm{d}V$. The molecular electrostatic potential ϕ is the addition of the molecular electric charge and of the nuclei α. Thus, the molecular electrostatic potential at point (x_1, y_1, z_1) can be expressed as [96]

$$\phi(x_1, y_1, z_1) = \sum_{\alpha} \frac{Z_\alpha}{r_{1\alpha}} - \iiint \frac{\rho(x_2, y_2, z_2)}{r_{12}} \mathrm{d}x_2\mathrm{d}y_2\mathrm{d}z_2 \tag{8}$$

where Z_α is the nuclei charge, r_{12} is the distance between points 1 and 2 and the integration is over the entire space. Likewise, the molecular electrostatic potential is calculated using the Gaussian 03 program.

As an example, the molecular electrostatic potentials of water are plotted as in Figure 33. Positive and negative potentials can be easily distinguished by the color: positive is blue and cyan, and negative is red and yellow. Although the regions around the nuclei show positive potential, the sites outside the molecule show potential either positive (blue) or negative (red) depending on electron distributions around the molecule. For the singlet ground state (Figure 33a). the negative potential is located at the lower sites of the molecule; while for the triplet excited state, which is less angular, the negative effect of the electrons can be seen on both sites of the molecule.

The molecular electrostatic potential can be modified by external fields or excitations. For instance, the molecule triflourobenzene is excited by two water molecules conveniently located to produce four possible inputs (Figure 34), Assuming positive voltage is "1" and negative is "0", then if we record the potential at the right site of the hydrogen atom, the truth table can be generated (Figure 35). Thus, the molecular system performs as an AND gate.

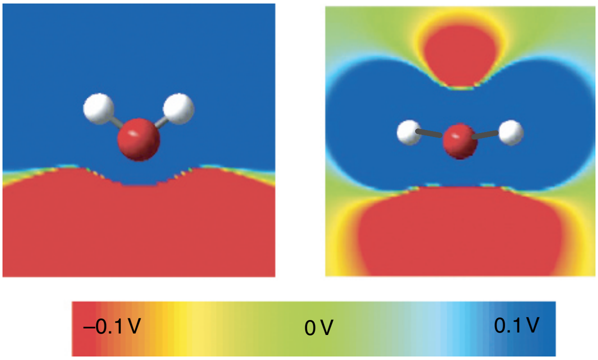


Figure 33 Molecular electrostatic potential [15]

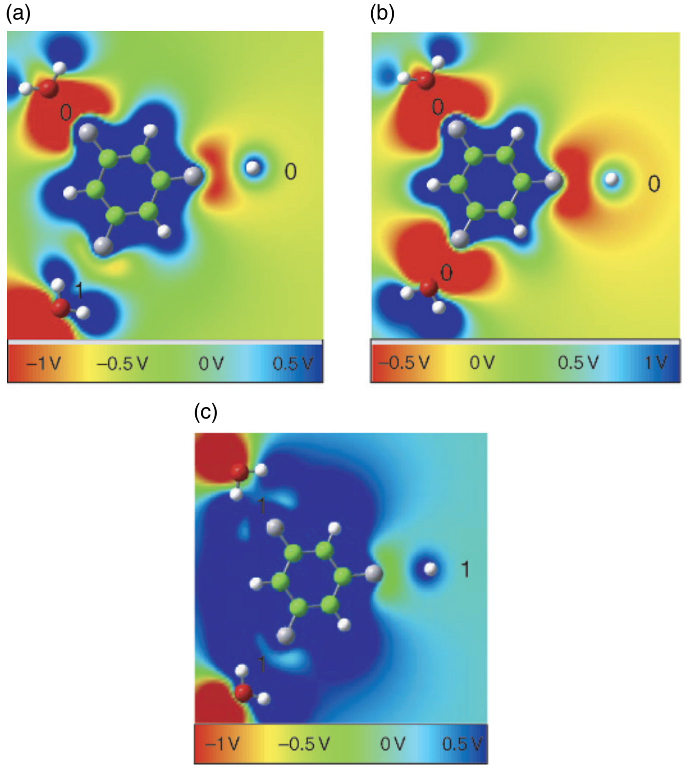


Figure 34 The implementation of a logical AND using a tri-flourobenzene molecule [15]

Input 1	Input 2	Output
1	1	1
1	0	0
0	1	0
0	0	0

Figure 35 Truth table of the molecular system shown in Figure 34 [15]

6. Conclusion and perspectives

Both AM and FM can be used to transmit information in molecular wires using molecular vibrations with a power dissipation of ~50 nW when working at 1 Tbps.

These kinds of vibrational movements are at the range of terahertz. The vibrational modes can not only be excited by an electromagnetic wave, but also be detected using infrared (IR) spectroscopy or Raman spectroscopy if their movements cause changes in electrostatic dipole (IR active) or in polarizability (Raman active).

When the signal is transferred using molecular vibrational modes, the atoms vibrate to a certain position, which may introduce a change in the MEP distribution of the whole molecular system and this change is subsequently transferred through vibronics.

Thus, by proper programming, the hundreds of molecules inside the nanoCell can be viewed as signal processing devices.

If information is transmitted using vibronics, the power dissipation could be evaluated from the energy that excites and keeps the molecular wire to vibrate although you do not need energy for a molecule to vibrate in its stationary states; molecules vibrate even at 0 K. Our simulation shows that the molecular wire dissipates 0.321 eV for transmitting one bit data using FM or 0.293 eV using AM. If the molecular wire transmits at 1 Tbps, the molecular wire will dissipate 51 nW using FM or 47 nW using AM, respectively.

Interestingly, the energy needed to excite the molecule with the two inputs is less than 0.5 eV. This implies that it may be possible to operate 200 millions of these gates at 10 GHz frequency with just less than 1 W. More precisely, $(2 \times 10^8 \text{gates}) \times (10^{10}\ \text{operation/s}) \times (0.5\,\text{eV/gate-operation}) \times (1.6 \times 10^9\,\text{J/eV}) = 0.16\,\text{J/s} = 0.16\,\text{W}$.

This is certainly a rough estimate, most likely within the same order of magnitude of the exact value. This estimation considers the energy to change states of the gate and the energy to excite their inputs. As gates are directly interconnected, losses in interconnections are perhaps a small fraction of the energy to change states. Nevertheless, even an error of one or two orders of magnitude underestimating the power consumption still represents an excellent result. Using a similar logic to calculate the power needed in the modern Pentium XE, assuming a gate capacitance of 40 aF, amazingly yields 125 W (most likely a lucky match) to which we need to add the energy dissipation in the interconnects, the energy needed to change the input states, and the utilization factor among others; all these should account for the total 130 W dissipation in the microprocessor.

The research work presented in this chapter focuses on the analysis of programmable molecular arrays, including electrical characterization of a nanoCell device, study of electron transport in discontinuous thin metal film as well as implementing molecular vibrational modes and molecular electrostatic potential as methods to process and transfer information using programmable molecular arrays. This research represents a major contribution to an in-depth investigation of the implementation of programmable molecular arrays.

First, the limitations by conventional silicon-based microelectronics in device fabrication and solid-state physics are outlined. New techniques such as single molecular electronic device, programmable molecular array, etc. are described that may help to overcome those limitations.

The electrical characteristics of programmable molecular array are thoroughly studied. The distinct features of programmable molecular array include the memory and switching phenomena and start-up transitional behavior. Effects of deposition of different molecules are investigated. It is found that the OPE molecules interlinking the gold islands do not contribute significantly to the electrical conductance. However, the alkane molecules create additional barrier for the electron transport through the islands. A programming procedure for the multi-leads molecular array is proposed. In this method, a matrix of truth valued between each pair of the leads is obtained and followed by a set of voltage pulses to program the molecular array.

Since the OPE molecules are found not to contribute significantly to the electrical conductance of the programmable molecular array, it is obvious that the discontinuous gold film is the fundamental conductor. After reviews of several theoretical models that explain the electrical conductance through a discontinuous metallic film, a model based on the clustering effect of gold atoms is proposed to account for the electrical behavior of discontinuous gold film in high electrical field. At low field, however, electrons have not overcome the barrier imposed by electron activation energy. So the discontinuous gold film exhibits a predictable current–voltage behavior in low field.

Finally, it is proposed that in the programmable molecular array, the molecular vibronics and MEP could replace the charge-current approach to transport and process information. The result of the simulation is provided as a primary proof for the concept. In addition, it is shown that the energy consumption using this approach is much smaller than the conventional charge-current-based device.

Much remains to be done in this topic of programmable molecular arrays. For instance, an effective electrical testing system has to be built to measure the conductance between each pair of the electrodes in the multi-leads programmable molecular array. Also, an electrical transducer needs to be developed in order to realize vibronics in the programmable molecular array. May this research work provide all of us several new ideas towards the realization of ultimate molecular computer in the future.

7. Acknowledgements

We thank the invaluable help of Mery Diaz Campos for her careful assistance in several aspects of this manuscript and we highly appreciate the support from the US Army Research Office (ARO), Defense Threat Reduction Agency (DTRA), and Defence Advanced Research Projects Agency (DARPA) as well as the research funds from the Texas A&M University.

References

[1] P. A. Derosa, V. Tarigopula, and J. M. Seminario, "Molecular electronics: Analysis and design of switchable and programmable devices using ab initio methods," in *Encyclopedia of Nanoscience and Nanotechnology*. New York: Dekker, 2004.

[2] J. M. Seminario, "A theory guided approach to molecular electronics," *Proc. IEEE Nanotech. Conf.*, vol. 3, pp. 75–78, 2003.

[3] P. A. Derosa, S. Guda, and J. M. Seminario, "A programmable molecular diode driven by charge-induced conformational changes," *J. Am. Chem. Soc.*, vol. 125, pp. 14240–14241, 2003.

[4] J. M. Seminario, A. G. Zacarias, and P. A. Derosa, "Analysis of a dinitro-based molecular device," *J. Chem. Phys.*, vol. 116, pp. 1671–1683, 2002.

[5] J. M. Seminario, P. A. Derosa, and J. L. Bastos, "Theoretical interpretation of switching in experiments with single molecules," *J. Am. Chem. Soc.*, vol. 124, pp. 10266–10267, 2002.

[6] J. M. Seminario, L. E. Cordova, and P. A. Derosa, "Search for minimum molecular programmable units," *Proc. IEEE Nanotech. Conf.*, vol. 2, pp. 421–424, 2002.

[7] J. M. Seminario, L. A. Agapito, and H. P. Figueroa, "Towards the design of programmable self-Assembled DNA-CNT: An approach to nanobiotronics," *Proc. IEEE Nanotech. Conf.*, vol. 2, pp. 287–290, 2002.

[8] J. M. Seminario, A. G. Zacarias, and P. A. Derosa, "Theoretical analysis of complementary molecular memory devices," *J. Phys. Chem. A*, vol. 105, pp. 791–795, 2001.

[9] J. M. Seminario and P. A. Derosa, "Molecular gain in a thiotolane system," *J. Am. Chem. Soc.*, vol. 123, pp. 12418–12419, 2001.

[10] J. M. Seminario, C. E. De La Cruz, and P. A. Derosa, "A theoretical analysis of metal–molecule contacts," *J. Am. Chem. Soc.*, vol. 123, pp. 5616–5617, 2001.

[11] J. M. Seminario and L. E. Cordova, "Toward multiple-valued configurable random molecular logic units," *Proc. IEEE Nanotech. Conf.*, vol. 1, pp. 146–150, 2001.

[12] P. A. Derosa and J. M. Seminario, "Electron transport through single molecules: Scattering treatment using density functional and green function theories," *J. Phys. Chem. B*, vol. 105, pp. 471–481, 2001.

[13] J. M. Seminario, A. G. Zacarias, and J. M. Tour, "Molecular current–voltage characteristics," *J. Phys. Chem. A*, vol. 103, pp. 7883–7887, 1999.

[14] J. M. Tour, M. Kosaki, and J. M. Seminario, "Molecular scale electronics: A synthetic/computational approach to digital computing," *J. Am. Chem. Soc.*, vol. 120, pp. 8486–8493, 1998.

[15] J. M. Seminario, Y. Ma, and V. Tarigopula, "The nanocell: A chemically assembled molecular electronic circuit," *IEEE Sensors*, in press.

[16] J. M. Seminario, L. Yan, and Y. Ma, "Scenarios for molecular-level signal processing," *Proc. IEEE*, vol. 93, pp. 1753–1764, 2005.

[17] Y. Ma and J. M. Seminario, "Transient behavior at the nanoscale," *J. Chem. Phys.*, vol. 124, pp. 171102-1-2, 2006.

[18] Y. Ma and J. M. Seminario, "The atomistic nature of transient and steady state responses," *J. Phys. Chem. B*, vol. 110, pp. 9708–9712, 2006.

[19] J. M. Seminario, L. Yan, and Y. Ma, "Encoding and transport of information in molecular and biomolecular systems," *Proc. IEEE Nanotech. Conf.*, vol. 5, pp. 65–68, 2005.

[20] L. Yan, Y. Ma, and J. M. Seminario, "Terahertz signal transmission in molecular systems," *Int. J. High Speed Electron. Syst.*, in press.

[21] Y. Ma, L. Yan, and J. M. Seminario, "Molecular-based processing and transfer of information in the terahertz domain for military and security applications," *Proc. SPIE*, vol. 6212, pp. 621204-1-8, 2006.

[22] L. Yan and J. M. Seminario, "Moletronics modeling towards molecular potentials," *Int. J. Quantum Chem.*, vol. 106, pp. 1964–1969, 2006.

[23] L. Yan, Y. Ma, and J. M. Seminario, "Encoding information using molecular vibronics," *J. Nanoscience Nanotech*, vol. 6, pp. 675–684, 2006.

[24] J. M. Seminario, L. Yan, and Y. Ma, Nano-detectors using molecular circuits operating at THz frequencies, *Proc. SPIE*, vol. 5995, 2005.

[25] J. M. Seminario, L. Yan, and Y. Ma, "Transmission of vibronic signals in molecular circuits," *J. Phys. Chem. A*, vol. 109, pp. 9712–9715, 2005.

[26] J. M. Seminario, P. A. Derosa, L. E. Cordova, and B. H. Bozard, "Molecular dynamics simulations of a molecular electronic device: The nanoCell," in J. Leszczynski, ed., *Comp. Chem.: Rev. Curr. Trends*, vol. 9, Singapore: World Scientific, 2005, pp. 85–119.

[27] J. M. Seminario, P. A. Derosa, B. H. Bozard, and K. Chagarlamudi, "Vibrational study of a molecular device using molecular dynamics simulations," *J. Nanoscience Nanotech*, vol. 5, pp. 1–11, 2005.

[28] J. M. Seminario, "Molecular electronics: Approaching reality," *Nature Materials*, vol. 4, pp. 111–113, 2005.

[29] Y. Ma, L. Yan, and J. M. Seminario, "Terahertz molecular electronics devices and systems," in R. J. Hwu, D. L. Woolard, and M. J. Rosker, eds, *Proc. SPIE*, vol. 5790, 2005, pp. 206–218.

[30] J. M. Seminario, Y. Ma, L. A. Agapito, L. Yan, R. A. Araujo, S. Bingi, N. S. Vadlamani, K. Chagarlamudi, T. S. Sudarshan, M. L. Myrick, P. E. Colavita, P. D. Franzon, D. P. Nackashi, L. Cheng, Y. Yao, and J. M. Tour, "Clustering effects on discontinuous gold film nanocells," *J. Nanoscience Nanotech.*, vol. 4, pp. 907–917, 2004.

[31] J. M. Seminario, P. A. Derosa, L. E. Cordova, and B. H. Bozard, "A molecular device operating at terahertz frequencies," *IEEE Trans. Nanotech.*, vol. 3, pp. 215–218, 2004.

[32] G. E. Moore, "Cramming more components onto integrated circuits," *Electronics*, vol. 38, 1965.

[33] G. E. Moore, "Lithography and future of Moore's law," presented at Proc. SPIE, 1995.

[34] S. A. Campbell, *The Science and Engineering of Microelectronic Fabrication*: Oxford University Press, 2001.

[35] "International technology roadmap for semiconductors," in Semiconductor Industry Association, 2005.

[36] J. A. Wheeler and W. Zurek, *Quantum Theory and Measurement (Princeton Series in Physics*, Princeton University Press, 1984.

[37] A. Aviram and M. A. Ratner, "Molecular rectifiers," *Chem. Phys. Lett.*, vol. 29, pp. 277, 1974.

[38] P. Sharma, "How to create a spin current," *Science*, vol. 307, pp. 531–533, 2005.

[39] S. A. Wolf, D. D. Awschalom, R. A. Buhrman, J. M. Daughton, S. V. Molnar, M. L. Roukes, A. Y. Chtchelkanova, and D. M. Treger, "Spintronics: A spin-based electronics vision for the future," *Science*, vol. 294, pp. 1488–1495, 2001.

[40] J. M. Tour, M. A. Reed, J. M. Seminario, D. A. Allara, and P. A. Weiss, "Molecular computer," in *US Patent 6,430,511*, 2002.

[41] J. M. Tour, *Molecular Electronics. Commercial Insights, Chemistry, Devices, Architecture and Programming*. New Jersey: World Scientific, 2003.

[42] J. M. Tour, W. L. VanZandt, C. P. Husband, S. M. Husband, L. S. Wilson, P. D. Franzon, and D. P. Nackashi, "NanoCell logic gates for molecular computing," *IEEE Trans. Nanotech.*, vol. 1, pp. 100–109, 2002.

[43] J. M. Tour, L. Cheng, D. P. Nackashi, Y. Yao, A. K. Flatt, S. K. St. Angelo, T. E. Mallouk, and P. D. Franzon, "NanoCell electronic memories," *J. Am. Chem. Soc.*, vol. 125, pp. 13279–13283, 2003.

[44] J. M. Seminario, L. E. Cordova, and P. A. Derosa, "An ab-initio approach to the calculation of current–voltage characteristics of programmable molecular devices," *Proc. IEEE*, vol. 91, pp. 1958–1975, 2003.

[45] J. M. Seminario, A. G. Zacarias, and J. M. Tour, "Molecular alligator clips for single molecule electronics. Studies of group 16 and isonitriles interfaced with Au contacts," *J. Am. Chem. Soc.*, vol. 121, pp. 411–416, 1999.

[46] D. P. Nackashi, N. H. Di Spigna, D. A. Winick, C. J. Amsinck, L. Cheng, J. M. Tour, and P. D. Franzon, "Discontinuous gold films for nanocell memories," Technical Proceedings of the 2004 NSTI Nanotechnology Conference and Trade Show, *Nanotech.*, vol. 3, pp. 45–48, 2004.

[47] J. Chen, W. Wang, M. A. Reed, A. M. Rawlett, D. W. Price, and J. M. Tour, "Room-temperature negative differential resistance in nanoscale molecular junctions," *Appl. Phys. Lett.*, vol. 77, pp. 1224–1226, 2000.

[48] J. Chen, M. A. Reed, A. M. Rawlett, and J. M. Tour, "Large on-off ratios and negative differential resistance in a molecular electronic device," *Science*, vol. 286, pp. 1550–1552, 1999.

[49] C. P. Husband, S. M. Husband, J. S. Daniels, and J. M. Tour, "Logic and memory with nanocell circuits," *IEEE Trans. Electron Dev.*, vol. 50, pp. 1865–1875, 2003.

[50] J. M. Seminario, Y. Ma, L. A. Agapito, L. Yan, R. A. Araujo, S. Bingi, N. S. Vadlamani, K. Chagarlamudi, T. S. Sudarshan, M. L. Myrick, P. E. Colavita, P. D. Franzon, D. P. Nackashi, L. Cheng, Y. Yao, and J. M. Tour, "Clustering effects on discontinuous gold film nanoCells," *J. Nanosci. nanotech.*, vol. 4, pp. 907–917, 2004.

[51] L. Cai, Y. Yao, J. Yang, D. W. J. Price, and J. M. Tour, "Chemical and potential-asisted asembly of thiolacetyl-terminated oligo(phenylene ethynylene)s on gold surfaces," *Chem. Mater.*, vol. 14, pp. 2905–2909, 2002.

[52] G. J. Kluth, M. M. Sung, and R. Maboudian, “Thermal behavior of alkylsiloxane self-assembled monolayers on the oxidized Si(100) surface,” *Langmuir*, vol. 13, pp. 3775–3780, 1997.

[53] N. M. Bashara and L. A. Weitzenkamp, “Conduction in very thin films at high electric fields,” *J. Appl. Phys.*, vol. 35, pp. 1983–1984, 1964.

[54] P. J. Dobson and B. J. Hopkins, “Electrical Conductance of Thin ultrahigh-vacuum-evaporated films of tungsten on glass,” *J. of Appl. Phys.*, vol. 39, pp. 3074–3081, 1968.

[55] R. Kiernan and D. W. Stops, “Tunnelling between metallic islands on a dielectric substrate,” *Nature*, vol. 224, pp. 907–908, 1969.

[56] J. E. Morris, “Non-ohmic properties of discontinuous thin metal films,” *Thin Solid Films*, vol. 11, pp. 81–89, 1972.

[57] C. A. Neugebauer and M. B. Webb, “Electrical conduction mechanism in ultrathin, evaporated metal films,” *J. Appl. Phys.*, vol. 33, pp. 74–82, 1962.

[58] K. Uozumi, M. Nishiura, and A. Kinbara, “On the field effect of the electrical conductance of discontinuous thin metal films,” *J. Appl. Phys.*, vol. 48, pp. 818–819, 1977.

[59] K. Uozumi, “A model to account for the field effect of the electrical conductance of discontinuous thin metal films,” *Thin Solid Films*, vol. 57, pp. 141–146, 1979.

[60] K. Uozumi, “Non-ohmic conductance of discontinuous thin gold films,” *Thin Solid Films*, vol. 89, pp. 45–48, 1981.

[61] M. Shin, S. Lee, K. W. Park, and E.-H. Lee, “Geometrically induced multiple coulomb blockade gaps,” *Phys. Rev. Lett.*, vol. 80, pp. 5774–5777, 1998.

[62] M. Shin, S. Lee, K. W. Park, and E. Lee, “Secondary Coulomb blockade gap in a four-island tunnel-junction array,” *Phys. Rev. B*, vol. 59, pp. 3160–3167, 1999.

[63] R. M. Hill, “Electrical conduction in ultra thin metal films. II. experimental,” *Proc. R. Soc. London, Ser. A*, vol. 309, pp. 397–417, 1969.

[64] R. M. Hill, “Electrical Conduction in Ultra Thin Metal Films. I. theoretical,” *Proc. R. Soc. London, Ser. A*, vol. 309, pp. 377–395, 1969.

[65] M. Amman, E. Ben-Jacob, and K. Mullen, “Charge solitons in 1-D array of mesoscopic tunnel junctions,” *Phys. Lett. A*, vol. 142, pp. 431–437, 1989.

[66] G. Y. Hu and R. F. O’Connell, “Exact solution for the charge soliton in a one-dimensional array of small tunnel junctions,” *Phys. Rev. B* vol. 49, pp. 16773–16776, 1994.

[67] C. B. Whan, J. White, and T. P. Orlando, “Full capacitance matrix of coupled quantum dot arrays: Static and dynamical effects,” *Appl. Phys. Lett.*, vol. 68, pp. 2996–2998, 1996.

[68] K. K. Likharev and K. A. Matsuoka, “Electron–electron interaction in linear arrays of small tunnel junctions,” *Appl. Phys. Lett.*, vol. 67, pp. 3037–3039, 1995.

[69] M. Shin, S. Lee, K. W. Park, and E.-H. Lee, “Additional Coulomb blockade and negative differential conductance in closed two-dimensional tunnel junction arrays,” *J. Appl. Phys.*, vol. 84, pp. 2974–2976, 1998.

[70] J. G. Simmons and R. R. Verderber, “New conduction and reversible memory phenomena in thin insulating films,” *Proc. R. Soc. London, Ser. A*, vol. 301, pp. 77–102, 1967.

[71] V. Cimrová and D. Neher, “Anomalous electrical characteristics, memory phenomena and microcavity effects in polymeric light-emitting diodes,” *Synth. Met.*, vol. 76, pp. 125–128, 1996.

[72] H. Kusano, N. Shiraishi, S. Hosaka, I. Kuruma, M. Kitagawa, K. Ichino, and H. Kobayashi, “Carrier transport mechanism of PVCz-based multi-layered electroluminescent devices,” *Synth. Met.*, vol. 91, pp. 341–343, 1997.

[73] J. Manca, W. Bijnens, R. Kiebooms, J. D’Haen, M. D’Olieslaeger, T. Wu, W. d. Ceuninck, L. d. Schepper, D. Vanderzande, J. Gelan, and L. Stals, “Effect of oxygen on the electrical characteristics of PPV-LEDs,” *Opt. Mater.*, vol. 9, pp. 134–137, 1998.

[74] S. Berleb, W. Brütting, and M. Schwoerer, “Anomalous current–voltage characteristics in organic light-emitting devices,” *Synth. Met.*, vol. 102, pp. 1034–1037, 1999.

[75] X. Xu, S. Yin, F. Kong, F. Zhang, K. Yao, and W. Huang, "Organic negative-resistance devices using PPV containing electron-transporting groups on the main chain," *J. Polym. Sci., Part B: Polym. Phys.*, vol. 39, pp. 589–593, 2001.

[76] E. Lebedev, S. Forero, W. Brütting, and M. Schwoerer, "Switching effect in poly(p-phenylenevinylene)," *Synth. Met.*, vol. 111–112, pp. 345–347, 2000.

[77] L. Ma, J. Liu, S. Pyo, and Y. Yang, "Organic bistable light-emitting devices," *Appl. Phys. Lett.*, vol. 80, pp. 362–364, 2002.

[78] L. Ma, Q. Xu, and Y. Yang, "Organic nonvolatile memory by controlling the dynamic copper-ion concentration within organic layer," *Appl. Phys. Lett.*, vol. 84, pp. 4908–4910, 2004.

[79] L. Ma, J. Liu, and Y. Yang, "Organic electrical bistable devices and rewritable memory cells," *Appl. Phys. Lett.*, vol. 80, pp. 2997–2999, 2002.

[80] L. Ma, S. Pyo, J. Ouyang, Q. Xu, and Y. Yang, "Nonvolatile electrical bistability of organic/metal-nanocluster/organic system," *Appl. Phys. Lett.*, vol. 82, pp. 1419–1421, 2003.

[81] L. D. Bozano, B. W. Kean, V. R. Deline, J. R. Salem, and J. C. Scott, "Mechanism for bistability in organic memory elements," *Appl. Phys. Lett.*, vol. 84, pp. 607–609, 2004.

[82] W. Brütting, H. Riel, T. Beierlein, and W. Riess, "Influence of trapped and interfacial charges in organic multilayer light-emitting devices," *J. Appl. Phys.*, vol. 89, pp. 1704–1712, 2001.

[83] W. Tang, H. Shi, G. Xu, B. S. Ong, Z. D. Popovic, J. Deng, J. Zhao, and G. Rao, "Memory effect and negative differential resistance by electrode-induced two-dimensional single-electron tunneling in molecular and organic electronic devices," *Advanced Materials*, vol. 17, pp. 2307–2311, 2005.

[84] S. M. Sze, *Physics of Semiconductor Devices*, 2nd edn, New York: Wiley, 1981.

[85] Intel, "Expanding Moore's Law," 2002.

[86] J. M. Seminario, R. A. Araujo, and L. Yan, "Negative differential resistance in metallic and semiconducting clusters," *J. Phys. Chem. B*, vol. 108, pp. 6915–6918, 2004.

[87] J. M. Seminario, Y. Ma, and V. Tarigopula, "The nanoCell: A chemically assembled molecular electronic circuit," *IEEE Sensors*, in press.

[88] S. P. Gurrum, S. K. Suman, Y. K. Joshi, and A. G. Fedorov, "Thermal issues in next-generation integrated circuits," *IEEE Transactions on Device and Materials Reliability*, vol. 4, pp. 709–714, 2004.

[89] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*. Oxford: Clarendon Press, 1990.

[90] L. Yan, Y. Ma, and J. M. Seminario, "Terahertz signal transmission in molecular systems," *Int. J. High Speed Electronics Syst.*, in press.

[91] J. M. Seminario, "Molecular devices operating at terahertz frequencies: Theoretical simulations and perspectives," presented at Proc. SPIE, 2004.

[92] E. C. Ifeachor and B. W. Jervis, *Digital Signal Processing: A Practical Approach*, 2nd Edn, Pearson Education Limited, 2002.

[93] J. H. McClellan, R. W. Schafer, and M. A. Yoder, *DSP First: A Multimedia Approach*: Tom Robbins, 1998.

[94] "Labview, Analysis Concepts," in *National Instruments*, 2004.

[95] L. Yan, Y. Ma, and J. M. Seminario, "Encoding information using molecular vibronics," *J. Nanosci. Nanotechnol.*, vol. 6, pp. 685–690, 2006.

[96] I. N. Levine, *Quantum chemistry*, Prentice-Hall, 2000.

Molecular and Nano Electronics: Analysis, Design and Simulation
J. M. Seminario (Editor)

Chapter 5

Modeling molecular switches: A flexible molecule anchored to a surface

Bidisa Das and Shuji Abe

Nanotechnology Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Umezono 1-1-1, Tsukuba 305-8568, Japan.
bidisa.das@aist.go.jp

1. Introduction

Miniaturization of electronic devices is continuing steadily and will hopefully approach the scale of molecules in the near future. Current research [1–5] in this field aims to incorporate a small number of active molecules into an electronic circuit and make them perform suitable electronic functions. Molecules are particularly interesting as active elements of electronic circuits, due to their tiny size, synthetic tailor-ability, properties of molecular recognition and molecular assembly and their dynamic stereochemistry. Though a large number of studies have already been carried out on the design and synthesis of molecular wires [6–9], rectifiers [10–13], transistors [14–16], switches [17–19], memories [20] etc., the fundamental understanding of the molecular processes that govern the working principle of such devices and the effective manipulation techniques are still not adequate to build efficient molecular-level electronic devices. Theoretical and computational design of functional molecules or molecular assemblies is a key to the development of such bottom-up nanotechnologies. Computational experiments allow us to examine structures quickly and easily, rejecting those which have obvious defects. By fully utilizing the potential of various theoretical modeling techniques that has been developed in the recent years, design and development of complex systems like molecular devices are feasible.

In this chapter we discuss molecular switches, focusing mainly on conductance switching systems. Switches are most basic components of an electronic circuit. A molecular switch is a nanoscale machine which switches reversibly between two or more states. Actually, any drastic change in the properties of a molecule due to some external stimulus may be considered as a switching phenomenon and possibly be utilized to design a molecular switch. Many molecules have different conformational or electronic states depending on its external conditions; such molecules are very promising candidates for the design and synthesis of molecular switches. A structural change of a molecule is analogous to the mechanical motion of a macroscopic system. Although a mechanical switch is usually too slow to be used in modern electronic devices of silicon technology, the situation may be completely different in the nanometer scale, where molecules can change its conformation rapidly enough to be used as a switching device. To be useful as a device the switching has to be reversible and triggered by a suitable external stimulus like light, electric field or magnetic field. One basis for such a device is an adsorbed molecule that can be electrically switched between two states, namely ON and OFF. This type of molecular switches has been designed and investigated, for example, in the case of [2]rotaxanes anchored to a metal surface [21, 22]. The molecule has a mobile macrocyclic part, and the two states correspond to two structures where the marcocycle resides in distinctly different positions of the molecule.

Perhaps the most difficult part in this research is how to control individual molecules. Break junctions [23–25], nano-gap electrodes [26–28], and crosswire assemblies [29] have been used for making electronic devices incorporating a small number of molecules. Nanoprobes such as scanning tunneling microscope (STM) and atomic force microscope (AFM) are powerful tools for the manipulation of a molecule adsorbed on surfaces using local forces. Our design of a reversible molecular switch [30] based on a polar amide molecule on surface is appropriate for conductance studies done by using STM.

2. Various types of molecular switches

There can be many types of interesting switching phenomena in molecules. A full description and survey of all the types is beyond the scope of this chapter; however, we try to outline briefly a few very interesting cases. Photoresponsive switching for molecules has been studied by many groups [31–33] and is an area of great interest. In this class of molecular switches, the molecule shows different optical properties for different conformations. Particularly promising photoresponsive molecules that have been used to design optical switches are azobenzene [34–36], carbocynine [37], diarylethenes bearing two thiophene-derived groups [38, 39] etc.

Molecular scale mechanical switching based on interlocked molecules mainly including rotaxanes and catenanes have been studied in details [21, 22, 40–43]. These interlocked systems generally consist of an electron-rich and an electron-deficient portion held together by non-covalent bonds. The magnitudes of the non-covalent bonding interactions that control the locations of the macrocycles in bistable catenanes and rotaxanes can be switched reversibly in solution by redox processes triggered by chemical, electrochemical and photochemical stimuli. Bistable rotaxanes are also known as molecular shuttles which consist of a molecular rod with two stations (recognition sites), onto which a macrocycle is threaded. The macrocyle is constrained to move within the thread due to the presence of bulky stoppers at the ends of the molecular rod portion. Presence of external stimulus initiates a redox reaction, as a result of which the macrocyle moves from one recognition site to the other reversibly. This type of molecular actuation gives rise to the OFF and ON states of the switch.

Conductance switching in molecules is our main topic of interest in this chapter. In molecular systems, conductance switching occurs when two of its conformations show distinctly different conducting properties which can be switched reversibly by means of external electric field. Experimental studies of conductance switching in molecules are very difficult, but with the development of scanning probe technologies, many studies in this field are now being reported. STM techniques allow for investigation of surface properties at very high resolution. Since an STM tip can be positioned with great precision, local electronic transport measurements are possible, enabling one to measure conductance of a single molecule [44–46]. A particularly interesting discovery in this field is the negative differential resistance (NDR) in ethynylphenyl-based organic molecules at room temperatures in self-assembled monolayer [47]. The molecule contains an ethynylphenyl backbone but a redox center is introduced in the middle benzene ring: an electron-withdrawing nitro ($-NO_2$) and an electron-donating amino ($-NH_2$) group. Current–voltage measurements of the device show NDR and a large ON–OFF peak-to-valley ratio in excess of 1000:1. The authors at first suggested a two-step reduction process that modifies charge transport through the molecule as the potential mechanism for this NDR effect. However, there is a debate [46, 48–51] on the origin of NDR in this molecule and it is not very well understood yet.

There can be two types of conductance switching behaviors, namely stochastic switching attributed to statistical fluctuations in the film or molecule [52–54] and switching due to an applied external voltage [55–59]. Random statistical switching, also known as stochastic switching, has only been observed using scanning probe techniques on individual molecules isolated in an insulating matrix [51–53, 60]. This is a transient, random process, independent of applied voltages and its origin had been debated over long time. Since this switching phenomenon is difficult to control, it is not certain how it can be utilized for designing molecular devices where the key idea is controllability. Studies of isolated individual molecules embedded in a matrix of alkanethiolate monolayers by Donhauser *et al.* [61, 62] had shown conductance switching of single phenylene ethynylene oligomers (OPE) embedded in matrices of alkanethiolates when studied using STM. The molecules switched reversibly between discrete states that differ in their apparent heights. Although the origin of the switching was not well understood, it was interpreted as a tilting of the inserted molecules or a change in the orientation of the molecule with respect to the STM tip. This type of motion can cause a change in the physical height of the molecule above the host matrix. The height is an important parameter in the case of STM experiments, where an exponential decrease in the tunneling current is observed due to an increase in the tip–sample distance. Therefore, a small increase in the height of the molecule on the surface can result in a large increase in the tunneling current. Weiss and co-workers [63] has recently studied six customized OPEs for their electronic properties using STM to test the true mechanisms of stochastic conductance switching. They tested for many previously suggested mechanisms, which included functional group reduction [64], functional group rotation [65], backbone ring rotation [66], neighboring molecule interactions [67, 68], bond fluctuations [60, 61, 69] and hybridization changes [61–63] by making appropriate modifications in the molecules. Finally, they concluded [63] that hybridization changes at the molecule–surface interface are responsible for the switching phenomenon. The hybridization change can occur through surface reconstruction or a change in the alignment of the molecule with the surface.

For strongly polar molecules, it is possible to reorient the dipole of the molecule by application of external bias voltage. Orientational switching caused due to the rotation of dipole as a result of applied electric field has been reported by Yanagi *et al.* [58]. They observed a reversible, orientational switching of chloro[subphthalocyaninato]-boron(III) molecule with STM. This molecule has a three-fold symmetric structure like shuttlecock with an axial chlorine head binding to the central boron atom. When adsorbed on Cu(100) surface, two orientations are possible: the axial chlorine atom upward or downward. After scanning at a positive or negative bias, the molecules were observed to switch to the upward or downward orientation, respectively. This clearly indicated that the electric field coupled with the dipole moment of the molecule strongly and could cause the flipping of the molecule on the surface. In another interesting study, Ishida *et al.* [70] observed apparent molecular motion induced by the polarity change of electric fields by STM. They used disulfide molecules containing a terphenyl moiety with a large dipole moment, embedded into alkanethiol self-assembled monolayers. From the STM measurements the authors concluded that the observed apparent height change was caused by the conductance change (rectification property) of the electrically active terphenyl moiety, although it could not be explained by a simple coupling between the electric field and the dipole moment. Recently Kitagawa *et al.* observed conductance switching of peptide helix bundles on a gold substrate by STM [59]. These are helical molecules with many amide groups linked by intramolecular hydrogen bonds and are capable of exhibiting two different lengths corresponding to an α-helix structure and a 3_{10}-helix structure. The conductance of the helix alternated between the two states by changing the polarity of applied bias. The conductance and the apparent molecular length were also observed to undergo stochastic changes with time. Since the molecules considered in this study are highly polar, the coupling of the dipole moment and the applied electric field may be an important factor controlling the switching.

There exist also a number of theoretical predictions and studies of NDR and associated conductance switching phenomena in different classes of molecules, and different mechanisms have been proposed. Seminario and co-workers have studied the electronic structure and geometry of the isolated OPE molecules and tried to explain the NDR mechanism found experimentally in OPEs by Reed and co-workers [47]. They proposed that NDR in these molecules is caused by the change of the electronic charge state of the molecule under increasing bias voltages and the resulting change of the molecular conformation due to the change of the charge state [71, 72]. The extended and the localized nature of the molecular orbitals under reduced and neutral conditions formed the basis of this study. Further analysis of this molecule sandwiched between two gold electrodes was performed by Stokbro *et al.* [73] using a combination of density functional and non-equilibrium Green function methods. They concluded that functional groups present in the OPEs have a stronger effect on the energetics of the monolayers than on the individual molecular orbitals responsible for current transport, hence a better understanding of the intermolecular interactions in such monolayers is important. Coherent electron transport study through a metal–molecule–metal junction consisting of photoactive azobenzene molecule is reported [74]. The conductance of the *cis* conformation of azobenzene molecule was found to be two orders of magnitude less than the conductance of the *trans* isomer. The *trans* isomer is expected to be a better conductor because of its planar orientations of the phenyl rings, giving rise to delocalized conduction channels. On the other hand, the conductance of the *cis* isomer is low because of different orientations of the molecular orbitals in the two rings. Another theoretical study of single molecule conduction switching of photochromic dithienylethene molecule is available [75]. It reports a large change in conduction due to optical switching of dithienylethene. The molecular switching process is found to produce a swapping of the highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) during the conformational change.

In a very simple theoretical analysis, Torisi and Ratner [76] have recently shown that ‘off-resonance rectification’ can be achieved by exploiting the conformational changes in molecules sandwiched between metal electrodes, driven by an external electric field. Molecules with polar groups were suggested as possible candidates, which rearrange in space by rotation around σ bonds. Our theoretical study for substituted benzamide molecule is in the same direction, where we utilize the conformational change due to bond rotation in applied electric fields to design a reversible molecular switch on metal surface [30].

3. Molecules with amide groups – Conductance switching in applied electric fields

Conductance switching in molecules can take place due to various reasons. For example, conductance switching can take place when *cis* azobenzene molecule is converted to *trans* isomer, or an open dithienylethene is converted to closed isomer due to photoexcitation [74, 75]. In both the situations, the conductance switching is governed by the change in the molecular orbitals. In metal–molecule–metal junction studies, nature of the molecular orbitals plays a very vital role. The π-type molecular orbitals that are extended over the full molecule can act as conduction channels for electrons in two-probe systems. However, if STM studies on SAMs are considered, conductance switching may occur if the height of the molecule from the substrate changes. The change in the height of the molecule can be due to isomerization process, bond rotation, change in titling angle of the molecule, etc. In STM studies an increase in the height of the molecule ensures a decreased tip–sample distance, thereby increasing the tunneling current exponentially. The STM image, of course, does not depend only on the tip–sample distance, but also on the electronic structure of the systems. We argue that the presence of a polar amide unit (–CONH) in a molecule not only makes it responsive towards applied electric field, it is flexible enough to be rotated at low available energies which can cause the molecular height to change drastically depending on the molecule (for adsorbed cases). This may result in conductance switching depending on the specific molecule concerned. In the next section, we discuss about a model molecule with one amide group on Au(111) surface in the presence of applied electric field. Situations can be more complicated if more number of amide groups are present in the molecule.

4. *N*-(2-mercaptoethyl)benzamide on Au(111): A reversible molecular switch

N-(2-mercaptoethyl)benzamide (see Figure 1) is a simple molecule with its dipole moment largely arising due to the presence of the polar amide moiety in it. The molecule also has a thiol end group, which ensures that it can chemisorb on Au(111) surface. The molecule can exist as *trans* amide and *cis* amide structures. We have studied this

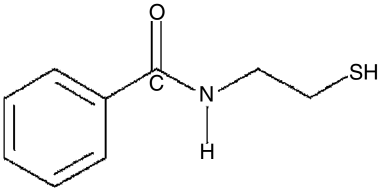


Figure 1 *N*-(2-mercaptoethyl)benzamide

molecule in free state and in adsorbed conditions. We have used a model surface of 21 Au atoms arranged in (111) fashion for adsorption studies. There are 20 atoms in one layer with only one atom added in the second layer below the adsorption site to mimic the hcp and fcc adsorption sites. Details of the calculations are discussed in [30].

4.1. Free and adsorbed conformations

In the optimized structure of *N*-(2-mercaptoethyl)benzamide, the C=O bond is *trans* to the N–H bond. The studies of the free molecules reveal that the *cis* amide conformation of *N*-(2-mercaptoethyl)benzamide is 6.38 kcal/mol less stable than the *trans* amide conformation. Conversion of *trans* amide structure to *cis* amide structure barely occurs at room temperature as the barrier height is about 15 kcal/mol, hence, only the *trans* conformation predominates in normal conditions. A rotation about the C–N bond adjacent to the CONH unit of the molecule leaves the free molecule nearly unchanged but reverses the orientation of the molecular dipole with respect to the surface. We found that the barrier height for the rotation of this bond is low (5.29 kcal/mol), and the rate of reaction at 300 K was estimated to be $1.1 \times 10^8\,s^{-1}$ using transition state theory [77, 78]. This essentially means that the rotation may occur at room temperature due to thermal fluctuations.

According to previous studies on the bonding of thiol molecules on Au(111) surface [79–85], there can be on-top, fcc-hollow, hcp-hollow, bridge, fcc-bridge and hcp-bridge adsorption sites for the sulfur atom in the thiol molecule to bond to the surface. Most studies indicated that upon adsorption of alkane thiols on Au(111) the S atom is bonded to two Au atoms on the surface (the bridge bond) with slight tilting towards the fcc site (called the fcc-bridge site) [79, 80, 82, 85]. The studies also suggest that the hcp-bridge and fcc-bridge sites have virtually same energy. We studied *N*-(2-mercaptoethyl)benzamide on the model 21 Au atom cluster (see Figure 2). The adsorbate preferred an hcp-bridged structure (Au-S: 2.75 Å) with a very slight displacement towards the hcp site. The S–C bond (1.85 Å) extends towards the fcc site. The molecule is largely tilted from the surface normal (z-axis) with a tilting angle (of the S–C bond) nearly 60° in all cases. The optimized, adsorbed structures are shown in Figure 2. It is evident from the figure that the molecule is not a perfect zig-zag structure like alkanethiols as reported before [80, 84]. The distortion from the usual structure occurs because the –CONH unit in *N*-(2-mercaptoethyl) benzamide tends to remain coplanar.

We found two conformations for *N*-(2-mercaptoethyl)benzamide molecule on Au(111) surface, viz. 'CO-up' and 'CO-down' structures (for details see Figure 2).

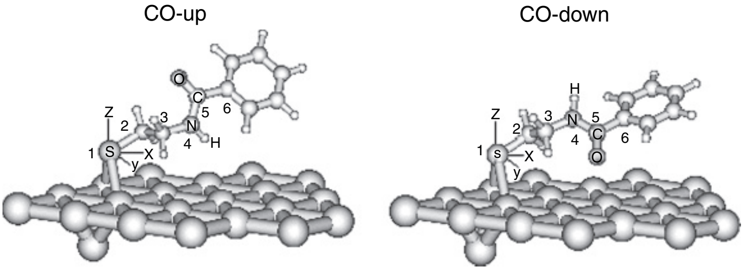


Figure 2 Optimized CO-up and CO-down structures of *N*-(2-mercaptoethyl)benzamide on model Au(111). The dipole moment of the CO-down structure is 4.49*D* (x: −0.84, y: 2.45, z: 3.68, with z-axis as surface normal) and that of the CO-up structure is 3.80*D* (x: 0.23, y: 2.84, z: −2.51, with z-axis as surface normal)

A rotation about the C3–N4 bond adjacent to the CONH unit of adsorbed molecule converts the CO-up conformer to the CO-down. The barrier height for the C–N bond rotation for the adsorbed molecule was roughly estimated to be ~5.3 kcal/mol. The energies of adsorbed CO-up and CO-down structures are nearly the same, with the CO-down structure more stable by merely 1.05 kcal/mol. The directions of dipole for the two conformations are reverse, which indicates that they respond to electric field in opposite manner. The perpendicular height of the adsorbed molecule from the Au(111) surface is calculated for the optimized structures. It is found that the CO-down structure is nearer to the surface in comparison to the CO-up structure and the benzene ring in the CO-down structure tends to be more parallel to the surface due to stronger surface–molecule interaction.

4.2. Bistability and hysteresis

The height of the molecule from the surface is considered as a relevant variable controlling the conductance switching in our study. Molecular height is obtained from the optimized structures. An external electric field is applied along z-axis for adsorbed CO-up and CO-down conformers. Charge transfer between the surface and the molecule was negligible in the concerned electric field strength.

The change in the height of the two conformers under the electric field is shown in Figure 3. For the CO-up structure, there is a rapid increase in height at positive fields above 1 V/nm and then it sharply falls to 0.8 nm at around an electric field of 3 V/nm.

This fall in height results from the fact that the CO-up conformer is converted to the CO-down with a rotation about C–N bond. The increase in the height of the molecule would mean an increase in tunnelling current in STM experiments, which would sharply fall when the height decreases. The CO-up structure remains at nearly the same height at negative fields till −2 V/nm. The optimized structures for the points marked in Figure 3 are displayed separately in Figures 4 and 5, showing how the tilting angle between molecule and the surface and the structure of the adsorbate changes in applied field. This is responsible for the increase in the height of the molecule. In the case of the CO-down conformer, however, the effect of electric field is just the reverse. In this case the height

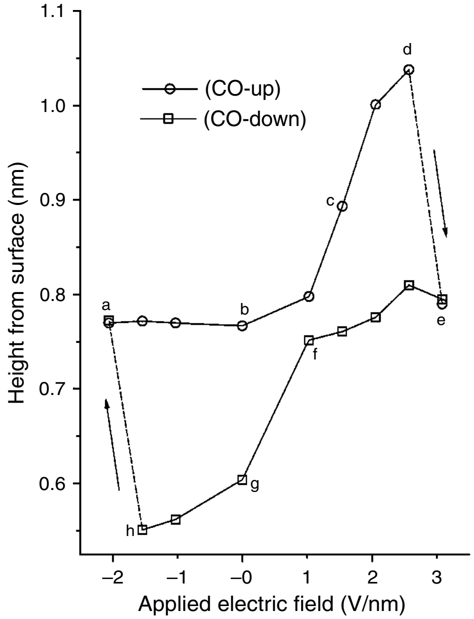


Figure 3 Effect of applied electric field on the molecular heights of CO-up and CO-down conformations of *N*-(2-mercaptoethyl)benzamide adsorbed on a model gold surface

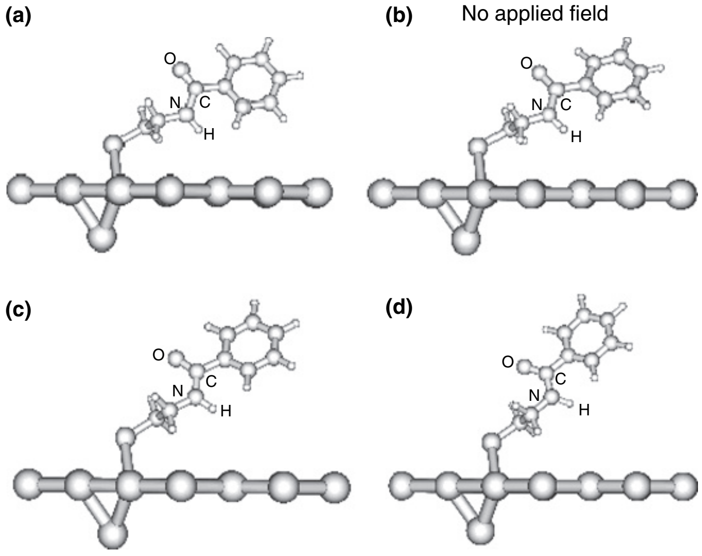


Figure 4 Selected optimized structures for CO-up conformation in applied fields. The points marked by **a, b, c** and **d** in the previous figure are shown here

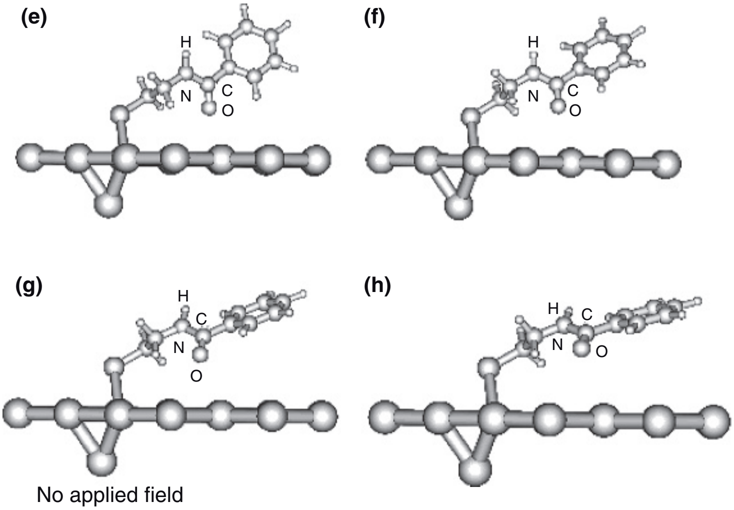


Figure 5 Selected optimized structures for CO-down conformation in applied fields. The points marked by **e, f, g** and **h** in Figure 3 are shown here

increases up to applied field 1 V/nm but remains almost unchanged above this field. At negative fields the height gradually decreases with a minimum at $-1.5\,\mathrm{V/nm}$. Then the molecule flips to the CO-up structure with a sudden jump in height. Hence, at $-2\,\mathrm{V/nm}$ there exists only the CO-up structure and at 3.1 V/nm there exists only the CO-down structure. In any applied fields between these two extremes, both the structures exist.

To understand the conformation changes under the electric field, we examined the change in the z-component of the dipole moment of the molecule under an electric field applied along z-direction. The calculated dipoles of the two conformers are shown in Figure 6. They are both increasing functions of the electric field, with the dipole moment of the CO-down structure larger than that of the CO-up structure. When the CO-up structure flips to the CO-down structure at around 3 V/nm, the dipole moment increases substantially, while the opposite process occurs at negative fields.

When no electric field is applied, the CO-up and the CO-down structures are similar in energy, with CO-down structure being more stable only by 1.05 kcal/mol. The dependence of the energy on the electric field is different for the two structures as shown in Figure 7. The CO-up conformation has a dipole pointing down towards the surface and stabilized more by an electric field in the negative z-direction. In the case of a positive electric field, the energy of the CO-up structure first increases slightly, but soon it starts decreasing, with a sharp fall around 2 V/nm. This drastic change in energy is caused by the flipping of the molecule to the CO-down orientation. In the case of the CO-down conformation the z-component of the dipole moment is positive, so that there is stabilization at positive fields. But the energy increases for fields in the negative z-direction. At about $-1.5\,\mathrm{V/nm}$ there is sudden decrease in energy because the molecule changes its conformation to CO-up. Only the CO-up structure exists below $-2.1\,\mathrm{V/nm}$ and only the CO-down structure exists above 3.1 V/nm. This is essentially a hysteresis phenomenon, where a bistability in the potential energy profile is reduced to a single minimum by the application of a threshold field. The single minimum outside

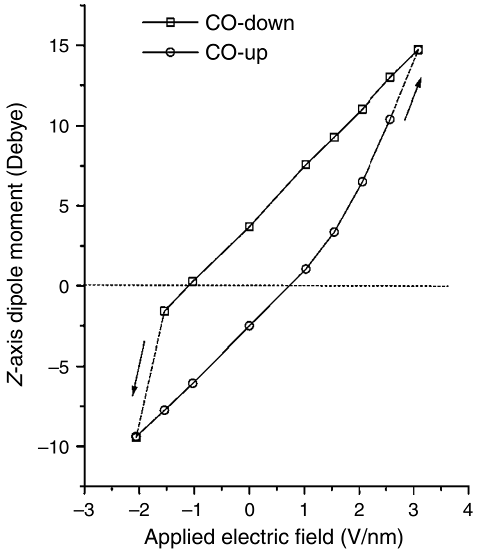


Figure 6 Effect of applied electric field on the z-component of dipole moment for CO-up and CO-down structures on model gold surface

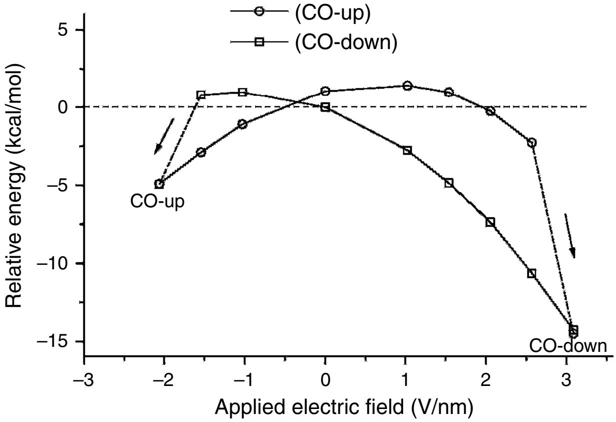


Figure 7 Effect of electric field on the energies of CO-up and CO-down conformations on gold surface. All the energies are relative to the energy of the CO-down structure at no field

the hysteresis region corresponds to the CO-down structure for positive fields and to the CO-up structure for negative fields. The unique hysteresis behavior shown in Figure 3 is a result of delicate balance between the deformation of the molecule and the direction of the dipole. The dipole of the molecule in this case is not along the molecular axis but depends on the orientation of the amide group. The amide moiety plays an important role because of its flexibility and polarity.

The CO-up and CO-down conformers of *N*-(2-mercaptoethyl)benzamide on Au(111) have different molecular heights and different dipole moments. Hence, the two conformations can act as high-conducting 'ON' or low-conducting 'OFF' states, which can be switched by means of an external electric field. In this case, the ON and OFF conformers are almost isoenergetic and the barrier height for the conversion of CO-up and CO-down conformers is not very large (nearly 6 kcal/mol). So there is not much control about the choice of the starting geometry due to thermal equilibrium at room temperature. In spite of the equilibrium between the two there can still be rectification. Figure 3 shows an overall increase in the height of the molecule at positive applied fields, which corresponds to the ON state of the molecule. The molecule remains to be OFF at negative applied fields. This is similar to the rectification of the current discussed by Troisi and Ratner [76]. Furthermore, the ON state is switched OFF by the application of external fields higher than 2.6 V/nm, corresponding to negative differential resistance. This shows that amide molecules can also be used for studying NDR mechanism.

At low temperatures, however, the ON and OFF states can be brought into separate observable states without thermal equilibrium. In this case the full hysteresis curve of Figure 3 can be followed. Alternatively, if the molecule is embedded in a matrix of self-assembled monolayer on the surface, the interactions with surrounding molecules may cause a substantial increase in the barrier height for the conversion. We can also expect a cooperative switching in the case of an ordered monolayer of the switching molecules on the surface. In the next section we relate this study for the model molecule with the conductance switching and rectifying behavior of an adsorbed azobenzamide molecule on Au(111) surface studied under applied bias voltage with STM at room temperatures [46].

5. Conductance switching in a photoisomeric azobenzamide molecule

The *trans–cis* isomerization is possibly one of most well-studied conformation changes caused by photoexcitation. A common example is azobenzene, which undergoes transformation from the more stable *trans* to the less stable *cis* conformation upon UV irradiation. Visible light irradiation or heating may be used for the reverse transformation. It is extremely interesting to study how the photoisomerizable azobenzene unit couples with the applied electric field. Better understanding of this process may allow the control of isomerization processes with applied electric field which may then have many important device applications.

The photoisomerization of *N*-(2-mercaptoethyl)-4-phenylazobenzamide (structure shown in Figure 8 and hereafter denoted as azobenzamide) was observed by Yasuda *et al.* for the first time using STM [46]. The molecule has a polar amide group with a thiol end for chemisorption on gold surface. The molecule embedded in *N*-dodecanethiol (C12) self-assembled monolayer (SAM) films formed on Au(111) substrate was observed by STM. The image of the molecule appeared bright under visible light (325 nm) and became dark under UV irradiation (450 nm). These two states corresponded to the *trans* and *cis* conformations of the azobenzene moiety present in the molecule. Then, the authors studied the effect of electrical excitation caused by STM tip on the individual

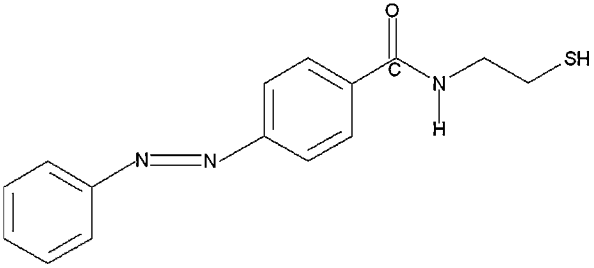


Figure 8 *N*-2-(mercaptoethyl)-4-phenylazobenzamide

azobenzamide molecules in the SAM without photoillumination. Studies were separately conducted under applied bias voltages for the molecules which were tightly packed and also the molecules which were relatively free (adsorbed in etch pits or at phase boundaries of the SAM). It was found that bright spots corresponding to azobenzamide molecules in tightly packed regions did not change with applied bias voltage but the molecules which were relatively free became dark at negative applied fields. The STM images of the same area with sample bias voltages of +1.0 and −1.0 V are shown in Figure 9.

In the case of an unchanged azobenzamide molecule, the I-V characteristics were symmetric for positive and negative voltages. In contrast, a drastic change was observed in the I-V curve measured over the molecule which changed in brightness. Tunneling current was almost flat between −1.0 and +0.5 V and rapid switching in the tunneling current between two I-V curves (high and low current states) was observed in the high positive voltage region. The turn-on voltage for the switching was around 0.5 V. The low current state I-V curve had a shape similar to that obtained for negative voltage, whereas the high current state I-V curve exhibited a characteristic similar to that obtained for the unchanged azobenzamide molecule, which indicated that the unchanged molecule was always in the high current state. The results had clearly shown that the molecules loosely surrounded by alkane thiol molecules changed their conformations between two distinct (high and low current) states during I-V measurement. The high and low current states were attributed to the *trans* and *cis* conformational structures of the azobenzamide molecule, considering the similarity with the photoinduced changes.

When the bias voltage was fixed and the tunneling current was measured, stochastic switching between two definite states at each bias voltage was observed. By analyzing the distribution of the residence time of the flip-flop motion, the lifetimes of the azobenzamide molecule in the two states were obtained for each voltage. The values were scattered over the range of 0.1–1 ms. The lifetimes of the two states showed opposite dependencies on the applied bias voltage. The lifetime of the high current and the low current states became longer and shorter, respectively, with an increase in the applied positive voltage. The low current state was stable in the low bias voltage region; barrier height for the flip-flop motion was anticipated to be high compared to the thermal energy at room temperature. Possible origin of the turn-on voltage for the flip-flop motion was not very clear. However, an interesting point, as the authors admitted, was that the low current state (which was attributed to the *cis* phase) was apparently the ground state not only for the negative bias voltage but also for the positive bias voltage, in contrast to usual knowledge.

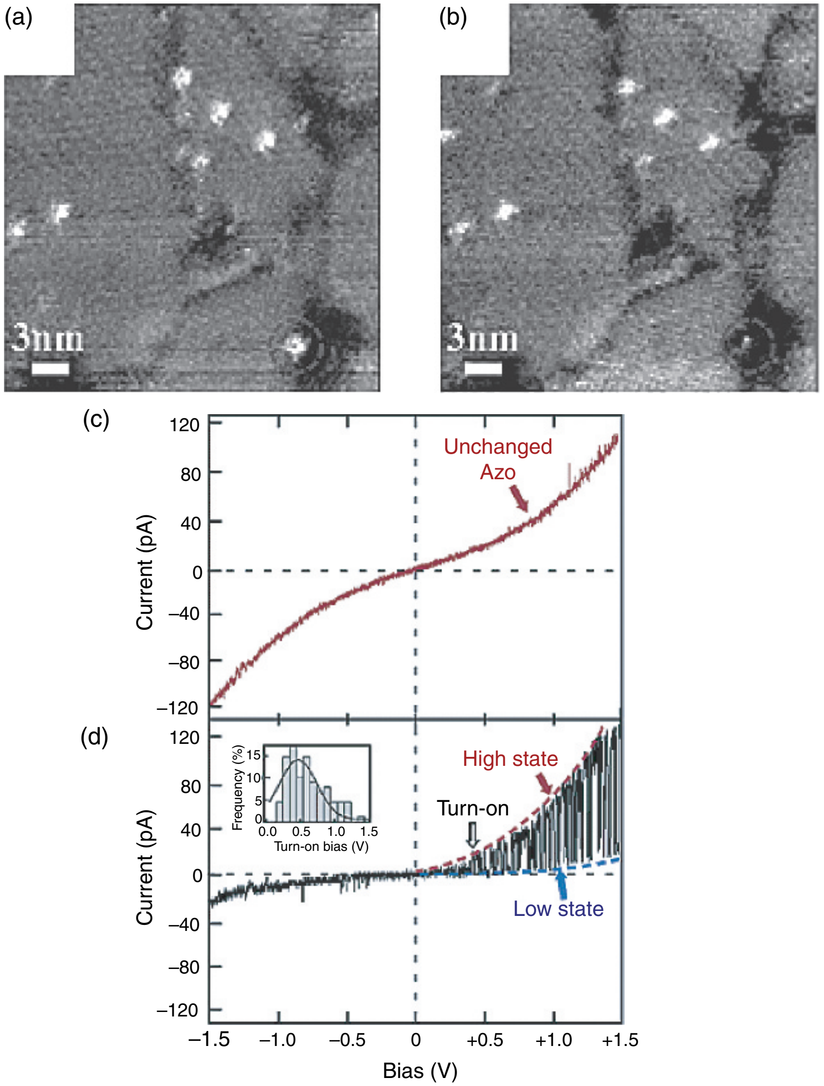


Figure 9 Typical STM images of azobenzamide-embedded C12 SAM film obtained at (a) $V_{\text{sample}} = +1.0\,\text{V}$, $I_{\text{tunnel}} = 10$ pA and (b) $V_{\text{sample}} = -1.0\,\text{V}$, $I_{\text{tunnel}} = 10$ pA. (c) I-V curve obtained over an azobenzamide molecule. (d) I-V curve measured over an azobenzamide molecule which changed in brightness in (b). Histogram of the turn-on voltage for the flip-flop motion is shown in the inset [46], Copyright (2003) American Chemical Society]

6. Possible conformational changes of the azobenzamide molecule

Although the main functional part of the azobenzamide molecule is the azobenzene moiety, it also contains an amide moiety and an alkanethiol moiety. The amide portion has a large polarity and must interact with the electric field strongly. Rotation of the amide unit can have interesting consequences as discussed in previous sections, therefore we studied in details [30] various conformational change processes possible in this molecule by ab initio quantum chemical calculations.

We first considered the *cis–trans* isomerization of the azobenzene moiety. The *cis* and the *trans* structures of free azobenzamide are optimized and the *trans* form is found to be more stable by 15.44 kcal/mol than the *cis* form. The transition state for the *cis–trans* isomerization is located at a barrier height of 37.5 kcal/mol from the *trans* isomer. This value is very similar to the barrier heights found for the *cis–trans* isomerization of the free azobenzene molecule in earlier studies [86–91]. Hence, the introduction of a long amide-substituted alkyl group in the azobenzene molecule did not influence its *cis–trans* isomerization energetics significantly. In order to see whether a reversal of stabilities between the *cis* and *trans* structures is possible for their adsorbed counterparts, we then checked relative stabilities of the various conformations of adsorbed molecule on a model surface using a five-atom Au cluster. It turned out that the *trans* azobenzene form was still the most stable structure on the five-atom Au cluster. The barrier height for conversion of *trans* azobenzene molecule to *cis* azobenzene for adsorbed case is not expected to change very differently from the free molecules as the isomerizable unit is far from the surface. Thus, theoretical studies did not support the assumption that the *cis* azobenzene conformation was the predominant structure under experimental conditions.

We then studied the cis–*trans* isomerization of the amide bond. As normally expected, a conformation with the *cis* amide unit and the *trans* azobenzene unit turned out to be less stable than the trans amide and trans azobenzene conformation by 4.58 kcal/mol. The barrier height for the amide bond to convert from *trans* to *cis* was calculated to be 14.51 kcal/mol for the free molecule. Thus, rapid interconversions at ordinary temperatures are not expected.

Finally we considered a rotation around the C–N bond adjacent to the amide group (rotation about C3–N4 shown in Figure 10). In contrast to the above two processes this barrier height is 5.68 kcal/mol. The result of this rotation is a nearly identical conformation with a change in the orientation of the C=O bond for free molecule. But, these two conformations are different when the molecule is adsorbed on a surface with a reversed orientation of the molecular dipole moment and a difference in molecular height from the surface.

7. Study of the azobenzamide molecule in applied electric field

To test how azobenzamide molecule responds to applied electric fields, we optimized the *trans* structures of the molecule on a five-atom Au cluster modeled as a surface in external electric fields theoretically [30].

7.1. Adsorbed conformations

The *trans–trans* (i.e., *trans*-azobenzene *trans*-amide) isomer of the molecule can exist as four different conformations when adsorbed: C=O and N=N both pointing towards the surface (downward); C=O and N=N both pointing away from the surface (upward); and C=O pointing downward but N=N pointing upward and vice versa. These structures are shown in Figure 10. Conformers **1** and **2** both have a CO-up structure with

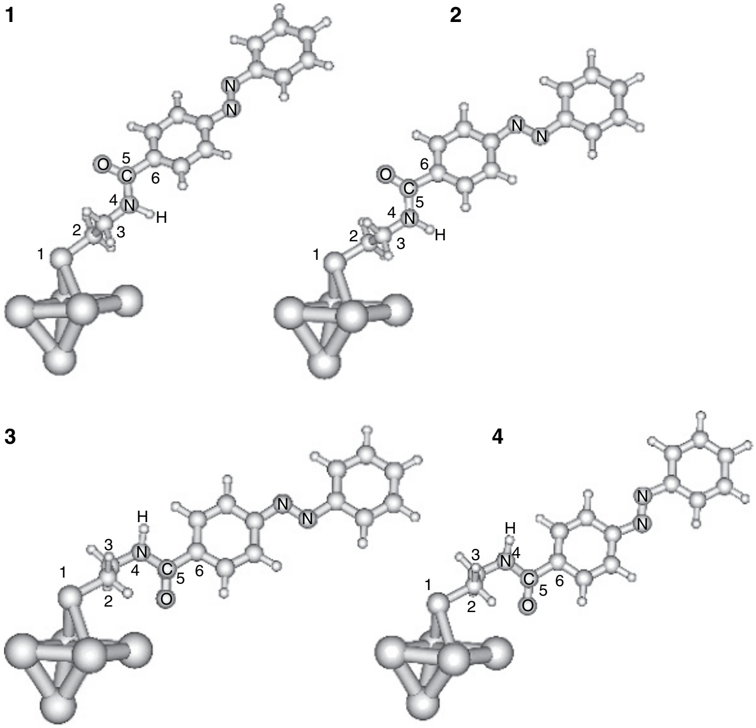


Figure 10 Optimized structures of *N*-(2-mercaptoethyl)-4-phenylazobenzamide on a five-atom gold cluster. **1:** C=O up and N=N up, **2:** C=O up and N=N down, **3:** C=O down and N=N down, **4:** C=O down and N=N up

the N=N bond pointing up and down respectively. Conformers **3** and **4** both have CO-down structures with N=N bond pointing down and up, respectively. The CO-down conformers (**3** and **4**) are slightly higher in energy (by 0.5 kcal/mol) than the CO-up conformers (**1** and **2**). Structure **3** has the smallest and structure **1** has the largest value of molecular heights from the surface amongst all the four conformers. Since all the four structures have similar energies, they can exist with almost equal probabilities on Au(111) at room temperature, if other interactions from the surroundings are not taken into account. Rotation about C3–N4 (see Figure 10) can allow interconversions between **1** and **3** or between **2** and **4**.

7.2. Electric-field-dependent conformation changes

All the four *trans*-amide, *trans*-azobenzene conformers were studied theoretically in the presence of an external electric field applied along *z*-axis. The dependence of the molecular height from the surface upon the electric field is shown in Figure 11. The heights of the CO-up structures (**1** and **2**) increase rapidly with the field applied in the positive *z*-direction. This is a result of the decrease in the tilt angle of the molecule. The CO-down structures (**3** and **4**) exhibit little change in height due to their dipole

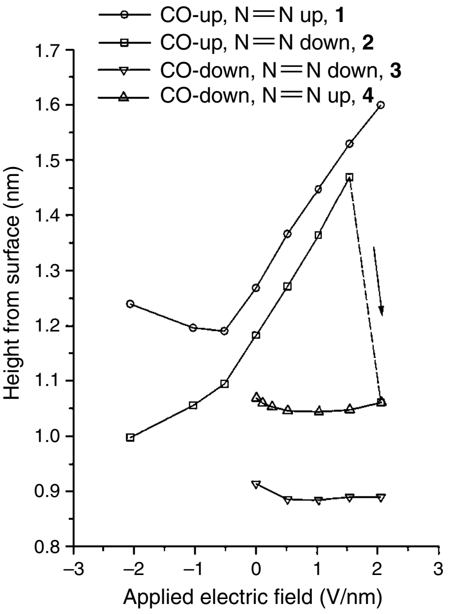


Figure 11 Effect of applied electric field on the height of adsorbed *N*-(2-mercaptoethyl)-4-phenylazobenzamide on the five-atom gold cluster

alignment for positive fields. The situation for this molecule is only partially similar to the case of *N*-(2-mercaptoethyl)benzamide, because now the dipole moment of the molecule not only is due to the CONH unit but also affected by the azobenzene unit. Therefore the two CO-up structures **1** and **2** behave quite differently under an external field. In the case of **2**, the CO-up and N=N down conformer, the molecule at first rises in height and then flips to **4** (CO-down, N=N up) after application of a field of ~1.5 V/nm. But in the case of **1**, the CO-up and N=N up conformer, the molecular height continues rising steadily even above ~2 V/nm, and the molecule becomes more and more vertical to the surface. We further checked the structure of **1** at external fields as strong as 4 V/nm and found that it did not flip and the structure became almost perpendicular to the surface. When conformer **2** (CO-up, N=N down) flips to **4** (CO-down, N=N up), there is a sudden fall in the height of the molecule from 1.47 to 1.07 nm. This can result in a huge change in conductance for STM experiments. The flipping is also accompanied by a sudden change in the dipole moment from 6.96*D* (z-component: 5.37*D*) to 12.36*D* (z-component: 10.88*D*). So in the positive field regime above ~1.5 V/nm there can be only three conformers present: **1**, **3** and **4**. Among these three structures, **1** has a much greater height from surface than the other two structures (see Figure 11). Hence **1** is likely to be much more conducting than **3** and **4**.

When the field is applied in the negative z-direction, the molecular height varies very slowly and the molecule tends to lie down on the surface, in analogy with the case of *N*-(2-mercaptoethyl)benzamide shown in Figure 3. In case of **1** the molecular height increases slightly in negative fields. Under negative fields, the conformers **3** and **4** are expected to be low in height and hence of low conducting type though we could not study **3** and **4** in negative fields due to strong oxygen–gold interactions arising from the

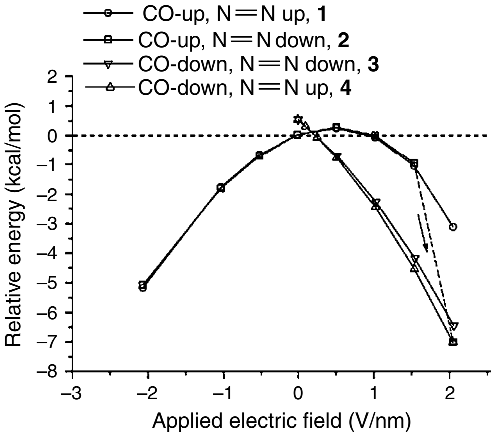


Figure 12 Effect of applied electric field on the relative energies of different conformations of *N*-(2-mercaptoethyl)-4-phenylazobenzamide on the five-atom gold cluster. All energies are relative to the most stable isomer **1**

small size of the metal cluster. A close inspection of the energy changes (see Figure 12) for all the four conformations reveals that the CO-down structures **3** and **4** behave in a similar fashion. The CO-up structures **1** and **2** also behave in a similar manner. At first the energy increase slightly with application of positive field but soon starts decreasing rapidly. At about 2 V/nm the energy of **2** suddenly falls sharply in contrast to **1** and it is converted to **4**.

7.3. Possible effects of intermolecular coupling

The existence of the high and low conformations (in terms of molecular heights) at positive fields shown in this theoretical study [30] is consistent with the stochastic switching between high and low conducting states observed in the STM image of azobenzamide molecules on Au(111) under positive sample bias voltages at room temperatures [46]. Therefore we believe that the rotation about the bond between the amide unit and the alkyl chain may be the dominant mechanism of the observed conductance switching. The *cis–trans* isomerization in the azobenzene unit may occur at much higher applied voltages, because the barrier height for this process is much higher.

However, if we look at the details of the experimental results, there are many issues to be resolved. First of all, the experiment indicates that the OFF state is more stable than the ON state without an electric field, as mentioned before. The observed lifetime is a decreasing function of the bias voltage for the low current state and an increasing function for the high current state. At small bias voltages, the low current state has a longer lifetime than the high current state. Therefore, if we consider these lifetimes as thermal transition rates, then the low current state must be the ground state and the high current state is a metastable state.

We considered theoretically an isolated molecule on a metal cluster, and thus the energetics of the switching states observed in the experiment [46] could not be captured.

The small barrier height for the rotation of C–N bond would imply that CO-up and CO-down conformations coexist at room temperature as mentioned before, and the conductance would be an average of thermally fluctuating high and low conducting states. However, the observed lifetime of switching was rather long, on the order of 1 ms [46]. This suggests that the barrier height is strongly enhanced by the effect of surrounding molecules. The importance of this effect is evident also from the fact that the switching phenomenon was observed only for molecules located in domain boundaries but not inside the domains of the SAM.

The relative energies of the various conformations shown in Figure 12 are expected to be affected by the intermolecular interaction. The structures with a smaller molecular height occupy a larger surface area, which requires more energy due to repulsive interaction with surrounding molecules. Thus, the upright conformations (**1** and **2**, in Figure 10) may become more stable than the relatively flat conformations (**3** and **4**) in the molecular matrix. Theoretical methods designed to take into account the intermolecular interactions on surfaces are likely to give results which reproduce the experimental results.

8. Other interesting molecules for switching

Amide molecules are very common in nature; all the protein molecules are actually amide-linked polymers. It is important to understand the effect of electric field on such amide molecules. The study of Kitagawa and co-workers had shown that such long chain helical molecules can also be used for conductance switching if they can be immobilized on metal surfaces [59]. However, due to intramolecular hydrogen bonding, the behavior of these interlinked amide groups is less obvious. In this case the applied electric field must break the hydrogen bonds and then reorient the dipole associated with the molecule. Here we mention about another molecule, tribenzyl-1-aza-adamantanetrione (shown in Figure 13). It is C_3 symmetric and is known to form self-assembled structures due to intramolecular forces [92]. This molecule has three amide groups arranged in a C_3 symmetric way. Application of electric field along the symmetry axis of the free molecule can cause the amide bonds to flip. Due to the presence of three amide bonds, the applied voltage necessary for such flipping process is expected to be higher than that required for benzamide molecules.

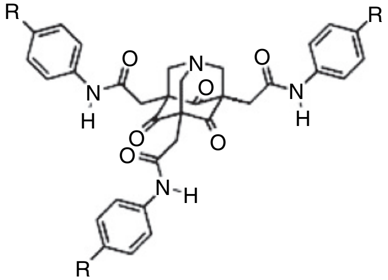


Figure 13 Tribenzyl-1-aza-adamantanetrione [92]

9. Concluding remarks

Amide molecules are an interesting class of flexible molecules, as they can undergo electric field–induced conformational switching processes. Our study shows that such switching may trigger a drastic conductance change, if the molecule is anchored to a surface. We have discussed this type of switching for amide molecules on a metal surface considering two examples. In the first case, we show the possibility of electric-field-dependent conformation switching in an adsorbed benzamide molecule. The conformational change involves the rotation of the amide unit in the molecule, which reverses the direction of the dipole and changes the height of the molecule. The conductance of the system is expected to be affected by this switching. Despite the small barrier height for interconversion between the conformers it shows a switching behavior as there is a net increase in conductance at positive applied electric fields which abruptly falls after a threshold field. This corresponds to NDR mechanism. The conductance change follows a hysteresis curve. Therefore one can think of a memory device consisting of a SAM of the switching molecules sandwiched between electrodes, where conformation switching of can be induced by applying a sufficient voltage. In the next case we have studied electric field–induced conformational changes in *N*-(2-mercaptoethyl)-4-phenylazobenzamide molecule on a gold cluster. We have qualitatively explained the voltage-induced phase switching of this molecule on Au(111) reported by Yasuda *et al.* [46] in terms of the rotation of the amide unit in the adsorbate molecule which may occur at room temperatures. The comparison of the theoretical results with the experiments also suggests that the kinetics of switching is likely to be strongly influenced by surrounding molecules in SAM. Finally, we have suggested a few model molecules with amide moieties where such kind of switching may occur.

Acknowledgements

The authors thank Dr. Tohru Nakamura and Dr. Satoshi Yasuda for many useful discussions. This work was partly supported by NEDO Nanotechnology Program (Synthetic Nano-Function Materials Project) and by MEXT Grant-in-Aid for Science Research (Priority Area: Super-Hierarchical Structures).

References

[1] A. H. Flood, J. F. Stoddart, D. W. Steuerman, J. R. Heath, *Science*, 306 (2004) 2055.
[2] R. L. Carroll, C. B. Gorman, Angew, *Chem. Int. Ed.*, 41 (2002) 4378.
[3] B. A. Mantooth, P. A. Weiss, *Proc. IEEE*, 91 (2003) 1785.
[4] X. Y. Zhu, *Surf. Sc. Rep.*, 56 (2004) 1.
[5] A. Nitzan, M. A. Ratner, *Science*, 300 (2003) 1384.
[6] D. L. Pearson, J. M. Tour, *J. Org. Chem.*, 62 (1997) 1376.
[7] L. Jones, J. S. Schumm, J. M. Tour, *J. Org. Chem.*, 62 (1997) 1388.
[8] E. A. Weiss, M. R.Wasielewski, M. A. Ratner, *Topics in Current Chemistry*, 257, (2005) 103.
[9] I. W. Tam, J. Yan, R. Breslow, *Org. Lett.*, 8 (2006) 183.
[10] A. Aviram, M. Ratner, *Chem. Phys. Lett.*, 29 (1974) 277.

[11] R. M. Metzer, B. Chen, U. Hopfner, M. V. Lakshmikantham, D. Vuillaume, T. Kawai, X. L. Wu, H. Tachibana, T. V. Huges, H. Sakurai, J. W. Baldwin, C. Hosch, M. P. Cava, L. Brehmer, G. J. Ashwell, *J. Am. Chem. Soc.*, 119 (1997) 10455.

[12] R. M. Metzger, *Chem. Rev.*, 103 (2003) 3803.

[13] G. J. Ashwell, W. D. Tyrrell, A. J. Whittam, *J. Am. Chem. Soc.*, 126 (2004) 7102.

[14] H. Park, J. Park, A. K. L. Lim, E. H. Anderson, A. P. Alivisatos, P. L. McEuen, *Nature*, 407 (2000) 57.

[15] C. Joachim, J. K. Gimzewski, *Chem. Phys. Lett.*, 265 (1997) 353.

[16] J-O Lee, G. Lientschnig, F. Wiertz, M. Struijk, R. A. J. Janssen, R. Egberink, D. N. Reinhoudt, P. Hadley, C. Dekker, *Nano Lett*. 3 (2003) 113.

[17] B. L. Feringa, *Molecular Switches*, Weily-VCH, Weinheim, 2001.

[18] Y. Liu, A. H. Flood, R. M. Moskowitz, J. F. Stoddart, *Chem. Eur. J.* 11 (2004) 369.

[19] A. R. Pease, J. O. Jeppesen, J. F. Stoddart, Y. Luo, C. P. Collier, J. R. Heath, *Acc. Chem. Res.*, 34 (2001) 433.

[20] T. Rueckes, K. Kim, E. Joselevich, G. Y. Tseng, C. L. Cheung, C. M. lieber, *Science*, 289 (2000) 94.

[21] Y. Luo, C. P. Collier, J. O. Jeppesen, K. A. Nielsen, E. Delonno, G. Ho, J. Perkins, H.-R. Tseng, T. Yamamoto, J. F. Stoddart, J. R. Heath, *Chem. Phys. Chem.*, 3 (2002) 519.

[22] C. P. Collier, J. O Jeppesen, Y. Luo, J. Perkins, E. W. Yong, J. R. Heath, J. F. Stoddart, *J. Am. Chem. Soc.*, 123 (2001) 12632.

[23] M. A. Reed, C. Zhou, C. J. Muller, T. P. Burgin, J. M. Tour, *Science*, 278 (1997) 252.

[24] J. Reichert, R. Ochs, D. Beckmann, H. B. Weber, M. Mayor, H. von Lohneysen, *Phys. Rev. Lett.*, 88 (2002) 176804.

[25] J. Reichert, H. B. Weber, M. Mayor, H. von Lohneysen, *Appl. Phys. Lett.*, 82 (2003) 4137.

[26] H. Park, A. K. L. Lim, A. P. Alivisatos, J. Park, P. L. McEuen, *Appl. Phys. Lett.*, 75 (1999) 301.

[27] J. Park, A. N. Pasupathy, J. I. Goldsmith, C. Chang, Y. Yaish, J. R. Petta, M. Rinkoski, J. P. Sethna, H. D. Abruna, P. L. McEuen, D. C. Ralph, *Nature*, 417 (2002) 722.

[28] Y. Otsuka, Y. Naitoh, T. Matsumoto, W. Mizutani, H. Tabata, T. Kawai, *Nanotechnology*, 15 (2004) 1639.

[29] J. G. Kushmerick, D. B. Holt, J. C. Yang, J. Naciri, M. H. Moore, R. Shashidhar, *Phys. Rev. Lett.*, 89 (2002) 86802.

[30] B. Das, S. Abe, *J. Phys. Chem. B*, 110 (2006) 4247.

[31] B. L. Feringa, *Acc. Chem. Res.*, 34 (2001) 504.

[32] F. Pina, J. C. Lima, A. J. Parola, C. A. M. Afonso, Angew. *Chem. Intl. Ed.*, 43 (2004) 1525.

[33] W. Hu, H. Nakashima, K. Furukawa, Y. Kashimura, K. Ajito, Y. Liu, D. Zhu, K. Torimitsu, *J. Am. Chem. Soc.*, 127 (2005) 2804.

[34] G. S. Kumar, G. C. Neckers, *Chem. Rev.*, 89 (1989) 1915.

[35] T. Hugel, N. B. Holland, A. Cattani, L. Moroder, M. Seitz, H. E. Gaub, *Science*, 296 (2002) 1103.

[36] F. Kulzer, M. Orrit, *Annu. Rev. Phys. Chem*, 55 (2004) 585.

[37] M. Heilemann. E. Margeat, R. Kasper, M. Sauer, P. Tinnefeld, *J. Am. Chem. Soc.*, 127 (2005) 3801.

[38] M. Irie, T. Fukaminato, T. Sasaki, N. Tamai, T. Kawai, *Nature*, 420 (2002) 759.

[39] T. Fukaminato, T. Sasaki, T. Kawai, N. Tamai, M. Irie, *J. Am. Chem. Soc.*, 126 (2004) 14843.

[40] N. N. P. Moonen, A. H. Flood, J. M. Fernández, J. F. Stoddart, *Topics in Current Chemistry*, 262 (2005) 99.

[41] P. M. Mendes, A. H. Flood, J. F. Stoddart, *Appl. Phys. A*, 80 (2005) 1197.

[42] S. Garandeé, S. Silvi, M. Venturi, A. Credi, A. H. Flood, J. F. Stoddart, *Chem. Phys. Chem.*, 6 (2005) 2145.

[43] D. W. Steuerman, H.-R. Tseng, A. J. Peters, A. H. Flood, J. O. Jeppesen, K. A. Nielsen, J. F. Stoddart, J. R. Heath, Angew. *Chem. Int. Ed.*, 43 (2004) 6486.

[44] L. A. Bumm, J. J. Arnold, M. T. Cygan, T. D. Dunbar, T. P. Burgin, L. Jones II, D. L. Allara, J. M. Tour and P. S. Weiss, *Science*, 271 (1996) 1705.

[45] M. Dorogi, J. Gomez, R. Osifchin, R. P. Andres, R. Reifenberger, *Phys. Rev. B*, 552 (1999) 9071.

[46] S. Yasuda, T. Nakamura, M. Matsumoto, H. Shigekawa, *J. Am. Chem. Soc.*, 125 (2003) 16430.

[47] J. Chen, M. A. Reed, A. M. Rawlett, J. M. Tour, *Science*, 286 (1999) 1550.

[48] J. M. Tour, A. M. Rawlett, M. Kozaki, Y. Yao, R. C. Jagessar, S. M. Dirk, D. W. Price, M. A. Reed, C-W. Zhou, J. Chen, W. Wang, I. Campbell, *Chem. Eur. J.*, 7 (2001) 5118.

[49] J. Chen, W. Wang, M. A. Reed, A. M. Rawlett, D. W. Price, J. M. Tour, *Appl. Phys. Lett.*, 77 (2002) 1224.

[50] A. M. Rawlett, T. J. Hopson, L. A. Nagahara, R. K. Tsui, G. K. Ramachandran, S. M. Lindsay, *Appl. Phys. Lett.*, 81 (2002) 3043.

[51] A. M. Rawlett, T. J. Hopson, I. Amlani, R. Zhang, J. Tresek, L. A. Nagahara, R. K. Tsui, H. Goronkin, *Nanotechnology*, 14 (2003) 377.

[52] J. M. Seminario, P. A. Derosa, *J. Am. Chem. Soc.*, 123 (2001) 12418.

[53] P. A. Lewis, C. E. Inman, Y. X. Yao, J. M. Tour, J. E. Hutchison, P. S. Weiss, *J. Am. Chem. Soc.*, 127 (2005) 17421.

[54] A. A. Dameron, J. W. Ciszek, J. M. Tour, P. S. Weiss, *J. Phys. Chem. B*, 108 (2004) 16761.

[55] J. Chen, J. Su, W. Wang, M. A. Reed, *Physica E*, 16 (2003) 17.

[56] A. S. Blum, J. G. Kushmerick, D. P. Long, C. H. Patterson, J. C. Yang, J. C. Henderson, Y. Yao, J. M. Tour, R. Sashidhar, B. R. Ratna, *Nature Materials*, 4 (2005) 167.

[57] T. G. Gopakumar, F. Muller, M. Hietschold, *J. Phys. Chem. B*, 110 (2006) 6051.

[58] H. Yanagi, K. Ikuta, H. Mukai, T. Shibutani, *Nano Lett.*, 2 (2002) 951.

[59] K. Kitagawa, T. Morita, S. Kimura, Angew. *Chem. Int. Ed.*, 44 (2005) 6330.

[60] R. A. Wassel, R. R. Fuierer, N. Kim, C. B. Gorman, *Nano Lett.*, 3 (2003) 1617.

[61] Z. J. Donhauser, B. A. Mantooth, T. P. Pearl, K. F. Kelly, S. U. Nanayakkara, P. S. Weiss, *Jpn. J. Appl. Phys.*, 41 (2002) 4871.

[62] Z. J. Donhauser, B. A. Mantooth, K. F. Kelly, L. A. Bumm, J. D. Monnell, J. J. Stapleton, D. W. Price (Jr), A. M. Rawlett, D. L. Allara, J. M. Tour, P. S. Weiss, *Science*, 292 (2001) 2303.

[63] A. M. Moore, A. A. Dameron, B. A. Mantooth, R. K. Smith, D. J. Fuchs, J. W. Ciszek, F. Maya, Y. Yao, J. M. Tour, P. S. Weiss, *J. Am. Chem. Soc* 128 (2006) 1959.

[64] J. M. Seminario, A. G. Zacarias, J. M. Tour, *J. Am. Chem. Soc.*, 122 (2000) 3015.

[65] M. D. Ventra, S. G. Kim, S. T. Pantelides, N. D. Lang, *Phys. Rev. Lett.*, 86 (2001) 288.

[66] J. Cornil, Y. Karzazi J. L. Bre'das, *J. Am. Chem. Soc.*, 124 (2002) 3516.

[67] N. D. Lang, P. Avouris, *Phys. Rev. B*, 62 (2000) 7325.

[68] J. M. Seminario, A. G. Zacarias, J. M. Tour, *J. Am. Chem. Soc.*, 120 (1998) 3970.

[69] G. K. Ramachandran, T. J. Hopson, A. M. Rawlett, L. A. Nagahara, A. Primak, S. M. Lindsay, *Science*, 300 (2003) 1413.

[70] T. Ishida, H. Fukushima, T. Tamaki, H. Tokumoto, *Jpn. J. Appl. Phys.*, 42 (2003) 5342.

[71] J. M.Seminario, P. A. Derosa, L. E. Cordova, B. H. Bozard, *IEEE Trans. Nanotech.*, 3 (2004) 215.

[72] J. M. Seminario, L. E. Córdova, P.A. Derosa, *Proc. IEEE*, 2003, 91(2003) 1958.

[73] J. Taylor, M. Brandbyge, K. Stokbro, *Phys. Rev. B*, 68 (2003) 121101.

[74] C. Zhang, M.-H. Du, H.-P. Cheng, X.-G. Zhang, A. E. Roitberg, J. L. Krause, *Phys. Rev. Lett*, 92 (2004) 158301.

[75] J. Li, G. Speyer, O. F. Sankey, *Phys. Rev. Lett.*, 93 (2004) 248302.

[76] A. Torisi, M. Ratner, *J. Am. Chem. Soc.*, 124 (2002) 14528.

[77] E. E. Nikitin, *Theory of Elemetary Atomic and Molecular Processes in Gases*, Clarendon Press, Oxford 1974.
[78] B. Das, K. L. Sebastian, *Chem. Phys. Lett.*, 330 (2000) 433.
[79] Y. Akinaga, T. Nakajima, K. Hirao, *J. Chem. Phys.*, 114 (2001) 8555.
[80] Y. Cao, Q. Ge, D. J. Dyer, L. Wang, *J. Phys. Chem. B*, 107 (2003) 3803.
[81] J. Nara, S. Higai, Y. Morikawa, T. Ohno, *J. Chem. Phys.*, 120 (2004) 6705.
[82] T. Hayashi, Y. Morikawa, H. Nozoye, *J. Chem. Phys.*, 114 (2001) 7615.
[83] R. D. Felice, A. Selloni, E. Molinari, *J. Phys. Chem. B*, 107 (2003) 1151.
[84] P. Cyganik, M. Buck, W. Azzam, C. Woll, *J. Phys. Chem. B*, 108 (2004) 4989.
[85] S. S. Jang, Y. H. Jang, Y.-H. Kim, W. A. Goddard III, A. H. Flood, B. W. Laursen, H.-R. Tseng, J. F. Stoddart, J. O. Jeppesen, J. W. Choi, D. W. Steuerman, E. Delonno, J. R. Heath, *J. Am. Chem. Soc.*, 127 (2005) 1563.
[86] R. Cimiragila, H.-J. Hofmann, *Chem. Phys. Lett.*, 217 (1994) 430.
[87] C. Angeli, R. Cimiragila, H.-J. Hofmann, *Chem. Phys. Lett.*, 259 (1999) 276.
[88] P. Cattaneo, M. Persico, *Phys. Chem. Chem. Phys.*, 1 (1999) 4739.
[89] T. Ishikawa, T. Noro, *J. Chem. Phys.*, 115 (2001) 7503.
[90] L. Gagliardi, G. Orlandi, F. Bernardi, A. Cembran, M. Garavelli, *Theo. Chem. Acc.*, 111 (2004) 363.
[91] A. Cembran, F. Bernardi, M. Garavelli, L. Gagliardi, G. Orlandi, *J. Am. Chem. Soc.*, 126 (2004) 3234.
[92] A. J. Lampkins, O. Abdul-Rahim, H. Li, R. K. Castellano, *Org. Lett.*, 7 (2005) 4471.

Molecular and Nano Electronics: Analysis, Design and Simulation
J. M. Seminario (Editor)

Chapter 6

Semi-empirical simulations of carbon nanotube properties under electronic perturbations

Yan Li[a] and Umberto Ravaioli[b]

[a] *Department of Chemistry, University of California, Davis, California, 95616, USA*
[b] *Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, Illinois, 61801, USA. ravaioli@uiuc.edu*

1. Introduction

Carbon nanotubes (CNTs) are tube-shaped nanostructures consisting of coaxial graphene cylinders, which usually measure a few nanometers in diameter and up to microns in length. Since their discovery in the early 1990s [1–3], CNTs have attracted enormous attention from researchers all over the world. The specific geometry as well as the unique electrical, mechanical, thermal, and chemical properties have made the CNT an ideal test ground for fundamental theories in low-dimensional nanostructures, and also a promising candidate for applications in the fast-developing field of nanotechnology and molecular electronics [4].

The past decade has seen considerable progress in the experimental aspects of CNTs, from synthesis, growth, separation, and analysis of the material to fabrication and control of CNT-based devices. Field emitters, field effect transistors, nano-electromechanical systems, AFM/STM probes, and biosensors are just a few examples of CNT applications successfully demonstrated in the lab. On the other hand, theoretical modeling and simulations have been a valuable complement of experiments to help analyze and explain experimental results as well as predict the behavior of the material or device under study. Most importantly, with the variety of well-developed numerical techniques and powerful computational platforms available today, it is now realistic to conceive and design novel CNT-based molecular devices from simulation and to predict numerically their performance as well as propose solutions for optimization, which could be increasingly important to guide experiments.

Depending on the complexity of the system and the physical properties of interest, one can employ theoretical approaches at different approximation levels to study CNT-based system. Here we focus on a practical, self-consistent tight-binding (TB) model to investigate the electronic properties of CNTs under external electronic perturbations, both in the infinitely long limit and in the finite-size limit. Despite its simplicity, the TB approach may include as many important physical details as do more sophisticated models with the right choice of empirical parameters.

Moreover, the physical interpretation of a TB model is amenable to intuitive connection with the physics, while its simple algorithm enables simulating systems of considerable size, which would be inaccessible for more advanced methods such as density function theory (DFT). In some situations, a multi-level approach combining methods at different approximation levels proves to be an efficient and accurate way to model the system [5].

This chapter is organized as follows. In Section 2, we briefly review the basics of CNTs and describe the self-consistent TB formalism. Next, we apply the TB model to investigate the electronic properties of CNTs, both in an infinite periodic system and in a finite-size system. In Section 3, we discuss the possibility of metal–semiconductor transitions (MSTs) in metallic nanotubes under angular perturbations. With the aid of group theory techniques and the analytical power of the TB derivation, we provide selection rules for subband coupling and estimate the magnitude of band gap openings as well as the Fermi velocity renormalization near the Fermi level. We also suggest an effective mechanism to enhance the MST by a combination of different forms of perturbations. Then, in Section 4, we study the finite-size effect on the structural and electronic properties of carbon nanotubes. By combining first principle calculations with classical molecular dynamics simulations, our model allows us to study the transport behavior of a water molecule or of an ion interacting with a short nanotube segment. We demonstrate the importance of the nanotube polarization effect and atomic partial charges in determining the energetics of the system, which may facilitate understanding and controlling the electronic behavior of carbon nanotubes in biological applications.

For simplicity, we only consider single-walled carbon nanotubes (SWNTs) in this chapter and mostly focus on armchair SWNTs (A-SWNTs), which possess the highest geometrical symmetry. Some conclusions can be easily extended to chiral SWNTs, e.g., through a more general $\mathbf{k}\cdot\mathbf{p}$ description of the electronic states, while for some other issues high order correction terms need to be incorporated to account for the chiral dependence [6].

2. Background

2.1. Geometry of CNTs

A single-walled nanotube can be described as a graphene sheet rolled up into a seamless cylinder along a certain direction defined by the chiral vector $\mathbf{C}_\mathrm{h}$, which can be decomposed into the unit vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ of a hexagonal lattice as shown in Figure 1:

$$\mathbf{C}_\mathrm{h} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 \text{ or } (n_1, n_2), \quad \text{with } n_1, n_2 \text{ being integers} \tag{1}$$

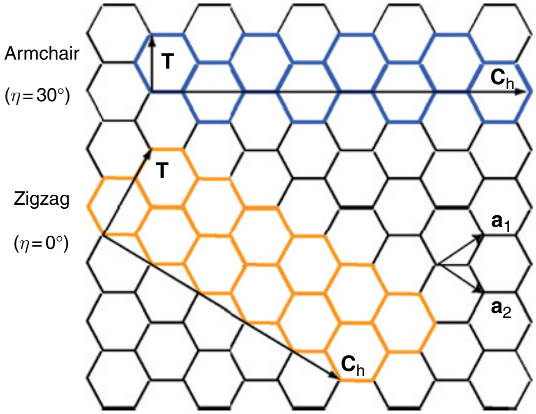


Figure 1 The unrolled graphene lattice and the unit cells of (4, 4) armchair and (5, 0) zigzag nanotubes. The chiral vector $\mathbf{C}_\mathrm{h}$ and translational vector $\mathbf{T}$ are also shown

Here, $a = |\mathbf{a}_{1,2}| = \sqrt{3}r_{\mathrm{CC}}$ is the lattice constant of the two-dimensional (2D) graphene, with $r_{\mathrm{CC}} = 1.42\,\text{Å}$, the C–C bond length [7]. The angle between $\mathbf{C}_\mathrm{h}$ and $\mathbf{a}_2$ is defined as the chiral angle η, where $0° \le \eta \le 30°$. The translational vector $\mathbf{T}$ is parallel to the tube axis and is normal to $\mathbf{C}_\mathrm{h}$. The rectangle $\mathbf{C}_\mathrm{h} \times \mathbf{T}$ defines the unit cell for the nanotube, as shown in Figure 1 for armchair and zigzag SWNTS, which have $\eta = 30°$ and $\eta = 0°$, respectively.

2.2. Tight-binding description

The electronic band structure of CNTs can be derived from that of the unrolled graphene sheet by using a zone-folding scheme [7]. For a 2D graphene layer, each carbon atom forms three in-plane σ-bonds hybridized in sp^2 configuration, while the 2p$_z$ orbital stands perpendicular to the plane and forms a covalent π-bond. The low energy (relative to the Fermi level) properties of 2D graphene are dominated by the delocalized π-electrons, and the electronic structure can be described by a single π-orbital Hamiltonian in the formalism of second-quantization as

$$H_0 = \sum_i \varepsilon_i c_i^+ c_i + \sum_{\langle i,j\rangle} \gamma_{ij} c_i^+ c_j \tag{2}$$

where c_i^+ and c_i are the creation and annihilation operators of the π-electron on the i-th atomic site, respectively. ε_i is the unperturbed onsite energy of π-orbitals in the periodic crystal potential, which is usually set to zero. γ_{ij} is the electron hopping integral (also called the "transfer integral") between atoms i and j, and the summation runs over all pairs of neighboring atoms within the proper cutoff of neighboring distance. Eigenstates of the Hamiltonian in Eq. (2) can be constructed from a linear combination of Bloch wave functions on the two non-equivalent sublattices of 2D graphene (denoted as A and B):

$$\begin{aligned}\Psi_{\mathrm{graphene}}(\mathbf{k},\mathbf{r}) &= C_\mathrm{A}(\mathbf{k})\Psi_\mathrm{A}(\mathbf{k},\mathbf{r}) + C_\mathrm{B}(\mathbf{k})\Psi_\mathrm{B}(\mathbf{k},\mathbf{r})\\ \Psi_\xi(\mathbf{k},\mathbf{r}) &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N} \mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}_{i\xi}}\varphi(\mathbf{r}-\mathbf{r}_{i\xi}), \quad (\xi = \mathrm{A}, \mathrm{B})\end{aligned} \tag{3}$$

where C_A and C_B are coefficients to be determined, $\varphi(\mathbf{r}-\mathbf{r}_{i\xi})$ is the atomic-like orbital centered at $\mathbf{r}_{i\xi}$ and i is summed over all A or B atoms. Below we take the first nearest-neighbor (NN) approximation and set $\gamma_{ij} \equiv \gamma_0 = -2.5\,\mathrm{eV}$ [7]. The overlap integral, $s \equiv \langle \varphi(\mathbf{r}-\mathbf{r}_{i\xi}) | \varphi(\mathbf{r}-\mathbf{r}_{j\xi}) \rangle$, is assumed to be zero for first-NNs. In the next section, we will show that the band structure near the Fermi level is modified only slightly when one includes the non-orthogonality of atomic orbitals ($s \neq 0$) or electron hopping integrals beyond the first-NN approximation. By substituting Eq. (3) into Eq. (2), and solving for $H_0\Psi(\mathbf{k},\mathbf{r}) = E(\mathbf{k})\Psi(\mathbf{k},\mathbf{r})$, one obtains the eigen coefficients as

$$\begin{pmatrix} C_A(\mathbf{k}) \\ C_B(\mathbf{k}) \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \pm f(\mathbf{k})/\,|f(k)| \\ 1 \end{pmatrix} \qquad (4)$$

$$f(\mathbf{k}) = \sum_{\lambda=1}^{3} \mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}_\lambda}$$

$$\begin{aligned} E_{\text{graphene}}(\mathbf{k}) &= \pm\gamma_0\,|f(\mathbf{k})| \\ &= \pm\gamma_0\sqrt{1+4\cos\frac{\sqrt{3}k_x a}{2}\cos\frac{k_y a}{2}+4\cos^2\frac{k_y a}{2}} \end{aligned} \qquad (5)$$

where the $\mathbf{r}_\lambda$'s correspond to bonding vectors that connect neighboring atoms.

When the graphene layer is rolled up into an infinitely long SWNT, the axial wave vector $\mathbf{k}_t$ remains continuous. Meanwhile, the angular wave vector $\mathbf{k}_c$ becomes quantized due to the periodic boundary condition in the circumferential direction: $\mathbf{k}_c \cdot \mathbf{C}_h = 2\pi m$, $(m = 1, 2, 3, \ldots)$, with m being the angular momentum. By substituting the discrete values of $\mathbf{k}_c$ into Eq. (5), and using the following coordinate transformation

$$\begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix} = \begin{pmatrix} \cos\left(\frac{\pi}{6}-\eta\right) & -\sin\left(\frac{\pi}{6}-\eta\right) \\ \sin\left(\frac{\pi}{6}-\eta\right) & \cos\left(\frac{\pi}{6}-\eta\right) \end{pmatrix} \begin{pmatrix} \hat{c} \\ \hat{t} \end{pmatrix} \qquad (6)$$

one immediately obtains the energy dispersion relations for 1D subbands of SWNTs, labeled by different angular momentum m. Theoretical analysis shows that an (n_1, n_2) SWNT is metallic (or quasi-metallic) when $n_1 - n_2$ is a multiple of 3; otherwise, the SWNT is semiconducting [7]. It turns out that for very narrow nanotubes, this simple classification breaks down and the solid-state properties of the nanotube becomes strongly dependent on the curvature effect and the σ–π hybridization [8, 9]. Since this is not the focus of this chapter, we will study relatively wide nanotubes of small curvature in the following.

2.3. Self-consistent formalism

In the presence of an external potential, electrons are driven by the electric field and redistributed on the nanotube surface. The charge/potential profile of the nanotube should be calculated self-consistently in order to capture the screening effect from the electrons. The total Hamiltonian is rewritten as

$$H = H_0 + \sum_i \left(U_i^{\text{ext}} + \sum_j U_{ij}^{\text{ind}} \rho_j^{\text{ind}} \right) c_i^+ c_i \tag{7}$$

where U^{ext} is the external potential and U^{ind} is the electron–electron Coulomb interaction from a non-uniform distribution of the induced charges, assuming a smoothed form as

$$U_{ij}^{\text{ind}} = \frac{1}{\sqrt{\left|\mathbf{r}_i - \mathbf{r}_j\right|^2 + U_0^{-2}}} \tag{8}$$

with e set to unity. U_0 is the on-site Hubband energy and its exact value is found to affect the quantitative results only slightly. As $|\mathbf{r}_i - \mathbf{r}_j|$ increases, the usual form of Coulomb interaction, i.e., $U_{\text{e-e}}(\mathbf{r}) \propto |\mathbf{r}|^{-1}$, is recovered. ρ_j^{ind} is the induced change of occupation number of π-electrons at site j, which is associated with the projection of the occupied eigenstate Ψ_λ on the j-th atomic orbital φ_j as

$$\begin{aligned} \rho_j^{\text{ind}} &= \left(\sum_{\lambda \in \text{occ}} 2 \left| \langle \varphi_j | \Psi_\lambda \rangle \right|^2 \right) - 1 \\ H\Psi_\lambda(\mathbf{r}) &= E_\lambda \Psi_\lambda(\mathbf{r}) \end{aligned} \tag{9}$$

Notice here that the eigen energies and eigen functions are no longer labeled by the axial wave vector **k**, because the translational invariance in the axial direction may not be preserved under external perturbations.

To ensure self-consistency of the total potential and charge distribution, the Hamiltonian in Eq. (7) is diagonalized iteratively. An initial guess of ρ_j^{ind} is first assumed, and Eqs (7) and (9) are solved to yield another set of ρ_j^{ind} constructed from the eigenvectors. At the next iteration, ρ_j^{ind} is obtained from a linear combination of values at the previous and current iterations. The procedure continues until the total potential and charges converge.

The Hamiltonian in Eq. (7) corresponds to an eigen problem of an $N_a \times N_a$ matrix for a system of N_a atoms, which becomes formidable with the growing size of the system. However, in many cases, the complexity of the problem can be effectively reduced by using the symmetry of the system. For example, when the angular dependence of the perturbation is negligible, the nanotube can be treated as a 1D system with uniform charge/potential distribution along the circumference. The size of the Hamiltonian is then reduced to $N_t \times N_t$ for each angular momentum m, with N_t being the total number of unit cells along the nanotube axis. There are also situations in which the external potential is only dependent on the angular coordinates of the nanotube while the axial periodicity is conserved. The eigen problem reduces to solving one $N_c \times N_c$ matrix at each $\mathbf{k}_t$ point, where N_c is the total number of carbon atoms within a unit cell. In either case, however, the induced Coulomb interactions U_{ij}^{ind} in Eq. (7) should be summed over all atomic sites $j = 1, 2, \ldots, N_a$.

3. Metal–semiconductor transition in carbon nanotubes

The metal–semiconductor transition (or metal–insulator transition) has been a subject of interest for decades [10]. It is well understood that MST is typically related to the breaking of a specific symmetry of the system, and CNTs are particularly interesting to study, due to their low dimensionality and special helical symmetry. A most intriguing case is given by the A-SWNTs, which have the highest geometry symmetry of all nanotubes (Figure 2(a)). In addition to the axial wave number k and angular momentum m, the symmetry group of A-SWNTs also contains vertical mirror planes (σ_v) and glide planes (σ'_v), horizontal mirror planes (σ_h) and rotoreflection planes (σ'_h), and two sets of horizontal C_2 rotation axes. The two crossing subbands, π and π^*, have *opposite* parities about σ_v and σ'_v, which make them robust against gap opening for typical non-chiral perturbations, such as curvature effect and many-body interactions. In order to mix the two subbands and lift the degeneracy at the Fermi point, symmetry about all

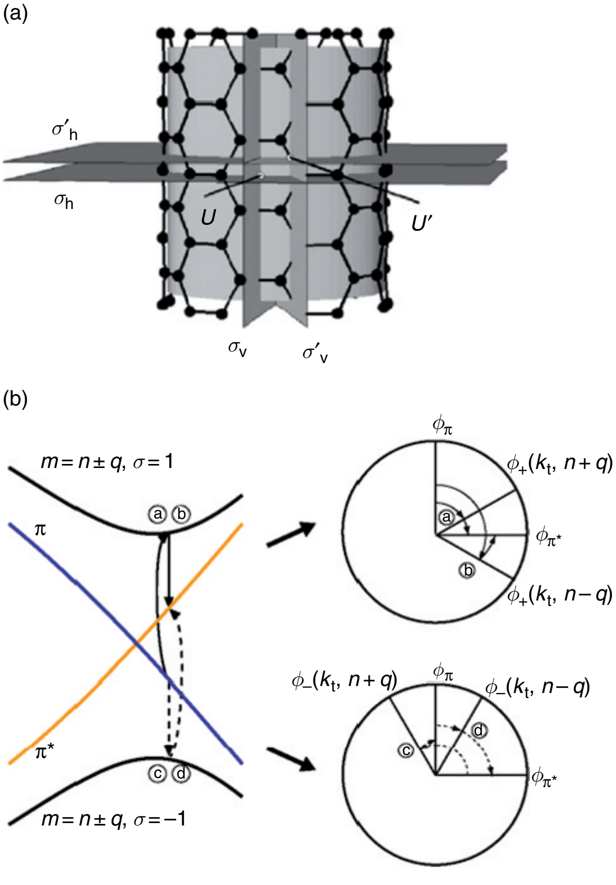


Figure 2 (a) Symmetry elements of A-SWNTs (see text) [23] and (b) Schematics of second order perturbation series between π and π^* subbands, and the corresponding phase angles of the four intermediate states [6]

vertical reflections and glide reflections must be broken simultaneously, which we refer to as the mirror symmetry breaking (MSB) rule.

The high symmetry of A-SWNTs is also responsible for the existence of ballistic (non-scattering) conductivity channels [11], which is very attractive for CNT-based electronic devices [12, 13]. Therefore, a method to modulate and control the conductance of A-SWNTs is particularly desirable from the point of view of applications. Many types of perturbations were studied previously, including intra-rope interactions [14], twisting or bending [15], squashing [16–19], applying uniform perpendicular electric fields [20, 21] and functionalizing the nanotubes [22]. Still, it remained an open question if MSB is a sufficient condition to induce a gap in A-SWNTs. In the following, we derive in a TB formalism the additional conditions for MST in A-SWNTs under external perturbations. For simplicity, the external perturbation is modeled as an angular potential consisting of a single angular Fourier component or as combinations of such potentials. It is found that the MST effect depends not only on the MSB but also on the selection rules between energy subbands of the A-SWNTs.

3.1. Model formulation

A non-orthogonal TB model ($s \neq 0$) is chosen in order to study the influence of the additional electron–hole symmetry on the band gap opening. The electronic states are obtained by solving the stationary Schrödinger equation

$$H\Psi_\lambda = E_\lambda S \Psi_\lambda \tag{10}$$

where H and S are the total Hamiltonian matrix and the overlapping matrix. First-NN approximation with hopping integral $\gamma_0 = -3.033\,\text{eV}$ and overlapping integral $s = 0.129$ are adopted [7].

The wave function of an unperturbed A-SWNT can be expressed as a linear combination of Bloch wave functions $\Psi_\text{A}(\mathbf{k},\mathbf{r})$ and $\Psi_\text{B}(\mathbf{k},\mathbf{r})$ as in Eq. (3), or equivalently as the combination of the two periodic functions $u_{\text{A,B}}(\mathbf{k},\mathbf{r})$, where $u_{\text{A,B}}(\mathbf{k},\mathbf{r})$ is such that $\Psi_{\text{A,B}}(\mathbf{k},\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{\text{A,B}}(\mathbf{k},\mathbf{r})$:

$$\begin{aligned} \Psi_\sigma(\mathbf{k},\mathbf{r}) &= \frac{\mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}}}{\sqrt{2[1-s\sigma|f(k)|]}}\left[\mathrm{e}^{i\phi_\sigma(\mathbf{k})}u_\text{A}(\mathbf{k},\mathbf{r}) + \mathrm{e}^{-i\phi_\sigma(\mathbf{k})}u_\text{B}(\mathbf{k},\mathbf{r})\right] \\ E_\sigma(\mathbf{k}) &= \gamma_0 \frac{-\sigma|f(\mathbf{k})|}{1-s\sigma|f(\mathbf{k})|} \end{aligned} \tag{11}$$

Here, $\sigma = \pm 1$ denotes the conduction and valence bands and $f(\mathbf{k})$ is defined in Eq. (4). The wave vector $\mathbf{k}$ is represented by an axial wave number k and an integer angular momentum m. The phase angle $2\phi_\sigma(\mathbf{k}) \equiv \arg[-\sigma f(\mathbf{k})]$ indicates the phase difference of the coefficient before u_A and u_B, and defines the pseudo-spinor symmetry of the state [11].

Now consider a perturbation H_1, which is uniform in the axial direction, for which k is conserved. The low energy behavior of π and π^* bands can be estimated from a 2×2 effective perturbation Hamiltonian matrix using nearly degenerate perturbation theory [6]:

$$H_{\mathrm{eff}} = \begin{bmatrix} E_{\pi}(k)+H_{\pi\pi}(k) & H_{\pi\pi^*}(k) \\ H_{\pi^*\pi}(k) & E_{\pi^*}(k)+H_{\pi^*\pi^*}(k) \end{bmatrix} \tag{12}$$

where the matrix element, $H_{\alpha\beta}(k)$ with $\alpha,\beta = \pi$ or π^*, can be represented by a perturbation series of different coupling order $\mu's$ as

$$\begin{gathered} H_{\alpha\beta}(k) = \sum_{\mu}\sum_{\{\Psi_i\}} H_{\alpha\beta}^{(\mu)}(\{\Psi_i\}) \\ H_{\alpha\beta}^{(\mu)}(\{\Psi_i\}) = \langle\neg\alpha|\, H_1 \,|\Psi_1\rangle \, \frac{\prod_{i=2}^{\mu-1}\langle\Psi_{i-1}|\, H_1 \,|\Psi_i\rangle}{\prod_{i=1}^{\mu-1}(-E_i)} \, \langle\Psi_{\mu-1}|\, H_1 \,|\neg\beta\rangle \\ \Psi_i \equiv \Psi_{\sigma_i}(k, m_i, \mathbf{r}), \quad E_i \equiv E_{\sigma_i}(k, m_i) \end{gathered} \tag{13}$$

One notes that all intermediate states Ψ_i $(i = 1, \ldots, \mu-1)$ are different from Ψ_{π} and Ψ_{π^*} by definition. Figure 2(b) illustrates an example of second order coupling between π and π^* subbands of an (n, n) A-SWNT through four different perturbation series. Also shown are the phase angles of the intermediate states relative to ϕ_{π} and ϕ_{π^*}, which determine their coupling strength with Ψ_{π} and Ψ_{π^*}.

3.2. Gapping of A-SWNTs

Assume that a scalar perturbation in the form of $H_1 = V_q\cos[q(\theta-\theta_0)]$ is applied to the (n, n) A-SWNT, where V_q is the magnitude of the potential. θ_0 is defined as the minimum angular offset of the vertical mirror planes or glide planes of the nanotube and those of the potential. Using the TB wave function and eigen energy in Eq. (11), the μ-th order perturbation matrix elements in Eq. (14) can be derived as [6]:

$$\begin{aligned} H_{\alpha\beta}^{(\mu)}(\{\Psi_i\}) &= \mathrm{e}^{-i\Delta\mu q\theta_0}\left(\frac{U_q}{2}\right)^{\mu} \frac{P_{\alpha\beta}(\{\Psi_i\})Q(\{\Psi_i\})}{\prod_{i=1}^{\mu-1}(-E_i^0)} \\ P_{\alpha\beta}(\{\Psi_i\}) &= \left[\prod_{i=1}^{\mu-1}\cos(\phi_{i-1}-\phi_i)\right]\cos(\phi_{\mu-1}-\phi_{\mu}+\Delta\mu q\pi/3n) \\ Q(\{\Psi_i\}) &= \prod_{i=1}^{\mu}\left[1-\frac{1}{2}s(\sigma_{i-1}\,|f_{i-1}|+\sigma_i\,|f_i|)\right] \end{aligned} \tag{14}$$

where the subscripts “0” and “μ” correspond to the initial state Ψ_{α} and final state Ψ_{β} respectively, with angular momentum $m_0 = m_{\mu} = n$. We stress that $E_i^0 = -\sigma_i\gamma_0|f(k, m_i)|$, since the factors $(1-s\sigma_i|f_i|)^{-1}$ in Eq. (11) are cancelled by those from the wave functions. $P_{\alpha\beta}(\{\Psi_i\})$ is the total phase of the perturbation series of given order μ while $Q(\{\Psi_i\})$ corrects for contributions from a nonzero orbital overlap s. The intermediate states, $\{\Psi_i \equiv \Psi_{\sigma_i}(k, m_i)\}$ $(i = 1, \ldots, \mu-1)$, satisfy the conservation law for the angular momentum m_i with the constraints:

$$\begin{aligned} m_{i-1} - m_i &= \pm q \quad (i = 1, \ldots, \mu-1) \\ m_{\mu-1} - m_{\mu} &= \pm q + \text{multiples of } 2n \end{aligned} \tag{15}$$

The value of $\Delta\mu$ in Eq. (14) is determined by satisfying the relation that $\sum_{i=1}^{\mu}(m_{i-1}-m_i)=\Delta\mu q+$ multiples of $2n=0$, i.e., $-\Delta\mu q=$ multiples of $2n$. Direct evaluation of Eq. (14) with all possible $\{\Psi_i\}$ sets is formidable. Nevertheless, one can get useful information by applying symmetry arguments, as shown below.

We first study the off-diagonal term $H^{(\mu)}_{\pi\pi^*}$, and replace m_i in Eq. (14) with $\tilde{m}_i = 2n-m_i$, which is allowed by the conservation of angular momentum. The energy denominators and the function Q remain unchanged while the sign of $\Delta\mu$ changes. By defining $\tilde{\Psi}_i \equiv \Psi_{\sigma_i}(k, 2n-m_i)$ and by using the relations between phase angles (see Figure 2(b)), one arrives at

$$\begin{aligned} &P_{\pi\pi^*}(\{\tilde{\Psi}_i\}) = -P_{\pi\pi^*}(\{\Psi_i\}) \\ &H^{(\mu)}_{\pi\pi^*}(\{\Psi_i\}) + H^{(\mu)}_{\pi\pi^*}(\{\tilde{\Psi}_i\}) \propto \left(\frac{U_q}{2}\right)^{\mu} \sin(\Delta\mu q\theta_0) P_{\pi\pi^*}(\{\Psi_i\}) \end{aligned} \tag{16}$$

Since the above relation is true for all possible sets of intermediate states at any coupling order, one can conclude the following. (1) The coupling between π and π^* is always zero if $\theta_0=0$ or when any vertical mirror plane or glide plane of the A-SWNT overlaps with that of the potential. This reproduces the MSB rule and is an explicit result of the mirror symmetry requirement [24]. (2) The coupling is zero if $\Delta\mu=0$, which results from mirror reflection symmetry of the energy bands: $E^0_\sigma(k,m)=E^0_\sigma(k,2n-m)$. This excludes the possibility of any nonzero second order contribution, i.e., $\mu=2$, $\Delta\mu=0$. In other words, a second order band gap is forbidden in A-SWNTs. (3) The next lowest possible $\Delta\mu$ satisfying the angular momentum conservation in Eq. (15) is given by

$$\mu_0 = \frac{2n}{\mathrm{gcd}(2n,q)} \tag{17}$$

in which gcd is the greatest common divisor. μ_0 is also the lowest contributing order of the perturbation series [24]. The very relations (1)–(3) also apply to tensor potentials, although the dependences of the corresponding phase angles are different [6].

The summation over all possible intermediate states can be simplified further by combing the original process ($\{\Psi_i\}$) with the reversal process ($\{\Psi^{\mathrm{R}}_i \equiv \Psi_{-\sigma_{\mu-i}}(k, 2n-m_{\mu-i})\}$). The sign change of the energy results in an extra factor of $(-1)^{\mu-1}$ in the denominator and the function Q changes accordingly. One can prove that

$$\begin{aligned} &P_{\pi\pi^*}(\{\Psi^{\mathrm{R}}_i\}) = P_{\pi\pi^*}(\{\Psi_i\}) \\ &H^{(\mu)}_{\pi\pi^*}(\{\Psi_i\}) + H^{(\mu)}_{\pi\pi^*}(\{\Psi^{\mathrm{R}}_i\}) \\ &\propto [1-(-1)^{\mu}] - s[1+(-1)^{\mu}]\sum_{i=1}^{\mu-1}\sigma_i|f_i| + O(s^2) \end{aligned} \tag{18}$$

Under the approximation of orthogonal basis ($s=0$), only the first term in Eq. (18) exists and it is nonzero only when μ is odd, i.e., when $\mu_0=$ odd. This constraint on μ_0 results from the invariance of the inner product of pseudo-spinors upon reversal operation [11], in combination with the electron–hole symmetry $E^0_{-\sigma}(k,m) = -E^0_\sigma(k,m)$.

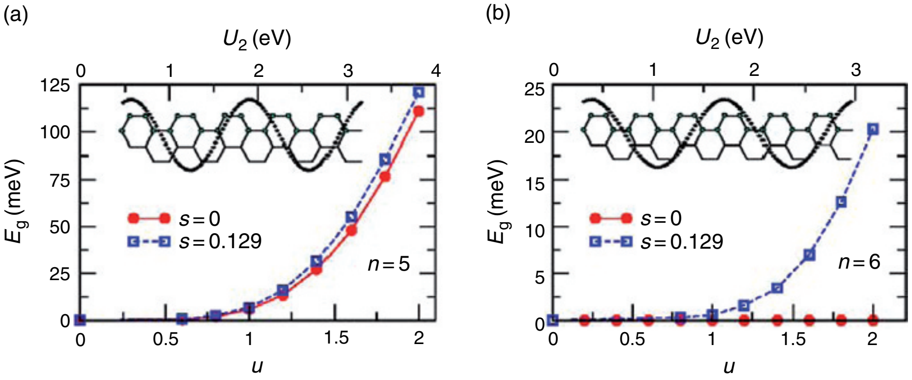


Figure 3 Band gap variation of (a) (5, 5) and (b) (6, 6) A-SWNTs as a function of the angular potential with $q = 2$. Insets: the unwrapped unit cell and schematics of the potential [6]

The latter, however, is not an intrinsic property of A-SWNTs, but rather due to the approximation of first-NN interaction. For instance, the energy band symmetry is broken when the second-NN hopping integral is included, or, if $s \neq 0$. Figure 3 plots the band gap opening at $q = 2$ calculated by the TB method with zero or finite overlap approximations, respectively. At $s = 0$, the (6, 6) A-SWNT remains metallic, because coupling order $\mu_0 = 6$ is forbidden. At nonzero s, a small band gap occurs and increases with the dimensionless potential, $u = U_q R/v_F$, as a power law. In contrast, the band gap curve of a (5, 5) A-SWNT only shows a slight increase at nonzero s, with corrections proportional to s^2.

An interesting feature for possible applications is that the potential has short oscillation period, $q = 2n$, which yields a coupling order $\mu_0 = 1$. Assuming $s = 0$, an analytical expression for the band gap can be obtained by the nearly degenerate perturbation theory:

$$E_g = \sqrt{3} U_o \sin(2n\theta_0) \tag{19}$$

Since π and π^* are now directly coupled, the band gap is linearly proportional to the perturbation and the relation in Eq. (19) holds up to a few electron volts. A potential of this form ($q = 2n$) requires changing the sign of the electrostatic potential alternatively on neighboring carbon atoms. One can possibly generate such perturbations by chemical/biological decoration of the tube or by using the high multipoles of very inhomogeneous potential [25].

The results discussed above can be generalized to arbitrary metallic nanotubes by expanding the TB electronic wave functions near the Fermi point while retaining the chiral-angle dependence in the phase angle $\phi(\mathbf{k})$ [6]. However, due to the lower symmetry, relations in Eqs (16) and (18) may not hold in nanotubes of other chirality, and a finite band gap may occur at a much lower perturbation order. For example, in a uniform electric field applied perpendicularly to the nanotube axis, an (n, n) armchair nanotube always remain metallic due to the high coupling order $\mu_0 = 2n$, while a second order band gap opens for an $(n, 0)$ metallic zigzag nanotube with $E_g \propto R^{-2}$.

3.3. Renormalization of the Fermi velocity

Except for a few special cases, for instance with q = multiples of $2n$, the coupling order between π and π^* is about of the same order n and the resulting band gap remains small. However, the diagonal coupling matrix elements in Eq. (12) are not necessary small. The same symmetry arguments used above can be applied here. For a scalar potential, $P_{\pi\pi}$ (or $P_{\pi^*\pi^*}$) remains the same upon mirror reflection or reversal operation so that

$$\begin{aligned} H^{(\nu)}_{\alpha\alpha}(\{\Psi_i\}) + H^{(\nu)}_{\alpha\alpha}(\{\tilde{\Psi}_i\}) &\propto \left(\frac{U_q}{2}\right)^{\nu} \cos(\Delta\nu q\theta_0) P_{\alpha\alpha}(\{\Psi_i\}) \\ H^{(\nu)}_{\alpha\alpha}(\{\Psi_i\}) + H^{(\nu)}_{\alpha\alpha}(\{\Psi_i^R\}) &\propto [P_{\alpha\alpha} - (-1)^{\nu} P_{\beta\beta}] \\ &\quad - s[P_{\alpha\alpha} + (-1)^{\nu} P_{\beta\beta}] \sum_{i=1}^{\nu-1} \sigma_i |f_i| + O(s^2) \end{aligned} \tag{20}$$

with $\alpha, \beta = \pi \,\text{or}\, \pi^*$, and $\alpha \neq \beta$. Since $\cos(\Delta\nu q\theta_0)$ is always unity when $\Delta\nu = 0$, the lowest contributing coupling order is therefore $\nu_0 = 2$. In an orthogonal basis $(s = 0)$, only the first term in Eq. (20) remains, which results in an energy shift for π and π^* subbands in the same direction when ν is odd and the opposite direction when ν is even. If $s \neq 0$, a relative shift between π and π^* subbands always occurs and their crossing point (the new Fermi point) is shifted.

The values of $H_{\pi\pi}$ and $H_{\pi^*\pi^*}$, although not contributing to the band gap opening, may significantly change the low energy density of states (DOS) or, equivalently, the Fermi velocity. By extracting the linear dependence on $|k - k_{\mathrm{F}}|$ from the second order perturbation term in Eq. (20), the renormalized Fermi velocity is estimated by

$$\bar{v}_{\mathrm{F}} \approx \left(1 - \frac{u^2}{q^2}\right) v_{\mathrm{F}} \tag{21}$$

which agrees very well with numerical results from TB calculations [6]. The low energy DOS is enhanced by the external potential, which is more evident for large radius A-SWNTs due to the power law dependence on $u = U_q R / v_{\mathrm{F}}$.

3.4. Combination of perturbations

Realistic perturbations usually have more than one dominating angular mode, and the interplay of different angular components may result in a stronger influence on the electronic properties of the nanotube. In the simplest case, only two angular components are present: $H_1 = V_1 \cos[q_1(\theta - \theta_1)] + V_2 \cos[q_2(\theta - \theta_2)]$. Here θ_1 and θ_2 are defined as the angular offsets of the vertical mirror (or glide) planes of the A-SWNT with regard to those of V_1 and V_2. The coupling order μ_0 is now a function of both q_1 and q_2. For $q_1 = 1, q_2 = 2$, the lowest-order nonzero coupling is of the 3rd order through $\psi_{\pi}(k) \rightarrow \psi_{\sigma 1}(k, m_1) \rightarrow \psi_{\sigma 2}(k, m_2) \rightarrow \psi_{\pi^*}(k)$ with 24 possible combination of intermediate states $\psi_{\sigma 1}(k, m_1)$ and $\psi_{\sigma 2}(k, m_2)$. By substituting these states into Eq. (14) the band gap is evaluated as [24]

$$E_g \approx 2|H^{(3)}_{\pi\pi^*}(k_{\mathrm{F}})| \propto \frac{u^3}{n^2} \sin(2\theta_{\mathrm{d}}) \tag{22}$$

where $\theta_{\mathrm{d}} = \theta_2 - \theta_1$ and the dimensionless potential u is defined as $(V_1^2 V_2)^{1/3} R/v_{\mathrm{F}}$. Equation (22) predicts that whenever mirror planes of the two potential components coincide ($\theta_{\mathrm{d}} = 0$), the band gap should be zero whether or not the MSB condition is satisfied. The maximum band gap occurs at $\theta_{\mathrm{d}} = \pm\pi/4$, which is plotted in Figure 4 for A-SWNTs of different radii. Since this combination always gives a secondary band gap ($E_{\mathrm{g}} \sim R^{-2}$), which could be hundreds of meV for A-SWNTs with a moderate radius, it may be an effective mechanism to induce MST effect and be used in tunable metallic field-effect transistor [25].

Another way to reduce the coupling order μ_0 is by combining perturbations of different types, e.g. an elastic radial deformation and a uniaxial electrostatic potential. Assume that a scalar potential (denoted as U) is applied on the A-SWNT together with a tensor perturbation (denoted as $* \propto \delta\gamma_q$) of same angular momentum q, but with an

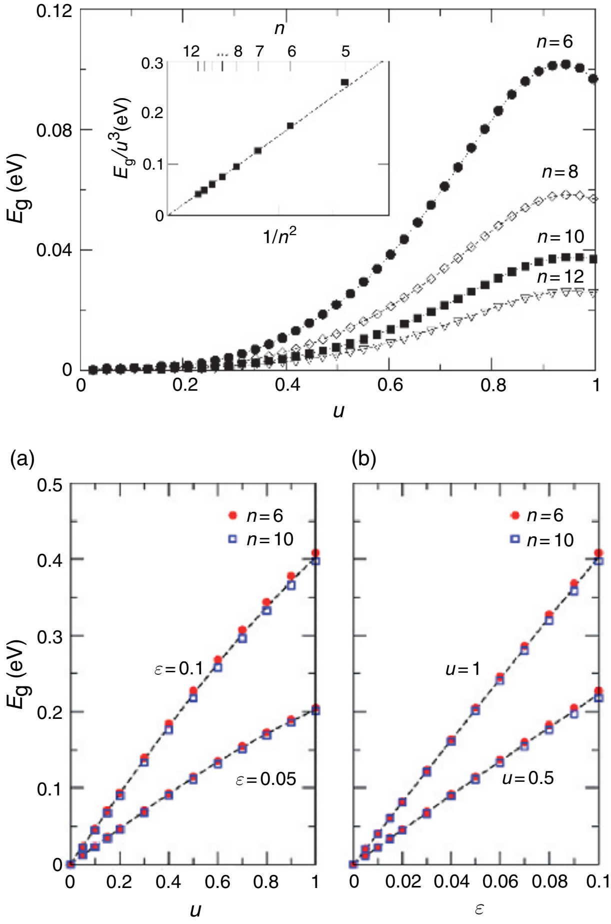


Figure 4 Top: band gap variation of A-SWNTs as a function of u under a scalar potential consisting of two Fourier components ($q_1 = 1$, $q_2 = 2$), with $V_1 = V_2$ and $\theta_{\mathrm{d}} = \pm\pi/4$ [24]. Bottom: band gap variation of A-SWNTs as a function of u and ε under the combined perturbation of an electrostatic potential and a radial deformation, both with angular momentum $q = 2$ [6]

angular difference θ_d between the mirror planes of these two components. It turns out that even if the second order contribution from either component is zero, the cross terms do not necessarily vanish. Summing up the cross terms, one obtains

$$E_g \approx 2|H^{(2)}_{\pi\pi^*,\mathrm{cross}}(k_\mathrm{F})| \propto u\varepsilon \sin(q\theta_\mathrm{d})\frac{|\gamma_0|}{q} \quad (q \ll n) \tag{23}$$

where $u = U_q R/v_\mathrm{F}$ and $\varepsilon \sim \delta\gamma_q/\gamma_0$ are the dimensionless potential and strain respectively. Similar dependence on θ_d is found as from the previous case of mixed scalar potential components, but here the band gap is almost independent of radius for fixed values of u and ε, as confirmed for the case of $q = 2$ in Figure 4.

The linear dependence of E_g on $u = U_q R/v_\mathrm{F}$ makes it promising to generate a substantial band gap in a large radius A-SWNT by using only a moderate external perturbations. Recently, it was found that the conductance of a carbon nanotube can be controlled by tuning the voltage of a local gate placed near mechanical defects on the tube, such as kinks or bends [26]. This scenario is reminiscent of the combination of scalar and tensor potentials of $q = 1$, and might be related to the resulting second order band gap.

4. Finite-size effect of carbon nanotubes

Carbon nanotubes, as building blocks of nanotechnology, hold great promise for nano-scale biological applications. It has been proposed that the small size and stable structure make short CNTs good candidates for artificial nano-scale channels of water [27–30], protons [31, 32], ions [33], or polymers [34]. Confinement of water molecules inside the hydrophobic interior of narrow nanotubes has been recently confirmed by experiments using X-ray diffraction [35] and neutron scattering [36]. On the other hand, as the length of a CNT is shortened to a few tens of the radius, the energy dispersion changes from a continuous spectrum to discrete energy levels while the electronic properties such as the dielectric response also change from 1D-like to 0D-like. At the same time, the terminations at the edges of a finite CNT result in appreciable buildup of partial charges, which plays an important role in the energetics of CNT-based channels. Therefore, studying the finite-size effect of CNTs is essential for understanding and controlling the electronic behavior of short CNTs in biological applications.

Below, we choose finite-length (6, 6) A-SWNT segments as our model system. The atomic partial charges on an unperturbed CNT are calculated using the DFT (B3LYP/6-31G*) method and assigned to the CNT as parameters in the simulation. The electronic properties are calculated within a self-consistent third-NN π-orbital TB method, which includes interactions up to the third-NNs with values derived previously [37]: $\gamma_0 = -2.97\,\mathrm{eV}$, $\gamma_1 = -0.073\,\mathrm{eV}$, $\gamma_2 = -0.33\,\mathrm{eV}$ for the first-, second- and third-NN hopping integrals. $U_0 = 14.6\,\mathrm{eV}$ is used for the on-site Hubband energy [38] in a smoothed form of the electron–electron Coulomb interactions as shown in Eq. (8). This description is then combined with classical molecular dynamics (MD) simulations to study the interaction between external molecules and the CNT channel, and the movement of the molecules through the channel. As shown below, this multi-level approach provides an efficient way to capture the electrostatics of the nanotube, which in turn determines the energetics of the system and the transport behavior of the confined molecules.

4.1. Structure and atomic partial charges

The finite-length (6, 6) A-SWNT segments are saturated with hydrogen atoms to avoid drastic geometry reconstruction caused by the dangling bonds [39]. C–C and C–H bond lengths in the ideal geometry are fixed at $d_{\mathrm{CC}} = 1.44\,\text{Å}$ [7] and $d_{\mathrm{CH}} = 1.09\,\text{Å}$ [40]. Segments of different lengths are label by l, the number of carbon atom sections. As shown in Figure 5, segments of odd and even l belong to D_{6h} and D_{6d} symmetry groups, respectively. The former contains a horizontal mirror reflection (σ_h) while the latter contains a rotation–reflection ($S_{2n} = C_{2n}\sigma_h$ with $n = 6$), which corresponds to a regular rotation about the axis by $2\pi/2n$ followed by the regular σ_h operation.

The corresponding selection rules for transitions between the electronic states are also slightly different for segments of these two different symmetry groups [41]. The atomic partial charges of A-SWNT segments, ρ_0, are determined under the restricted electrostatic potential (RESP) fitting scheme by matching the surrounding electrostatic potential (ESP) built by ρ_0 to the ESP from the DFT method [42]. For the range of segment lengths we studied ($l = 2$ to 12), it is found that due to its relative lower electronegativity, each hydrogen loses 0.13–0.14 electrons, which are mostly picked up by the outermost carbon atoms while the charges on other carbon atoms decay toward the tube center rapidly, as demonstrated in Figure 6 for $l = 12$.

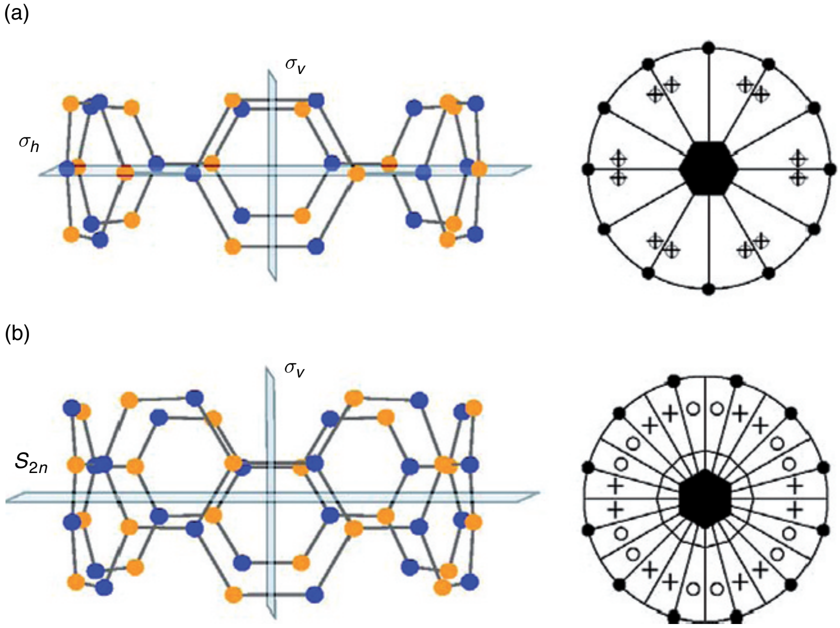


Figure 5 Structure of (6, 6) A-SWNT segments with $l = 3$ (top) and $l = 4$ (bottom) and stereographs of the corresponding D_{6h} and D_{6d} symmetry groups

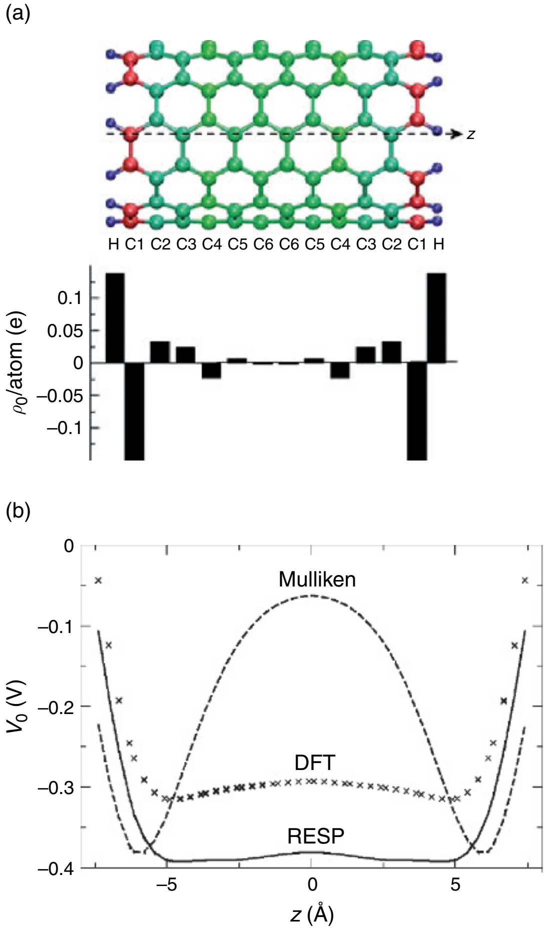


Figure 6 (a) The structure of a 12-section (6, 6) A-SWNT and the atomic partial charges from RESP population analysis. (b) Corresponding electrostatic potential along the tube axis in comparison with results from DFT calculation and from Mulliken population analysis

Also shown is the ESP profile calculated from the RESP charges, which features a wide potential well of about $-0.32\,\mathrm{eV}$ inside the A-SWNT and agrees reasonably well with the DFT ESP profile except for a slight overall shift. Also plotted in Figure 6(b) is the ESP calculated from the Mulliken population analysis, which produces an artificial barrier of $0.32\,\mathrm{eV}$ in the middle of the tube, resulting in distorted channel energetics. A further analysis indicates that the atomic partial charges mostly reside in the s orbitals of hydrogen atoms (s_H) and the σ orbitals of carbon atoms [41]. These highly localized charges are not expected to move during weak electronic perturbations, but the resulting local dipoles at the edges contribute significantly to the electrostatics of the short CNT and may, for instance in a CNT-based molecular channel, greatly influence the entrance and permeation of molecules through the channel. Geometry optimization at the B3LYP/6-31G* level shows a contraction of C–C bond length from 1.44 Å to

1.36–1.39 Å while the changes in the C–H bond and other C–C bonds are quite small (<2%) [42]. The overall magnitude of atomic partial charges is also decreased in the optimized structure.

4.2. Band gap oscillation

An infinitely long A-SWNT is metallic with two subbands crossing at the Fermi level. As the tube length is reduced, the energy levels become quantized and a finite gap opens between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO). The energy gaps of nanotubes with ideal geometry computed with B3LYP and TB method are shown in Figure 7 as a function of tube length. Both curves exhibit similar oscillation patterns, and clearly fall into three different classes characterized through lengths $l = 3s+w$, with s being an integer and $w = 0, \pm 1$.

It was shown earlier that including the third-NN interaction is necessary to lift the degeneracy of HOMO and LUMO states at $w = -1$ and reproduce the nonzero band gaps from DFT calculations [43]. The energy gap predicted by TB method is systematically lower than the B3LYP result by ~1 eV at short lengths, because the TB parameters used here were derived by fitting the band structure of graphene to pure DFT results [37], with the latter known for underestimating the band gap [40]. On the other hand, as a hybrid DFT variant, the B3LYP method should improve the accuracy of energies by using a mixture of the Hatree–Fock exchange and the DFT exchange-correlation functional [44]. The differences of gap minima in the two descriptions are caused by the neglect of curvature effect in our TB model, which has a much less important influence on the dielectric response. Upon structure optimization, the energy distance between HOMO and LUMO is systematically larger, as a result of reducing the total energy of the system.

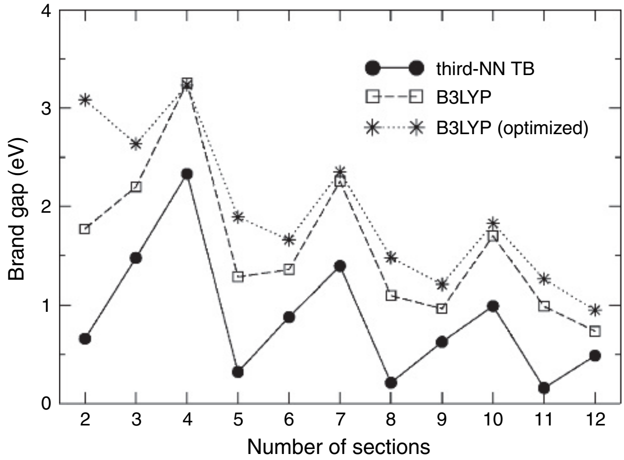


Figure 7 Band gap variation of a (6, 6) nanotube as a function of tube length [42]

4.3. Dielectric responses

For infinitely long CNTs, the static dielectric functions can be expressed in the Fourier space as [45]:

$$\begin{aligned} \varepsilon_m(k) &= 1 + \chi_m(k) G_m(k) \\ G_m(k) &= 2 I_m(k) K_m(k) \end{aligned} \tag{24}$$

where the density response function $\chi_m(k)$ and the electrostatic kernel $G_m(k)$ are labeled by the axial wave number k and the angular momentum m. I and K are the modified Bessel functions of the first and second kind. The electronic potential imposed by uniform electric fields parallel and perpendicular to the axis corresponds to $m = 0,\ k \to 0$ and $m = \pm 1,\ k = 0$, respectively. Electronic transitions contributing to dielectric response must satisfy the selection rules $|k_1 - k_2| = k$ and $|m_1 - m_2| = m$. For finite-length CNTs, the axial wave vector becomes quantized but similar selection rules apply; while on the other hand, the overall screening ability is expected to be much weakened due to long range nature of the Coulomb interaction.

To investigate the effect of size on the dielectric behavior of A-SWNTs, we estimate the averaged screening constants by turning on a uniform weak electric field parallel and perpendicular to the nanotube axis and extract the corresponding changes in the ESP profile from the middle region of the nanotube. Results are plotted in Figure 8 as a function of index l.

As expected, the parallel dielectric constant is drastically reduced, compared to the bulk values, as the screening in the parallel direction mostly arises from transitions between electronic states with small difference in the axial wave vector, i.e., with small Δk with length-induced limitation $\Delta k \sim 1/\mathrm{L}$. Both TB and B3LYP methods show very similar oscillation patterns in the period of three sections, except that values of ${\varepsilon_{||}}^{\mathrm{TB}}$ are about 10% higher than results from ${\varepsilon_{||}}^{\mathrm{B3LYP}}$. This can be partly attributed to the fact that the pure DFT method, used in the TB parameterization, usually overestimates the dielectric constant. The oscillation in $\varepsilon_{||}$ though, is not directly correlated with that of the band gap shown in Figure 8, as can be seen by comparing $\varepsilon_{||}$ and the inverse of ${E_{\mathrm{g}}}^{-1}$. This independence on E_{g} can be understood from the special vertical mirror reflection symmetry (σ_v) of A-SWNTs, which is preserved in the finite segments (see Figure 5). The transition from HOMO to LUMO, which possesses opposite parities with regard to σ_v, is forbidden by conservation requirement of the parities.

For dielectric response to fields applied perpendicularly to the tube axis, both TB and B3LYP methods predict similar oscillation patterns of the corresponding dielectric constant $\varepsilon_\perp$ while small differences arise from neglecting the contribution of σ electrons and the curvature effect in the TB description. The inset of Figure 8(b) shows $\varepsilon_\perp^{\mathrm{TB}}$ for A-SWNT segments up to 30 sections long and one can clearly discern the splitting of the values into two tails: one converging monotonically to the bulk value ($\varepsilon_\perp \sim 5$ [24]) while the other converging in an oscillating fashion and more quickly.

As shown in Figure 8, energy optimized A-SWNT segments always exhibit a lower screening capability in both directions. This is because structure optimization always tends to increase the energy separation between occupied states and unoccupied states, resulting in larger electronic transition energies and smaller dielectric constants.

So far, the dielectric response has been evaluated under weak external electric field and the nanotube is assumed to be charge neutral. As pointed out for the case of infinitely

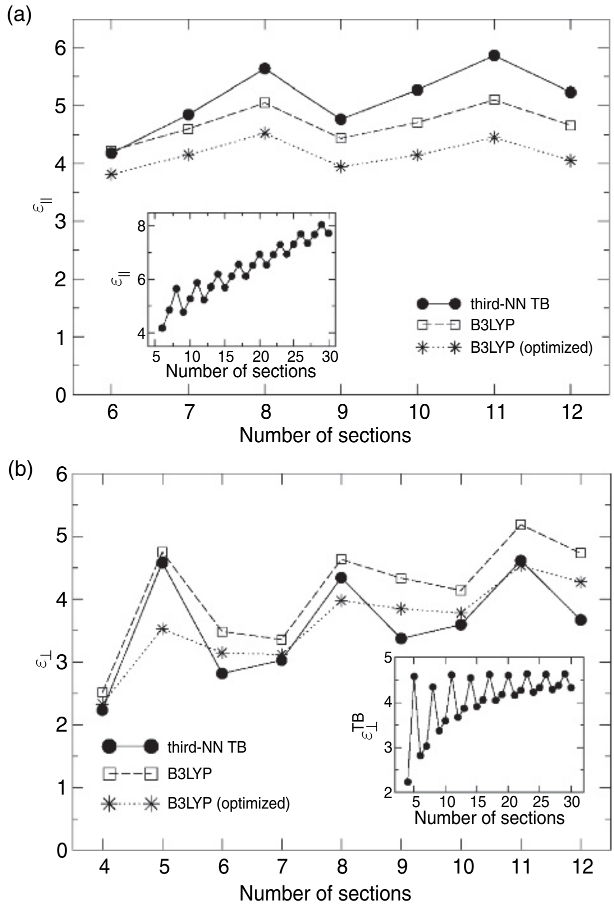


Figure 8 Parallel dielectric constant $\varepsilon_\|$, and the perpendicular dielectric constant, $\varepsilon_\perp$, for (6, 6) A-SWNT segments with different length, calculated from both TB and B3LYP methods. Insets: the TB dielectric constant for A-SWNTs with up to 30 sections

long nanotubes, the dielectric constant is expected to be enhanced upon injection of charges into the nanotube or under strong electric field [20, 46]. It would be interesting to see how the dielectric behavior of finite-length nanotubes changes under similar conditions, as it is usual practice in CNT applications to dope the nanotube or to modulate the device behavior by applying electric fields, e.g., through a gate or a tip.

4.4. Examples: Short CNTs in biological applications

The nanotube models used in most existing MD simulations have an incomplete description of nanotube electrostatics, which may affect accurate modeling of transport through the nanotube channel. For example, most models treat the nanotube carbon atoms as neutral, thus neglecting the local dipoles formed at the nanotube edge due to charge transfer and the polarization effect from the delocalized π electrons. Both are important sources of electrostatic interaction. Attempts have been made to employ *ab initio* MD methods [31, 32] to study water and proton transport through nanotubes, but the computational cost limits their applications in MD settings. The self-consistently TB approach discussed above, when combined with classical MD simulations and using parameters from first principle calculation, may serve as an efficient and sufficiently reliable way to characterize the CNT-based channel and account for the electron degrees of freedom of the system [5, 42, 43].

As a demonstration, here we show examples based on this semi-empirical description to study the interaction of a single water molecule or a potassium ion with a finite-length nanotube channel in biological applications. The atomic partial charges and the polarization of the nanotube are found to be important in determining the energetics of such system and the movement of water/ions through the channel.

We look first at a system consisting of a single water molecule and an open-ended 12-section (6, 6) A-SWNT with ends saturated with hydrogen atoms [42]. The interaction between the water molecule and the nanotube includes the short-range van der Waals (vdW) interaction calculated with a standard Lennard–Jones potential, $U_{\mathrm{L-J}}$, the Coulomb interaction arising from the atomic partial charges q_{0i}, U_0, and the interactions with the induced charges δq_{i} on the A-SWNT, U_{ind}, where the index i runs over all atomic sites of the CNT. Geometry and atomic partial charges of the water molecule are assumed to be those of the TIP3P model [47], i.e., with $d_{\mathrm{OH}} = 0.9572\,\text{Å}$, $\theta_{\mathrm{HOH}} = 104.52°$, and $q(\mathrm{O}) = -2q(\mathrm{H}) = -0.834e$. Amber 94 force field parameters [48] are used for the C-O Lennard–Jones potential:

$$U_{\mathrm{L-J}}(r) = 4.0\varepsilon_{\mathrm{CO}}[(\sigma_{\mathrm{CO}}/r)^{12} - (\sigma_{\mathrm{CO}}/r)^{6}] \tag{25}$$

with $\varepsilon_{\mathrm{CO}} = 0.1143\,\mathrm{kcal/mol}$ and $\sigma_{\mathrm{CO}} = 3.2752\,\text{Å}$. The atomic partial charges on the CNT are determined by RESP scheme at the B3LYP/6-31G* level on an isolated CNT and fixed during the simulation: $q_0(\mathrm{H}) = 0.138308e$, $q_0(\mathrm{C1}) = -0.176750e$, $q_0(\mathrm{C2}) = 0.032725e$, $q_0(\mathrm{C3}) = 0.023788e$, $q_0(\mathrm{C4}) = -0.022954e$, $q_0(\mathrm{C5}) = 0.006498e$ and $q_0(\mathrm{C6}) = -0.001615e$, where the carbon atoms are counted from the edge towards the center. The induced charges are calculated self-consistently through the TB method, treating electric fields from the water dipole as external perturbations in the TB Hamiltonian.

Figure 9(a) illustrates the total interaction energy and its components as a single water molecule with fixed orientation moves along the nanotube axis. $U_{\mathrm{L-J}}$ is found to be almost symmetric with a simple potential well of $-4.34\,\mathrm{kcal/mol}$ in the middle of the tube. In contrast, U_0 is anti-symmetric due to the nonzero axial dipole moment of the water molecule, $\mu_z \approx 2.35\,\mathrm{Debye}$. The edge dipoles on the nanotube strongly affect the energetics of water entering the tube: on the left (right) side, the water molecule encounters a shallow potential well (barrier) of $0.44\,\mathrm{kcal/mol}$ followed by a high potential barrier (well) of $2.97\,\mathrm{kcal/mol}$. In the middle of the tube, the potential is nearly vanishing. U_{ind} has a similar shape as $U_{\mathrm{L-J}}$, but is about one order of magnitude smaller.

In summing up all three contributions, the total potential, U_{tot}, exhibits several characteristic features. At the edges, the shape of U_{tot} is dominated by U_0, and the resulting energy barrier from the edge dipoles may slow down the transport of water through the

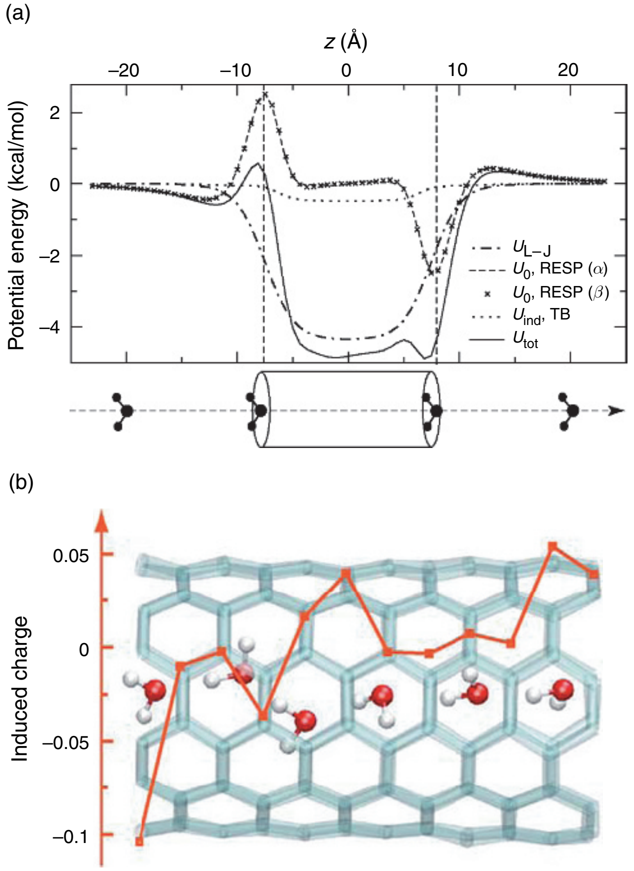


Figure 9 (a) Interaction energies between an ideal 12-section (6, 6) A-SWNT and a single water molecule of fixed orientation at various positions along the nanotube axis α and β refer to different RESP fitting schemes [42]. (b) A snapshot of MD simulation with a 6-water-chain inside the same nanotube. The overlaying curve plots the summed induced charges of each carbon section of the nanotube [43]

nanotube. With a different termination, e.g., with atoms or functional groups other than hydrogen atoms, the shape and magnitude of U_0 and even the dynamics of the whole system may be quite different. Therefore it is extremely important to include the proper atomic charges and the corresponding electrostatic potential term U_0 in a realistic model of CNT channels. In the middle of the tube, the major contribution to U_{tot} stems from $U_{\text{L-J}}$, while U_{ind} further lowers the potential well by about 0.5 kcal/mol.

Though small, the polarization effect from U_{ind} is not negligible when considering the filling of the CNT channel by a one-dimensionally ordered chain of water molecules instead of just a single molecule. Figure 9(b) is a snapshot taken from a classical MD simulation with six water molecules inside the same A-SWNT. Due to the confinement of the nanotube, these water molecules are arranged in a single file and their dipole moments are preferably aligned along the tube axis. According to previous MD studies, filling of the hydrophobic interior of a narrow nanotube results from a delicate energy balance and is subject to change under external perturbations as small as few $k_B T$ [28], with k_B being the Boltzman constant and T the room temperature. We found that U_{ind} between the six-water chain confined inside a short nanotube and the images charges on the nanotube surface (Figure 9(b)) amounts to about the same energy scale ($\sim 6\,k_B T$), while the total dipole moment of the water chain is reduced from 12.23 to 5.24 Debye. A similar yet stronger screening effect has also been observed in an *ab initio* calculation [32].

When the molecule interacting with the nanotubes is charged, the situation becomes drastically different [5]. Figure 10(a) shows the interaction energies of a potassium ion (K^+) with the same nanotube as a function of the axial location of the ion. The Lennard–Jones potential between the K^+ ion and the nanotube is modeled using the CHARMM27 force field, with $\varepsilon_{\text{KH}} = 0.043749\,\text{kcal/mol}$, $\varepsilon_{\text{KC}} = 0.078039\,\text{kcal/mol}$, $\sigma_{\text{KH}} = 2.7473\,\text{Å}$, and $\sigma_{\text{KC}} = 3.3464\,\text{Å}$ [49, 50]. In contrast with the case of a water molecule, the electrostatic interaction terms now dominate over the Lennard–Jones potential in the whole range. The Coulomb interaction from the atomic partial charges, U_0, takes a symmetric shape and is attractive inside the nanotube and repulsive outside the nanotube, with an energy barrier of about 7.02 kcal/mol from the edge dipoles. The polarization effect from π electrons further lowers the potential by about 25 kcal/mol. Most importantly, U_{ind} effectively reduces the access barrier to 1.05 kcal/mol, making it possible for a monovalent cation to enter the nanotube simultaneously. The interaction energies along the x axis in the tube center are also plotted in Figure 10(b), which indicates a local maximum at the origin and a minimum at a radial distance of about 1.3 Å, as a competing result of a more repulsive vdW potential and more attractive Coulomb interaction when the ion moves towards the wall of the CNT.

To study the dynamics of the K^+ ion inside the short CNT channel, we simulate the system by employing the polarizable nanotube model under the micro-canonical ensemble with the nanotube fixed at its ideal geometry. The K^+ ion is initially placed still on the z-axis $(0, 0, -8.803\,\text{Å})$, close to one end of the nanotube. Due to the attraction from the image charges on the nanotube, the K^+ ion starts to oscillate. The instantaneous induced charges at each atomic sites on the nanotube are updated constantly during the MD simulation, while the initial atomic partial charges are always fixed.

During the simulation of 4.84 ps (with a time step of 2 fs), the ion finishes two complete cycles at a frequency of 0.43 THz, which agrees well with the estimated value of 0.41 THz obtained by fitting the harmonic component of the interaction potential. The intervals used for numerical fitting are indicated by the dashed lines in Figure 10. If the K^+ ion is set off on the z-axis by 1.0 Å, for example, the axial oscillation of the ion is then accompanied by fluctuations in the radial direction, resulting in a slightly decreased oscillation frequency [5]. Energy dissipation, e.g., through interaction with the phonon modes of the nanotube, should dampen the oscillation observed in the above QM/MM simulation. A complete characterization of this nano-oscillator will require knowledge of the vibrational modes as well as the dynamic dielectric response of the nanotubes. Nevertheless, this simple CNT-based terahertz (THz) oscillator scheme could conceivably lead to a THz wave-detector at room temperature, and there should be room to improve and modify the interaction strength and oscillation dynamics by selecting CNT geometry, terminations, and the type of interacting ions or molecules.

In the water/CNT or K^+/CNT complex, the atomic partial charges on the nanotube edges constitute a major part of the total interaction energy, while the polarization from

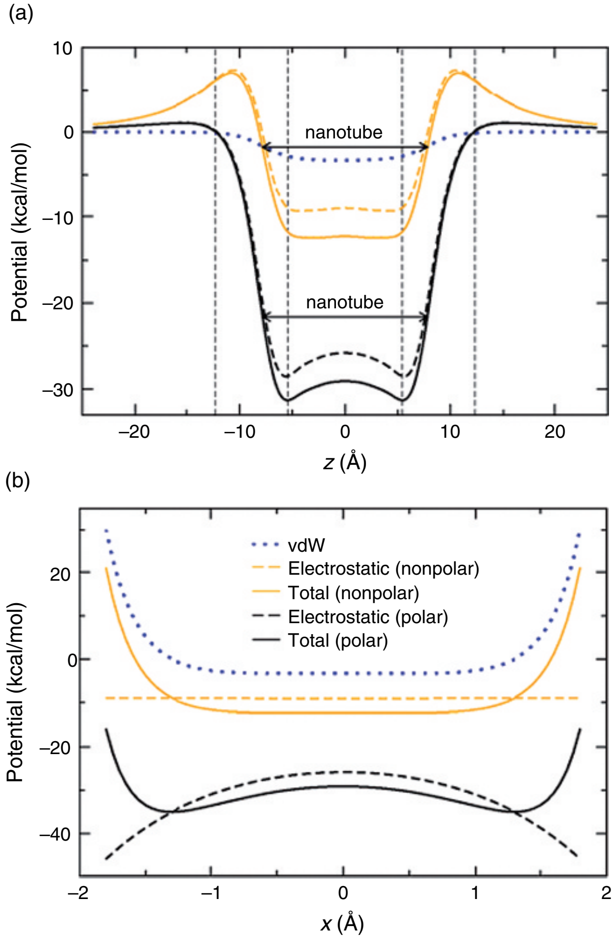


Figure 10 Interaction energy between the K^+ ion and the nanotube along (a) the z axis and (b) the x axis. Comparison is made between a nonpolar model (without contribution of U_{ind}) and a polar model (with contribution of U_{ind})

π electrons considerably lowers the electrostatic energy. These properties should affect the motion of charged or polar molecules entering, moving through and reorienting inside the CNT channel. We have provided the proof of concept that a multiple-level approach combining a quantum mechanical model (including first principle and semi-empirical methods) and molecular mechanical simulations provides a fast yet efficient way to capture the main features of such systems at a relatively low computational cost.

5. Conclusion

In this work, we have applied a semi-empirical TB method to study both infinite and finite-size CNTs under electronic perturbation. The inclusion of charge self-consistency and the combination with methods at different approximation levels make it an accurate and efficient way to treat the electron degrees of freedom of the CNT-based system. We have demonstrated that the low dimensional (0D and 1D) characteristics of geometric and electronic properties make CNTs a promising element with unique properties for applications in electronic transport devices as well as in biological systems involving nano-scale confinement in molecular channels.

Acknowledgements

This work was supported by the National Science Foundation through grants NCN EEC-0228390 and NSF CCR 01-21616, and by the US Army Research Office DURINT program through contract SIT 527826-08. We also acknowledge Deyu Lu, Slava V. Rotkin and Klaus Schulten for their collaboration in related research projects.

References

[1] S. Iijima and T. Ichihashi, *Nature*, 363 (1993) 603.
[2] S. Iijima, *Nature*, 354 (1991) 56.
[3] D. S. Bethune, *et al.*, *Nature*, 363 (1993) 605.
[4] P. Avouris, *Acc. Chem. Res.*, 35 (2002) 1026.
[5] D. Lu, Y. Li, U. Ravaioli, and K. Schulten, *Phys. Rev. Lett.*, 95 (2005) 246801.
[6] Y. Li, U. Ravaioli, and S. V. Rotkin, *Phys. Rev. B*, 73 (2006) 035415.
[7] R. Saito, G. Dresselhaus, and M. S. Dresselhaus, *Physical Properties of Carbon Nanotubes*, Imperial College Press, London, 1998.
[8] N. Hamada, S. I. Sawada, and A. Oshiyama, *Phys. Rev. Lett.*, 68 (1992) 1579.
[9] I. Cabria, J. W. Mintmire, and C. T. White, *Phys. Rev. B*, 67 (2003) 121406.
[10] N. F. Mott, *Metal-Insulator Transitions*, Taylor and Francis, London, 1974.
[11] T. Ando, T. Nakanishi, and R. Saito, *J. Phys. Soc. Jpn.*, 67 (1998) 2857.
[12] S. V. Rotkin and S. Subramoney (eds), *Applied Physics of Nanotubes: Fundamentals of Theory, Optics and Transport Devices*, Springer-Verlag, Berlin, 2005.
[13] M. S. Dresselhaus, P. Avouris, and G. Dresselhaus (eds), *Carbon Nanotubes: Synthesis, Structure, Properties, and Applications*, Springer-Verlag, Berlin, 2001.
[14] P. Delaney, H. J. Choi, J. Ihm, S. G. Louie, and M. L. Cohen, *Nature*, 391 (1998) 466.
[15] C. L. Kane and E. J. Mele, *Phys. Rev. Lett.*, 78 (1997) 1932.
[16] C. J. Park, Y. H. Kim, and K. J. Chang, *Phys. Rev. B*, 60 (1999) 10656.
[17] P. E. Lammert, P. H. Zhang, and V. H. Crespi, *Phys. Rev. Lett.*, 84 (2000) 2453.
[18] O. Gülseren, T. Yildirim, S. Ciraci, and Ç. Kılıç, *Phys. Rev. B*, 65 (2002) 155410.
[19] J. Q. Lu, *et al.*, *Phys. Rev. Lett.*, 90 (2003) 156601.
[20] Y. Li, S. V. Rotkin, and U. Ravaioli, *Nano Lett.*, 3 (2003) 183.
[21] J. O'Keeffe, C. Y. Wei, and K. J. Cho, *Appl. Phys. Lett.*, 80 (2002) 676.
[22] Y.-S. Lee, M. B. Nardelli, and N. Marzari, *Phys. Rev. Lett.*, 95 (2005) 076804.
[23] M. Damnjanović, I. Miloševć, T. Vuković, and R. Sredanović, *Phys. Rev. B*, 60 (1999) 2728.
[24] Y. Li, S. V. Rotkin, and U. Ravaioli, *Appl. Phys. Lett.*, 85 (2004) 4178.
[25] S. V. Rotkin and K. Hess, *Appl. Phys. Lett.*, 84 (2004) 3139.
[26] M. J. Biercuk, N. Mason, J. M. Chow, and C. M. Marcus, *Nano Lett.*, 4 (2004) 2499.
[27] W. H. Noon, K. D. Ausman, R. E. Smalley, and J. Ma. *Chem. Phys. Lett.*, 355 (2002) 445.
[28] G. Hummer, J. C. Rasaiah, and J. P. Noworyta, *Nature*, 414 (2001) 188.
[29] R. J. Mashl, S. Joseph, N. R. Aluru, and E. Jakobsson, *Nano Lett.*, 3 (2003) 589.

[30] F. Zhu and K. Schulten, *Biophys. J.*, 85 (2003) 236.
[31] C. Dellago, M. M. Naor, and G. Hummer, *Phys. Rev. Lett.*, 90 (2003) 105902.
[32] D. J. Mann and M. D. Halls, *Phys. Rev. Lett.*, 90 (2003) 195503.
[33] S. Joseph, R. J. Mashl, E. Jakobsson, and N. R. Aluru, *Nano Lett.*, 3 (2003) 1399.
[34] C. Wei and D. Srivastava, *Phys. Rev. Lett.*, 91 (2003) 235901.
[35] Y. Maniwa, *et al.*, *J. Phys. Soc. Jpn.*, 71 (2002) 2863.
[36] A. I. Kolesnikov, *et al.*, *Phys. Rev. Lett.*, 93 (2004) 035503.
[37] S. Reich, J. Maultzsch, C. Thomsen, and P. Ordejón, *Phys. Rev. B*, 66 (2002) 035412.
[38] K. Esfarjani and Y. Kawazoe, *J. Phys.: Condens. Matter*, 10 (1998) 8257.
[39] S. Hou, Z. Shen, X. Zhao, and Z. Xue, *Chem. Phys. Lett.*, 373 (2003) 308.
[40] A. Rochefort, D. R. Salahub, and P. Avouris, *J. Phys. Chem. B*, 103 (1999) 641.
[41] Y. Li, Ph.D Thesis, University of Illinois at Urbana-Champaign, 2005.
[42] D. Lu, Y. Li, U. Ravaioli, and K. Schulten, *J. Phys. Chem. B*, 109 (2005) 11461.
[43] D. Lu, Y. Li, S. V. Rotkin, U. Ravaioli, and K. Schulten, *Nano Lett.*, 4 (2004) 2383.
[44] A. D. Becke, *J. Chem. Phys.*, 98 (1993) 5648.
[45] T. Ando, *J. Phys. Soc. Jpn.*, 66 (1997) 1066.
[46] Y. Li, S. V. Rotkin, and U. Ravaioli, *Proc. Third IEEE Conf. Nanotechnol.*, 1 (2003) 1.
[47] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, *J. Chem. Phys.*, 79 (1983) 926.
[48] W. D. Cornell, P. Cieplak, C. I. Bayly, and P. A. Kollman, *J. Am. Chem. Soc.*, 115 (1993) 9620.
[49] D. Dragoman and M. Dragoman, *Prog. Quant. Electron.*, 28 (2004) 1.
[50] D. Beglov and B. Roux, *J. Chem. Phys.*, 100 (1994) 9050.

Molecular and Nano Electronics: Analysis, Design and Simulation
J. M. Seminario (Editor)

Chapter 7

Nonequilibrium Green's function modeling of the quantum transport of molecular electronic devices

Pawel Pomorski, Khorgolkhuu Odbadrakh, Celeste Sagui,
and Christopher Roland

Center for High Performance Simulations (CHiPS) and Department of Physics,
North Carolina State University, Raleigh, NC 27695-8202,
USA. roland@gatubela.physics.ncsu.edu

1. Introduction

The recent advent of *molecular electronic systems* has opened up a new frontier, whose aim is the ultimate miniaturization of electronic systems [1–7]. The current–voltage (I-V) characteristics of such *atomic* and *molecular systems*, along with the more complex *Nanometer Electro-Mechanical Systems* (NEMS), hold the promise of revolutionary new devices for ultra-sensitive probes and detectors, very high speed and ultra-large density electronic components, and the possibilities of novel logic layouts. Indeed, demonstrations of molecular-based logic gates and nonvolatile random access memory have already been realized, and point to the exciting possibility of molecular computing [1, 2, 4]. However, many challenges remain before molecular electronic systems can become a viable technology. In very general terms, these include a fundamental understanding of molecular electronic systems, the synthesis and cost-effective assembly of nanostructures into desirable, functional arrangements, and the establishment of good, reliable macro-to-nanoscale contacts. In this short review, we focus on theoretical developments associated with the first of these challenges, aimed primarily at first principles approaches to the modeling of quantum transport for *real* molecular electronic systems.

Aside from the technological impetus, there is a great interest in understanding quantum transport from a fundamental point of view. For instance, at what length scales, and in what manner, do the macroscopic transport laws break down and quantum effects take over? Although the trend in device miniaturization has been going on for about 60 years [8], experimental answers to these questions have only become available during the 1980s and 1990s with the advent of *mesoscopic* [9], and now *nanoscale* systems [10]. Mesoscopic systems are typically in the 100 nm to 1 μm range, and were first studied intensely in the 1980s. The outstanding example here is the *two-dimensional* electron gas, which is a carefully manufactured system in which a small subset of electrons is trapped in a layer of a heterojunction between two semiconducting materials. The electrons move through in this layer over relatively long distances ($\sim$1 μm), while maintaining phase-coherence and without undergoing significant scattering [9]. While mesoscopic devices dominated the 1980s, the 1990s have seen the rise of molecular or nanoscale devices, which directly manipulate the electronic states of individual atoms and molecules to form the device. This field has benefited considerably from the development of self-organized structures such as carbon nanotubes [11], and other nanowires [12]. Devices such as nanotube-based field-effect transistors [13, 14], intramolecular metal-semiconductor diodes [15, 16], and intermolecular-crossed nanotube-nanotube diodes [17] have already been produced in the laboratory, and analyzed. Currently, the transport properties of similar structures based on different materials such as metals, semiconductors, polymer, organic compounds, and biomolecules are all the subject of intense scientific investigations [18].

In the field of molecular electronics, experimental work has clearly demonstrated that many of the important device characteristics relate specifically to the strong coupling between the *atomic* and the *electronic* degrees of freedom. Hence, the accurate prediction of the properties of atomic and molecular scale devices – including the true I-V curves with as few adjustable parameters as possible – still represents a formidable challenge. However, recent algorithmic advances related to the application of nonequilibrium Green's function (NEGF) methodology to the transport problem [18–21], combined with power of fast supercomputers, mean that theory is poised to make significant inroads into this important problem. The purpose of this chapter is to give a short review of this new methodology [20, 21], and use selected case studies to illustrate some of the very basic principles behind quantum transport [22–25].

2. Review of methodology

At the nanometer length scale, devices are typically much smaller than the mean-free path length of the electrons, which move ballistically through the system [18]. To illustrate the complexity of the quantum transport problem, consider a typical two-probe molecular device shown in Figure 1. The device consists of an Si_{10} cluster contacted to two metallic electrodes, assumed to be aligned along the z-axis, which extends to two electron reservoirs at $z = \pm\infty$, where the current is collected. A gate electrode, which couples capacitively with potential V_g to the molecule, may also be present. It is clear that accurate transport properties cannot be described semi-classically, but must be derived from a quantum mechanical treatment which, at the level of density functional theory (DFT), includes the atomic orbitals, the exchange-correlation interactions, the core-valence interactions, the coupling to the electrodes, and the effects of any externally imposed fields. Finally, the microscopic system must be coupled to macroscopic reservoirs in order to investigate the electron transport. Furthermore, if the electrochemical potentials $\mu_{l,r} + V_{l,r}$ with bias voltage $V_{l,r}$ of the left (l) and right (r) electrodes are not equal, then the device is actually in a nonequilibrium state.

To properly deal with transport, three aspects are essential. First, one must deal with an *open* quantum system within the DFT formalism. This is different from conventional

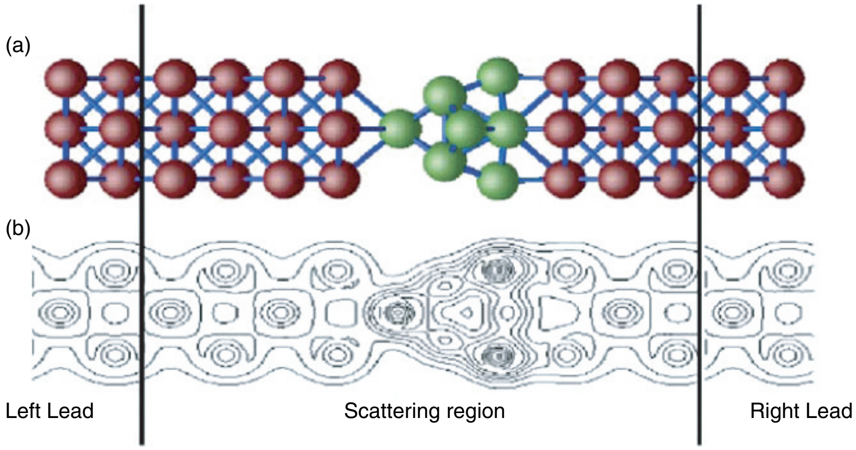


Figure 1 (a) Plot of the Al(100)/Si_{10} molecular device. (b) Contour plot of a slice of the equilibrium charge distribution. Note the almost perfect match across the boundaries of the scattering region and the electrodes

DFT simulation schemes, which are based on plane-wave [26] or real-space methods [27] which deal with a *finite* or *isolated* system, such as a collection of atoms or molecules or with a *periodic supercell* type of structure. By contrast, the device shown in Figure 1 has *open* boundaries provided by the long electrodes, which maintain a specific chemical potential due to the external bias voltages. Hence, one must find new ways of making this "infinite" open problem into a tractable finite-size problem. Second, one has to calculate the device Hamiltonian $\mathbf{H}[\rho(\mathbf{r})]$ within DFT using the correct charge distribution $\rho(\mathbf{r})$, which must be constructed under a finite bias with the correct open boundary conditions. For open systems, the charge distribution is constructed both out of the *scattering states* connecting $z = -\infty$ to $z = +\infty$ and out of the *bound states* which may exist within the molecular region. Finally, one needs an efficient numerical procedure in order to simulate molecular devices with a large number of atoms.

To date, all the theoretical approaches to the accurate modeling of quantum transport of molecular devices fall into four main categories: semi-empirical methods [28–33], supercell methods [33–38], open-jellium Lippman-Schwinger approach [39–45], and the recently developed NEGF approach [20, 21]. Fully recognizing the important contributions made by all of these approaches, we briefly outline their pros and cons along with a short discussion of pending challenges.

Semi-empirical [28–33] methods are typically *non-self-consistent*, and are based on parameterized tight-binding type of Hamiltonians for bulk and isolated molecular systems. These parameters, in general, cannot account for important factors such as the external bias or gate potentials. In addition, one can expect difficulties with the alignment of the Fermi levels of the electrodes and the molecular region, because true self-consistency is lacking. Semi-empirical methods are, however, relatively simple and easy to implement, which accounts for their widespread popularity. Indeed, in situations where there is relatively little charge transfer, they often provide good semi-quantitative insight. For instance, they have recently been used quite extensively in exploring transport through carbon nanotube–based systems [33].

Supercell methods [34–38, 46] are based on *ab initio* solutions to the Kohn–Sham (KS) equations using periodic boundary conditions. Once these are solved, the effective device potential is joined to perfect electrodes, and the scattering states determined via a recursive technique. This is a first principles method and, as such, it is parameter-free. However, because the effective potential is derived from a problem with periodic boundary conditions, these methods cannot describe systems with different electrodes and, more importantly, they cannot deal with systems under an external bias. These methods are for the most part suitable for calculating transport coefficients such as the conductance within the linear response regime.

In the Lippman-Schwinger-based open-jellium approach [39–45], the leads are described in terms of a jellium model. The KS equations are then solved self-consistently for the open structure, and the charge density is constructed from the scattering states of the device. As such, this method – which has been pioneered by Lang [39] – correctly describes the open boundary conditions, and may be used to calculate I-V curves. There are, however, two problems inherent in this approach which appear to be difficult to overcome. For molecular devices, it is not enough to simply use the scattering states to construct the charge density and the potential – the bound states which exist inside the device must also be accounted for in order to achieve a truly self-consistent solution. Second, the use of the jellium model for the leads does not account for any effects resulting from the realistic atomic structure of the leads. For instance, associated with a set of atomic leads will be a band-structure, which determines the wavevectors and energies of the electrons moving through the device. Clearly, a realistic description of molecular electronic systems must, in some fashion, account for the features of the leads.

The most recent advances in the theory of molecular electronics have been based on combining the NEGF formalism with DFT-based simulations [20, 21]. Over the past few years, this approach has been applied to an ever growing number of physical systems, including fullerenes, nanotubes, clusters, metallic nanowires, and organic compounds, primarily in a two-probe geometry [47–55]. Roughly speaking, the main advantages of the NEGF–DFT approach are: it allows for (i) a proper treatment of the open-boundary conditions for a quantum system under a bias voltage; (ii) a fully atomistic treatment of the electrodes; and (iii) a self-consistent calculation of the charge density via NEGFs, thereby incorporating the effects of both the bound and the scattering states in the system. Moreover, because this approach is based on real-space grids, the entire procedure may be parallelized enabling the treatment of large systems.

The key feature of this NEGF formalism is that the self-consistent charge density is not constructed out of the eigenstates of the system. Rather, the charge density is determined via the Keldysh NEGFs [18–21], which provides an efficient framework for dealing with an open quantum system. Roughly speaking, the self-consistent charge density ρ may be constructed from the NEGF $G^<$ via:

$$\rho = -\frac{\mathrm{i}}{2\pi}\int \mathrm{d}E G^{<}(E) \tag{1}$$

with

$$G^{<} = G^{R}\sum\nolimits^{<}[f_l, f_r]G^{A} \tag{2}$$

with $G^{\mathrm{R,A}}$ denoting the retarded/advanced Green's functions of the device. The quantity $\Sigma^{<} = -2\mathrm{i}\,\mathrm{Im}(f_{\mathrm{l}}\Sigma^{\mathrm{l}} + f_{\mathrm{r}}\Sigma^{\mathrm{r}})$ represents the injected charge from each of the electrodes, and is defined in terms of the self-energies (Σ) of each of the leads, and the distribution functions $f_{\mathrm{l,r}}$ describing the occupation of the eigenstates of the electrodes. Note that in the absence of a bias voltage, one can compute the linear conductance coefficients using the *equilibrium* Green's function G^{R}, instead of $G^{<}$. However, in the presence of a bias voltage, $G^{<}$ must be used because the system is *out of equilibrium*. For a system at equilibrium, all states below the the chemical potential μ will be filled, such that $f_{\mathrm{l}}(E) = f_{\mathrm{r}}(E)$. For those states, $\mathrm{Re}[G^{<}(E)] = 2\mathrm{Im}[G^{\mathrm{R}}(E)]$, so that:

$$\rho = \frac{2}{\pi}\int_{-\infty}^{\mu_{\min}} \mathrm{d}E G^{\mathrm{R}}(E) + \frac{1}{\pi}\int_{\mu_{\min}}^{\mu_{\max}} \mathrm{d}E G^{<}(E) \tag{3}$$

where $\mu_{\min} = \min(\mu_{\mathrm{l}} + V_{\mathrm{l}}, \mu_{\mathrm{r}} + V_{\mathrm{r}})$ and $\mu_{\max} = \max(\mu_{\mathrm{l}} + V_{\mathrm{l}}, \mu_{\mathrm{r}} + V_{\mathrm{r}})$. The evaluation of these integrals is complicated by the fact that they contain contributions from both the *scattering* states – i.e. eigenstates with a continuous spectrum which correspond to electrons with wave-functions extending infinitely into the leads – and the *bound* states, which are states of discrete energy with wave-functions localized in the central scattering region and decaying into the leads. Such bound states arise if the molecules in the central region have states with energies below the propagating threshold of the leads, if there are mismatches in the symmetries of the wave-functions, or when there are bandgaps present in both of the lead electrodes. The contributions of these bound states to the charge density may be significant, and generally speaking should not be ignored. To calculate their contribution, we note that it has been shown that G^{R} has poles at the bound state energies that lie below the real energy axis in the complex plane. Hence, G^{R} is analytic above the real axis. Thus, a convenient way of dealing with them is to integrate the first term of equation (3) along a semi-circle in the upper half plane of the complex plane, starting from some minimum energy that lies below all the states, and ending on the real axis at $\mu_{\min}$, as shown in Figure 2. Numerically, very accurate integration is achieved by means of Gaussian quadrature with a relatively modest number of points. The presence of bound states between $\mu_{\min}$ and $\mu_{\max}$ is actually problematic, giving rise to singularities in $G^{<}$ which manifest themselves as convergence problems. Fortunately, most systems investigated to date are free from

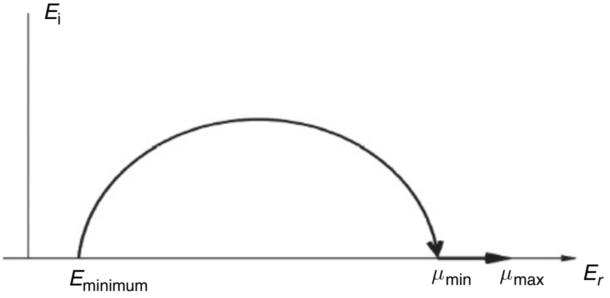


Figure 2 Schematic integration pathway in the complex plane used to evaluate the charge density

this problem, and may otherwise be treated with a special symmetry decomposition scheme [24].

To calculate the required quantities, the electronic wave-functions are represented with a minimal $\{s, p, d\}$ linear combination of atomic orbitals (LCAO) basis set [55], along with standard pseudopotentials [56]. This gives a description similar to that used for the SIESTA code [21], except that the terms are augmented to include the effects of the open system Poisson's equation with the proper boundary conditions. The latter is solved by means of standard multigrid methods [57]. Ultimately, such a description gives a set of tight-binding-like matrix elements for both the Hamiltonian and the overlap matrices. To handle the contributions of the leads, the respective self-energies are calculated by means of the method of Sanvito and coworkers [58]. Construction of the retarded (advanced) Green's functions from these quantities is standard, and proceeds by direct matrix inversion.

The current is then evaluated with the Landauer formula:

$$I = \frac{2e}{h} \int_{\mu_{\min}}^{\mu_{\max}} \mathrm{d}E(f_l - f_r)T(E) \tag{4}$$

where $T(E)$ is the transmission probability given by:

$$T(E) = 4Tr\left[\mathrm{Im}\left(\sum\nolimits^l\right) G^R \mathrm{Im}\left(\sum\nolimits^l\right) G^A\right] \tag{5}$$

It should be emphasized that since the current is calculated from a self-consistent analysis, the quantities inside the trace are all functions of the bias potential, so that *gauge-invariant* I-V curves are obtained. This is important because all the relevant physics of the devices can only depend on the voltage differences [59]. At equilibrium, the current is proportional to the conductance G:

$$I = G(\mu)(\mu_l - \mu_r) \tag{6}$$

which when evaluated at the Fermi level μ of the device gives $G(\mu) = (2e^2/h)T(\mu)$.

3. Transport through prototypical molecular electronic devices

To illustrate the power of the methodology, along with the main physics behind molecular transport, we briefly discuss four examples taken from primarily our own work [22–24]. Specifically, transport through small Si clusters and finite nanotubes, carbon nanotube capacitance, and an organic molecule that acts as a molecular diode will be considered.

As a first example. we consider the I-V characteristics of small Si-clusters in contact with two Al-electrodes, as shown in Figure 3. The clusters display *metallic* I-V characteristics with a linear regime about the origin and significant nonlinearities setting in at a higher bias. Specifically, for Si_4, there is evidence of strong negative differential resistance (NDR) with the current decreasing with increasing bias voltage. For the Si_7 clusters, this effect is considerably less pronounced. How can one explain the

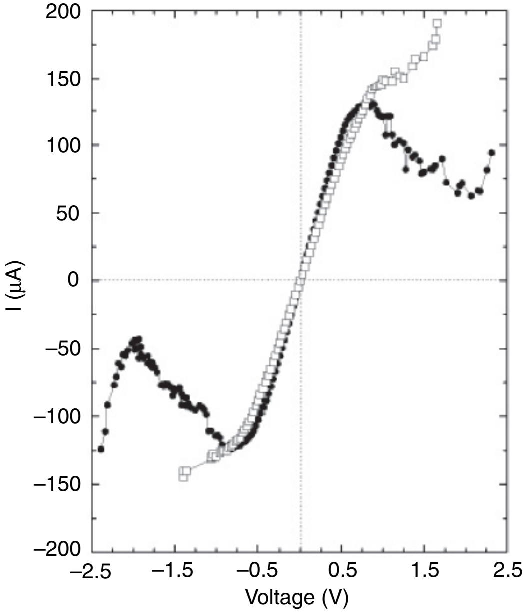


Figure 3 Sample I-V curves for Si nanoclusters between Al-leads: filled circles, Si_4; squares, Si_7

shape of these I-V curves? To understand these, we have analyzed the transmission coefficients $T(E, V_b)$, which give us the contribution to the current as a function of the electron energy and bias voltage. The shape of this function is determined by two concepts, namely the band structure of the leads and the renormalized molecular levels that mediate the transmission process.

Figure 4 shows $T(E)$ (right panel) for both Si_4 and Si_7, along with the band structure (left panel) of the (100)Al for wavevectors in the device direction. On the band structure, the wave vectors that couple to the scattering states of the system are marked with filled circles, such that the size of the circles indicate the relative importance of that state to the transmission process. What is observed? For Si_4, there are two bands that significantly contribute to the transmission near the Fermi level, which here has been shifted to $E = 0$. The contributions from the other bands are considerably less, and hence $T(E)$ takes on a value greater than two (units G_o). Similarly, for Si_7, more bands contribute to the transmission giving a larger value of ~3.5. The linearly increasing part of the I-V curves is now readily understood: as the bias voltage increases, the relative position of the band structure of the left- and right-electrodes are shifted apart by the bias. Between $E = 0$ and ~0.46 eV, the current increases because more and more conduction density of states are included in the integration window. However, for $V > 0.46$ eV, there are no additional conduction density of states because no scattering states in that energy region couple to the cluster. Clearly, $T(E)$ must decrease precipitously at this point, and the current begins to drop. The system therefore displays NDR unit ~2 eV, when transmission can take place through other subbands and the current begins to recover. This behavior is to be contrasted with that of the Si_7 cluster. For that cluster, there

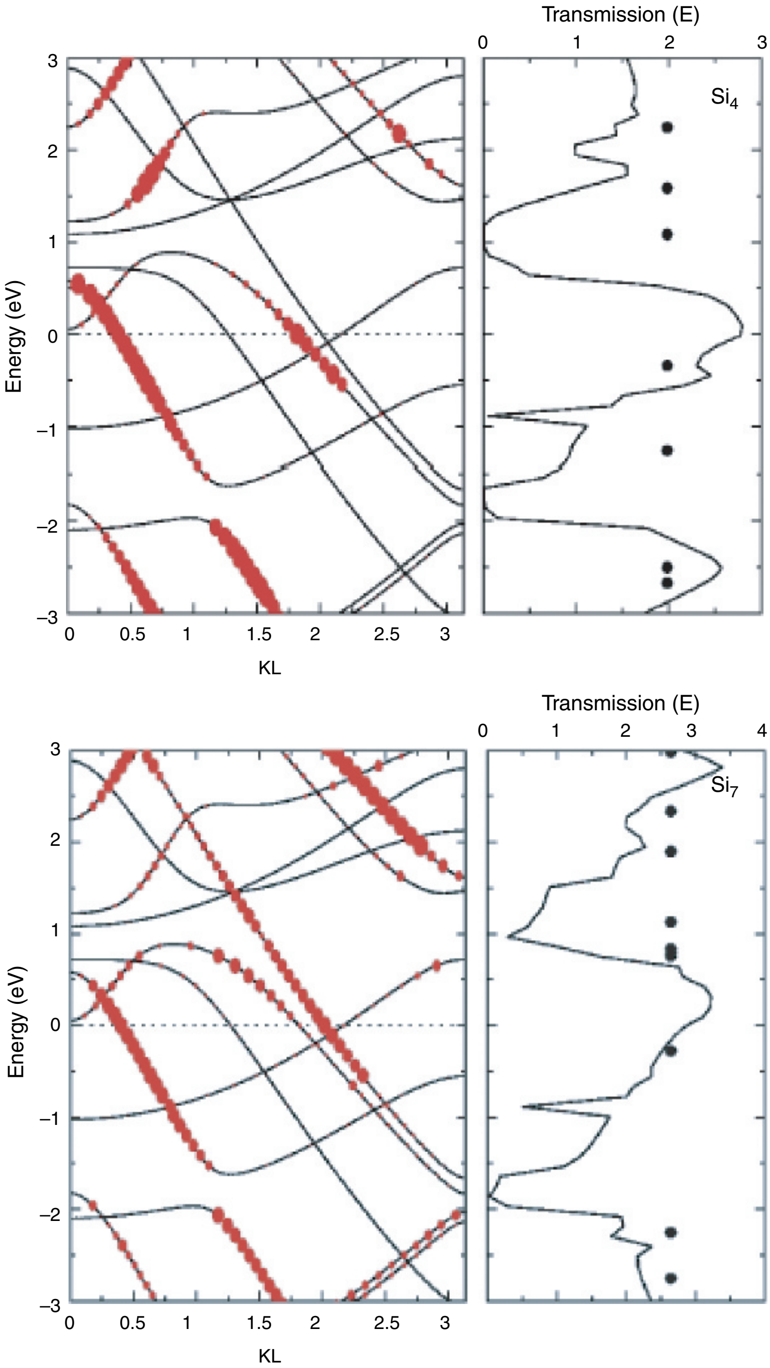


Figure 4 Bandstructure of Al(100) electrodes (left panel) and $T(E)$ (right panel). On the bandstructure, the wavevectors corresponding to the scattering states are marked with filled circles, with the size indicating their relative importance for the transmission process. On the right panel, the RMLs are marked with filled circles

is always at least one band that conducts significantly. Hence, although $T(E)$ does decrease significantly, it does not go to zero. This decrease in T manifests itself through a decrease in the slope of the I-V curves. These considerations clearly emphasize the importance of keeping a fully atomistic description of the leads.

The aspect that determines the shape of $T(E)$ are the molecular levels that mediate the transmission process. First, consider an isolated molecular cluster. For such a system, it is relatively easy to find the molecular levels by simply diagonalizing the Hamiltonian of the system, thereby identifying the HOMO/LUMO levels. These may or may not, however, be exactly the levels relevant to the transmission process. Because of the coupling between the electrodes and the molecule, the character of these levels will change considerably in ways that are not known *a priori*. Furthermore, the energy spectrum for the open quantum system will be continuous and broadened because of the contacts to the leads. To understand these changes, we have focused on calculating the renormalized molecular levels (RMLs), which we have approximated as follows. After a self-consistent field calculation of the Kohn-Sham equations is completed, one obtains the self-consistent potential and the matrix elements of the Hamiltonian. By diagonalizing the *submatrix* of the Hamiltonian associated with the molecule (e.g., the Si-clusters), one obtains a set of new levels – the RMLs – which are characteristic of the changes induced in the molecular levels of the isolated cluster through the coupling to the leads. These RMLs are plotted as filled circles on the $T(E)$ of Figure 4. Roughly speaking, the presence of peaks in $T(E)$ correspond to the location of an RML. Hence the position of an RML marks a broad region of energy over which there is significant coupling between the electronic states of the clusters and the leads. Generally speaking, the six states closest to the HOMO/LUMO gap of the original molecular levels of the isolated clusters provide the majority of the contribution to the RMLs.

The second example we will consider is that of short carbon nanotube, which will serve to illustrate the important role played by *evanescent* waves for molecular electronic systems [23, 60]. The device consists of an (8, 0) single-walled carbon nanotube between two 5×5 Al(100) electrodes. Figure 5 shows the calculated I-V curves for four semiconducting nanotubes of increasing length. The short tubes display more or less linear behavior, characteristic of metallic systems. As the length of the nanotube is increased, the I-V curves flatten and the system acquires much more of a semiconducting nature. To gain further insight into the transport, Figure 5 also shows the transmission spectra of these devices. This is a somewhat complicated function, reflecting both the bandstructure of the leads and the nanotube levels that mediate the transport. As a generic feature, it is evident that $T(E)$ does develop a gap about the Fermi level, that progressively widens and deepens as the nanotube length is increased.

To understand the origins of these features, we consider the electronic states of these devices. As the central part of the nanotube becomes very long, one can expect that the electronic states in this region approach that of a very long, periodic nanotube without any Al-leads. An examination of the coupling between the nanotube and the electrodes does indeed show that the most significant perturbations are confined to the atomic planes adjacent to interface: aside from the interface dipole, only a small amount of charge flows into the nanotube.

Thus, away from the edges of the leads, the system is approximately translationally invariant, and so the electronic states of this device should be similar to states of an infinitely long nanotube with a similar charge distribution. This has been checked

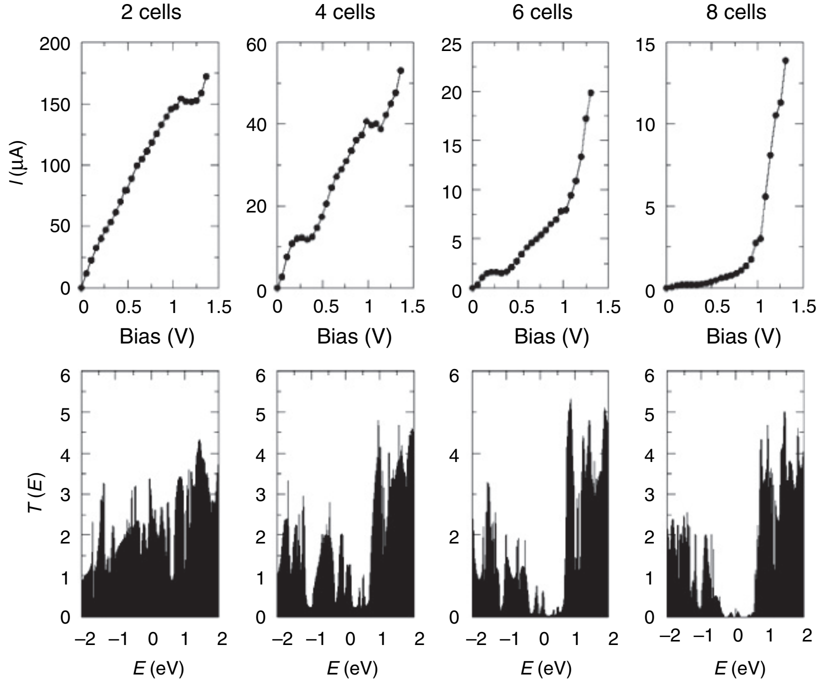


Figure 5 I-V curves (top panel) for nanotube device, consisting of 2–8 unit cells, respectively. The transmission spectrum for each corresponding device is also shown (bottom panel). Note the development of the gap about $V = 0$, as the number of unit cells is increased

explicitly through a calculation of the renormalized bandstructure of the system. Given a self-consistent Hamiltonian matrix, we select the submatrix corresponding to the different unit cells of the device, and calculate the bandstructure by standard means. This renormalized bandstructure reflects the influence of the leads, the bias voltage, and other external effects. Generally, within a constant energy shift, there is excellent agreement between the shape of the renormalized bandstructure and that of an infinite carbon nanotube. What is important here is that this bandstructure retains the characteristic of a semiconducting system with no propagating modes near the Fermi energy. This suggests the presence of evanescent modes inside the semiconducting nanotube, with very long decay lengths responsible for the metallic transport characteristics of the finite-sized nanotubes.

These evanescent modes give rise to a finite density of states inside the nanotube, which acquire charge from the Al-electrodes forming metal induced gap states (MIGS). The presence of these MIGS is signalled by nonzero values of $T(E)$ in the gap region, as shown in Figure 5. Of course, for a very long nanotube coupled perfectly to ideal leads, transmission would be given by the number of propagating modes present at any given electron energy. Evanescent modes, which decay a finite distance away from the electrodes, would then not play a role. In contrast, for the finite-sized nanotube system considered here, the evanescent modes play a crucial role because of tunneling through

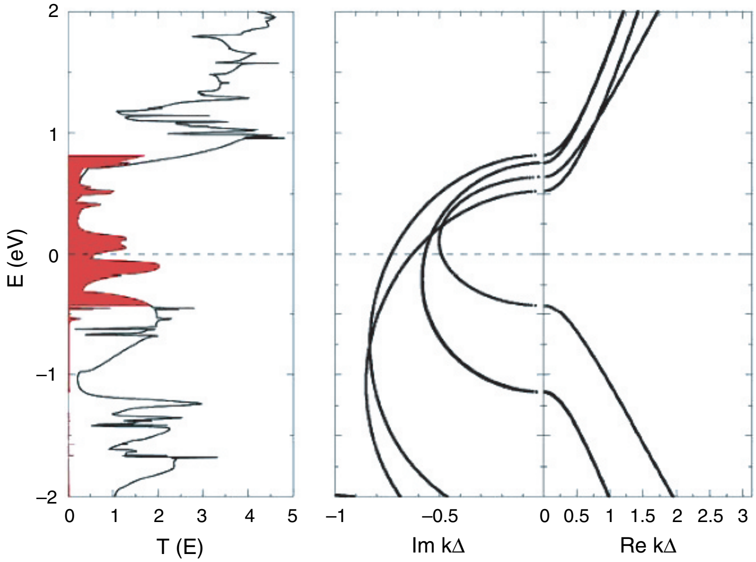


Figure 6 Transmission and complex band structure for a nanotube device with a length of four unit cells. Because of the symmetry, only the negative branch of the complex band structure is shown. The contributions of the evanescent bands to $T(E)$, which dominate the gap region, are shaded

these modes. Clearly, their contributions will be most important inside the band gap, where no propagating modes are present.

To quantify this aspect, we have decomposed $T(E)$ into its contributions from the different modes, and plotted the results along in Figure 6. This figure clearly shows that the evanescent mode contribution dominates the transmission process accounting for almost 100% of $T(E)$ within the gap region. The origins of the evanescent modes may be understood in terms of the complex band structure of the nanotubes. The evanescent modes are associated with the imaginary branch of the complex bandstructure. Essentially, the complex branch consists of a pair of "loops", that "connect" different valence and conduction bands of the real branch of the bandstructure. For the (8, 0) nanotube, $\mathrm{Im}(k)$ near the Fermi level is about $k \sim 0.3$ (in units of the inverse unit cell length Δ). This is important, because the decay length of the evanescent modes may be estimated as $\lambda \sim 2\pi/k \sim 6$–$10$ unit cells, in agreement with the numerical data.

We now turn to the case of an polyphenylene-based molecule (Figure 7), which can be shown to act as a molecular diode. The polyphenylene molecule consists of two parts: a donor and an acceptor part, each with their own molecular levels. The two parts of this organic molecule are separated from each other by means of a central CH_2 pair, which acts as an effective electronic barrier. The molecule is attached to two Au leads. Figure 8 shows the calculated I-V curves, which show that the device acts as an effective molecular diode. Thus, the current is mostly small, until the bias voltage approaches $-2\,\mathrm{V}$, at which point the current begins to rise steeply. The workings of this device can be understood from a straightforward analysis of the molecular levels of the device. For the most part, the donor and the acceptor portions of the molecule

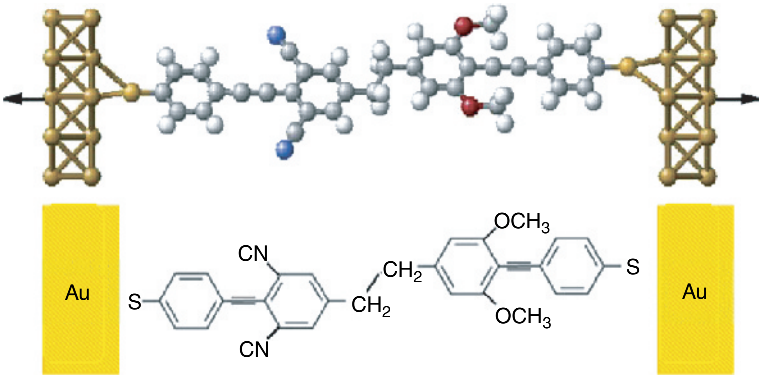


Figure 7 Schematic of a polyphenylene-based molecule that acts as a molecular diode. Note the central pair of CH_2, which separates the donor and the acceptor parts of the molecule

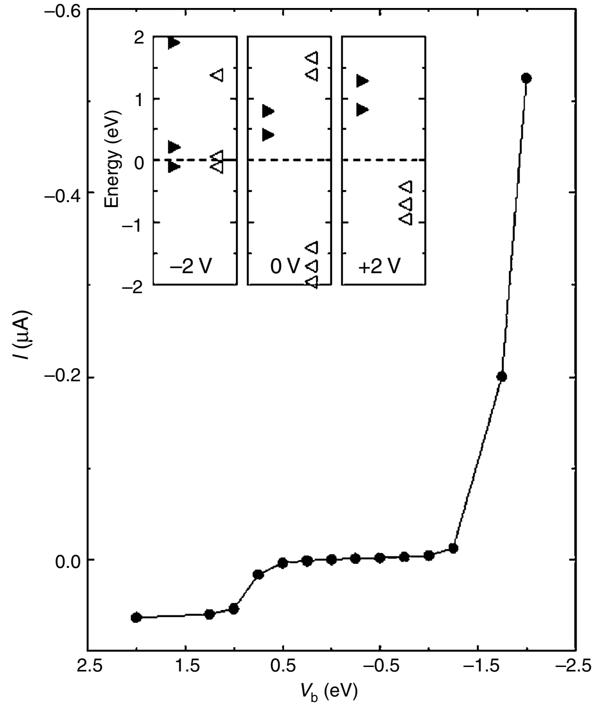


Figure 8 I-V characteristics of the molecular diode shown in the previous figure. The inset shows the positioning of the most important molecular levels of the two parts of the molecule for different bias voltages. Note that the levels align and current flows when $V_b = -2\,V$; otherwise, they do not align and current is low

have levels that are well separated from each other. Hence, electrons tunneling through the CH_2 barrier have nowhere to go, and so the current is small. However, under the influence of a bias of about $-2\,\mathrm{V}$, the levels are shifted into alignment. Electrons can therefore tunnel through these levels and then into the device electrodes signaled by the subsequent current rise.

As a final illustrative example, we discuss recent investigations of the capacitance of carbon nanotube systems [24]. This problem differs from the previous examples insofar that it depends on the *induced rearrangement of charges*, rather than on the direct flow of current. For general nanoscale systems, it has been theoretically predicted [61–63] and experimentally verified [64] that the capacitance of molecular scale conductors shows distinct, nonclassical behavior, as previously anticipated from general considerations [63]. Specifically, at the nanoscale, the screening length of the material is often comparable to the dimensions of the system, so that the classical concept of electrostatic capacitance is no longer adequate. Instead, one makes use of the notion of the the electrochemical capacitance $\mathrm{C}_{\alpha\beta} = \mathrm{e}(\mathrm{d}Q_\alpha/\mathrm{d}\mu_\beta)$, which is a measure of the charge variation $\mathrm{d}Q_\alpha$ on conductor α when the electrochemical potential of the reservoir connected to conductor β is changed by a small amount $\mathrm{d}\mu_\beta$. $C_{\alpha\beta}$ takes quantum effects into account, and may behave qualitatively and quantitatively very different when compared to its classical counterpart [64].

To calculate the capacitance, use was made of the NEGF–DFT-based formalism. However, here the focus is on calculating the charge as a finite bias is applied to the reservoirs, i.e., $\Delta Q = Q(V + V_\mathrm{b}) - Q(V)$, and then calculate the capacitance from $V_\mathrm{b} = e\mathrm{d}\mu$. As a further feature, it has been useful to use Dirichlet boundary conditions for the electrostatic potential at the walls of the finite-sized calculational box, which corresponds to surrounding the system with a metal container. This serves to terminate any field lines which may escape from the nanotubes [65]. For a container infinite in size, the results would reduce to that of a system in free space.

Figure 9 shows an example of a two-probe system [24], in which a capped (5, 5) nanotube is inserted at a finite distance into an open (12, 12) tube, with the central axis coinciding. The system now consists of two semi-infinite nanotube leads, and a central region containing the junction of the two tubes. Note that distance between the carbon atoms is quite large – at least 9.1 A, so that the tubes are *de facto* separated, with no dc current flowing between them. Of course, the tubes are coupled through influence of the electrostatic potential. Details of the charging of the nanotubes, along with a histogram of the change in the total charge accumulated on each nanotube ring, is shown in Figure 9. In particular, it is noted that the (12, 12) tube acquires a very large amount of charge on its terminal ring due to the presence of the dangling bonds.

For this particular nanotube case, the charge accumulation is essentially linear as a function of the bias voltage. The calculated values of the capacitance matrix are $C_{(5,5)(5,5)} = 0.105\,\mathrm{aF}$, $C_{(12,12)(5,5)} = -0.455\,\mathrm{aF}$, $C_{(5,5)(12,12)} = -0.451\,\mathrm{aF}$, and $C_{(12,12)(12,12)} = 0.156\,\mathrm{aF}$. Figure 9 also displays coefficient $C_{(5,5)(12,12)}$ as a function of the nanotube insertion depth, which ultimately matches that of two nested carbon nanotube shells. From the figure, it is also evident that nanotube, to which the bias is applied, gains charge along its entire length. This "self-charging" is due to the capacitive coupling between the nanotube and the surrounding metal container. It therefore follows, that the values of $C_{(\mathrm{m,m})(\mathrm{m,m})}$ calculated will increase linearly in size, as more and more of the

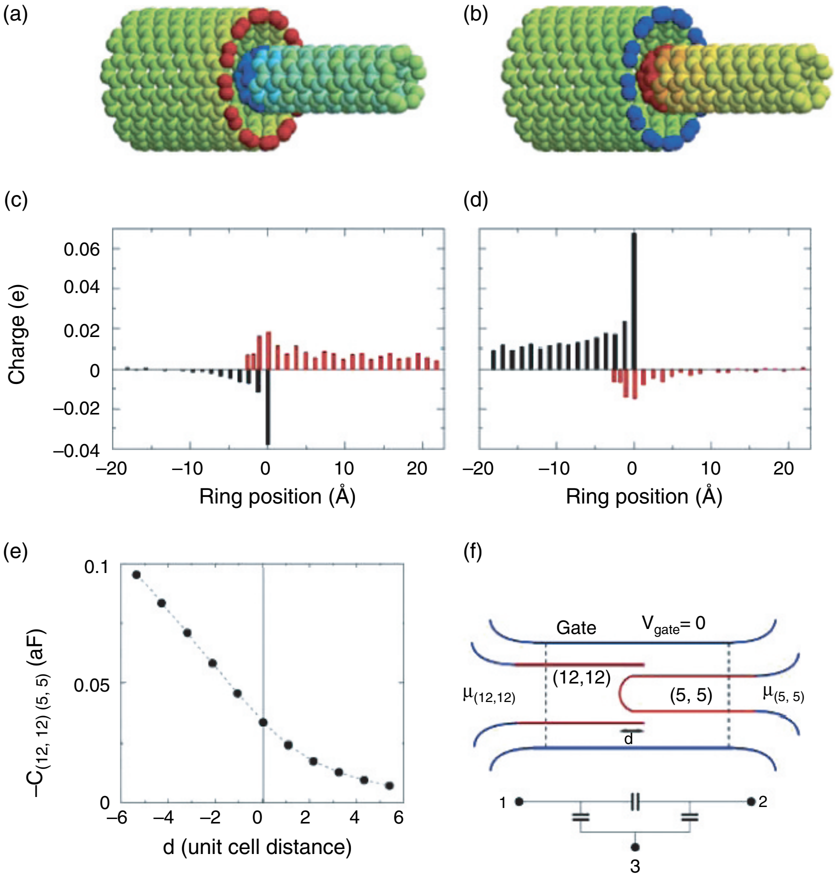


Figure 9 (a–d) Nanotube charging of the (12, 12)/(5, 5) biased junction for a central simulation box of 458 atoms. A +0.272 eV bias is applied on the right (5, 5) tube in (a, c), and on the left (12, 12) tube in (b, d). Upper panels (a, b) show charge accummulation with a temperature color scheme (blue is charge addition, red is charge depletion). Lower panels (c, d) display histogram plot of the charge accummulated on the nanotube rings, with the black (red) bars showing charge accummulation on the (12, 12) capped (5, 5) tubes, respectively. Panel (e) shows capacitance matrix element as a function of the relative position d of the two nanotubes. The origin $d = 0$ is chosen such that the longitudinal positions of the terminal rings of both tubes coincide, and negative d indicates the overlap between the two tubes. Finally, (f) illustrates the computational configuration and the equivalent circuit device

charge density is included. To understand this, it is useful to consider the equivalent circuit of the device shown in Figure 9(f). In addition to the capacitance between the nanotubes in the central cell, there are the additional capacitances between the nanotubes and the container. For a system enclosed in an infinitely large box, the capacitive coupling between the leads/box will vanish, and all charge variations will occur in the junction only. All capacitance matrix elements would then be equal in magnitude. However, finite-size investigations show that the values of the $C_{(m,m)(n,n)}$ capacitance coefficients are in reasonable agreement with their large-box limits. In addition to the investigation of the two-probe nanotube system presented here, the capacitance of nested concentric nanotubes, a special (12, 0)/(6, 6) nanotube junction, and the action of nanotubes as probes over Al (100) surfaces have all been considered [24].

4. Current theoretical developments

The NEGF–DFT formalism for quantum transport is poised to make important contributions towards calculating the device properties of real materials systems, especially when combined with the power of parallel supercomputers. While the first implementations discussed in this review have focused on issues related to the I-V curves of two-probe devices, one can expect a host of new developments such that calculations based on *three-* and *four*-probe molecular devices, current-induced forces, spin-dependent, and time-dependent devices may all be expected to emerge shortly. Here, we briefly mention a few important issues related to the last two developments.

The incorporation of *spin-dependent* effects opens up the important field of nanoscale magnetoelectronics or "spintronics" [66]. In magnetoelectronics, both the *charge* and the *spin* degrees of freedom are utilized for the operation of a functional device. Such effects are of course extremely important from technological point of view. Spintronic devices have already been used to construct nonvolatile random access memory, magnetic tunnel junctions, and magnetic transistors. Clearly, the ability to investigate the functionality of such devices at the nanoscale will prove to be an important future direction.

In discussing devices, it has tacitly been assumed that we are dealing with steady-state dc currents through molecular systems. However, the *dynamic response* of molecular devices is also significant in its own right. Time-dependent information is, of course, an intrinsic part of nonequilibrium devices such as electron pumps and turnstiles. Dynamic information also plays an important role when discussing high-speed applications, the nonequilibrium charge distributions, dynamic couplings, and simply the transport dynamics of a conductor. In this case, the problem is complicated by the fact that time-dependent electromagnetic fields can take the system out of equilibrium, with the possibility of inducing *displacement* currents. The inclusion of these displacement currents is absolutely crucial, as only then can the total current be conserved and the system display gauge invariance. Recently, the NEGF formalism has been extended to cover the case of the dynamic conductance [67]. This formalism shows that the dynamic conductance develops both real and imaginary frequency-dependent (ω) components, i.e., $G(\omega, V_b) = \mathrm{d}I(\omega, V_b)/\mathrm{d}V_b$, which may be of a *capacitive* or *inductive* nature, depending on the frequency. To date, this formalism has only been used to investigate transport through carbon nanotubes in the wideband limit, and hence almost nothing is known about the dynamic response of other molecular electronic systems. Such investigations may be expected to emerge in the near-term future.

5. Summary

In summary, we have briefly reviewed recent theoretical developments for simulating the quantum transport properties of *real* materials systems. The method is based on combining the Keldysh NEGF formalism with the power of DFT simulations, and has the advantages of: properly treating *open* quantum systems, allowing for a full atomistic treatment of the leads, the *self-consistent* calculation of the charge density, and – when combined with the power of state-of-the art supercomputers – the ability to treat large-scale molecular electronic systems. This formalism has proven itself to be reliable, and is believed to have full predictive power within the context of DFT. Hence, it is not only being applied to an ever growing list of materials system, but also the main product of new simulation-based venture capital firm *Atomistix* [68].

Acknowledgements

We gratefully acknowledge the financial support from NSF Grant No.0312105 and DOE.

References

[1] C. P. Collier, E. W. Wong, M. Belohradsky, F. M. Raymo, J. F. Stoddart, P. J. Kuekes, R. S. Williams, and J. R. Heath, *Science* **285**, 391 (1999); C. P. Collier, G. Mattersteig, E. W. Wong, Y. Luo, K. Beverly, J. Sampaio, F. M. Raymo, J. F. Stoddart, and J. R. Heath, *Science* **289**, 1172 (2000).

[2] M. A. Reed, C. Zhou, and C. J. Muller, *Science* **278**, 252 (1997); J. Chen, M. A. Reed, and A. M. Rawlett, *Science* **286,** 1550 (1999).

[3] C. Joachim, J. K. Gimzewski, R. R. Chlitter, and C. Chavy, *Phys. Rev. Lett.* **74**, 2102 (1995); J. K. Gimzewski and C. Joachim, *Science* **283**, 1683 (1999).

[4] T. Rueckes, K. Kim, E. Joselevich, G. Y. Tseng, C. L. Cheung, and C. M. Lieber, *Science* **289**, 94 (2000).

[5] S. J. Tans, M. H. Devoret, R. M. Alwin, and H. Dai, *Nature* **386,** 474 (1997); S. J. Tans, J. Verschueren, R. M. Alwin and C. Dekker, *Nature* **393,** 49 (1998).

[6] J. W. G. Wildoeer, L. C. Venema, and A. G. Rinzler, *Nature* **391**, 59 (1998).

[7] Y. Xue, S. Datta, S. Hong, R. Reifenberger, J. L. Henderson, and C. P. Kubiak, *Phys. Rev. B* **59**, 7852 (1999).

[8] Semiconductor Industry Association, *International Technology Roadmap for Semiconductors* (2001), http://public.itrs.net.

[9] See, for example, D. K. Ferry and S. M. Goodnick, *Transport in Nanostructures*, Cambridge University Press (Cambridge, 1997).

[10] See, for example, J. R. Heath and M. Ratner, *Physics Today* **53**, Vol. 5, 43 (2003).

[11] See, for example, M. S. Dresselhaus, G. Dresselhaus, and P. C. Eklund, *Science of Fullerenes and Carbon Nanotubes* (Academic Press, 1996); J. Bernholc, D. Brenner, M. Buongiorno Nardelli, V. Meunier, and C. Roland, *Annu. Rev. Mater. Res.* **32**, 347 (2002).

[12] See, for example, Y. Cui and C. Lieber, *Science* **291**, 851 (2001).

[13] J. Tans, R. M. Verschueren, and C. Dekker, *Nature* **393**, 49 (1998).

[14] R. Martel, T. Schmidt, H. R. Shea, T. Hertel, and P. Avouris, *Appl. Phys. Lett.* **73**, 2447 (1998).

[15] J. Hu, M. Ouyang, P. Yang, and C. Lieber, *Nature* **399**, 48 (1999).

[16] Z. Yao, H. W. C. Postma, L. Balents, and C. Dekker, *Nature* **402**, 273 (1999).

[17] M. S. Fuhrer, *et al.*, *Science* **288**, 494 (2000).

[18] S. Datta, *Electronic Transport in Mesoscopic Systems* (Cambridge University Press (1995).

[19] A. P. Jauho, N. S. Wingreen and Y. Meir, *Phys. Rev. B* **50**, 5528 (1994).

[20] J. Taylor, H. Guo and J. Wang, *Phys. Rev. B* **63**, RC 121104, 245407 (2001).

[21] M. Brandbyge, J. L. Mozos, P. Ordejon, J. Taylor and K. Stokbro, *Phys. Rev. B* **65**, 165401 (2002).
[22] C. Roland, V. Meunier, B. Larade, and H. Guo, *Phys. Rev. B* **66**, 035332 (2002).
[23] P. Pomorski, C. Roland, and H. Guo, *Phys. Rev. B* **70**, 115408 (2004).
[24] P. Pomorski, C. Roland, H. Guo and J. Wang, *Phys. Rev. B* **67**, RC 161404 (2003); P. Pomorski, L. Pastewka, C. Roland, H. Guo and J. Wang, *Phys. Rev. B* **69**, 115418 (2004).
[25] C. C. Kaun, B. Larade, and H. Guo, *Phys. Rev. B* **67**, 121411 (2003).
[26] R. Car and M. Parrinello, *Phys. Rev. Lett.* **55**, 2471 (1985).
[27] E. L. Briggs, D. Sullivan, and J. Bernholc, *Phys. Rev. B* **62**, 5471(R); *Ibid.* **54**, 1462 (1996).
[28] P. L. Pernas, A. Martin-Rodero, and F. Flores, *Phys. Rev. B* **67**, 121411 (1990).
[29] W. Tian and S. Datta, *Phys. Rev. B* **49**, 5097 (1994).
[30] A. L. V. de Parga, O. S. Hernan, R. Miranda, A. Yeyati, N. Mingo, A. Martin-Rodero, and F. Flores, *Phys. Rev. Lett.* **80**, 357 (1998).
[31] E. G. Emberley and G. Kirczenow, *Phys. Rev. B* **60**, 6028 (1999).
[32] M. Brandbyge, N. Kobayashi, and M. Tsukada, *Phys. Rev. B* **60**, 17064 (1999).
[33] H. Mehrez, J. Taylor, H. Guo, J. Wang, and C. Roland, *Phys. Rev. Lett.* **84**, 2682 (2000); C. Roland, M. Buongiorno Nardelli, J. Wang, H. Guo, Y. Wei, and H. Mehrez, *Phys. Rev. B* **63**, 195412 (2001); D. Orlikowski, H. Mehrez, J. Taylor, H. Guo, J. Wang, and C. Roland, *Phys. Rev. B* **63**, 155412 (2001).
[34] C. C. Wan, J. -L. Mozos, G. Taraschi, J. Wang, and H. Guo, *Appl. Phys. Lett.* **71**, 419 (1997); C. C. Wan, J. L. Mozos, J. Wang, and H. Guo, *Phys. Rev. B* **55**, 13393 (1997).
[35] J. Wang, H. Guo, J. -L. Mozos, C. C. Wan, G. Taraschi, and Q. Zheng, *Phys. Rev. Lett.* **80,** 4277 (1998); G. Taraschi, J. -L. Mozos, C. C. Wan, H. Guo, and J. Wang, *Phys. Rev. B* **58**, 13138 (1998).
[36] J. L. Mozos, C. C. Wan, G. Taraschi, J. Wang, and H. Guo, *Phys. Rev. B* **56**, 4351(R) (1997); J. L. Mozos, C. C. Wan, G. Taraschi, J. Wang, and H. Guo, *J. Phys. Cond. Matter* **10**, 2663 (1998).
[37] H. J. Choi and J. Ihm, *Phys. Rev. B* **59**, 2267 (1999).
[38] H. J. Choi, J. Ihm, Y. Y. Yoon, and S. G. Louie, *Phys. Rev. B* **60**, 14009 (1999).
[39] N. D. Lang, *Phys. Rev. B* **52**, 5335 (1995); A. Yazdani, D. M. Eigler, and N. Lang, *Science* **272**, 1921 (1996).
[40] K. Hirose and M. Tsukada, *Phys. Rev. B* **51**, 5278 (1995).
[41] U. Landman, R. N. Barnett, A. G. Scherbakov, and Ph. Avouris, *Phys. Rev. Lett.* **85**, 1958 (2000).
[42] N. D. Lang and Ph. Avouris, *Phys. Rev. Lett.* **81**, 3515 (1998).
[43] M. Di Ventra, S. T. Pantelides, and N. D. Lang, *Phys. Rev. Lett.* **84**, 979 (2000).
[44] N. D. Lang and Ph. Avouris, *Phys. Rev. Lett.* **81**, 3515 (1998).
[45] S. N. Yaliraki, M. Kemp, and M. A. Ratner, *J. Am. Chem. Soc.* **121**, 3428 (1999).
[46] H. J. Choi, J. Ihm, S. G. Louie, and M. L. Cohen, *Phys. Rev. Lett.* **84**, 2917 (2000).
[47] B. Larade, J. Taylor, and H. Guo, *Phys. Rev. B* **64,** 075420 (2001).
[48] C. Roland, B. Larade, J. Taylor and H. Guo, *Phys. Rev. B* **65**, RC 041401 (2002).
[49] J. Taylor, M. Brandbyge, and K. Stokbro, *Phys. Rev. Lett.* **89**, 138301 (2002).
[50] J. L. Mozos, P. Ordejon, M. Brandbyge, J. Taylor, and K. Stokbro, *Nanotechnology* **13**, 346 (2002).
[51] K. Stokbro, J. L. Mozos, P. Ordejon, M. Brandbyge, and J. Taylor, *Comp. Mater. Sci.* **27**, 151 (2003).
[52] H. Mehrez, A. Wlasenko, B. Larade, J. Taylor, P. Grutter, and H. Guo, *Phys. Rev. B* **65**, 195419 (2002).
[53] B. Larade, J. Taylor, Q. R. Zheng, H. Mehrez, P. Pomorski, and H. Guo, *Phys. Rev. B* **64**, 195402 (2002).
[54] C.-C. Kaun, B. Larade, H. Mehrez, J. Taylor, and H. Guo, *Phys. Rev. B* **65**, 205416 (2002); C.-C. Kaun, B. Larade, and H. Guo, *Phys. Rev. B* **67**, 121411 (2003).

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

[56] D. R. Hamann and M. Schutter and C. Chiang, *Phys. Rev. Lett.* **43**, 1494 (1982).

[57] A. Brandt, *Math. Comput.* **31**, 333 (1977).

[58] S. Sanvito, C. J. Lambert, J. H. Jefferson, and A. M. Bratkovsky, *Phys. Rev. B* **59**, 11936 (1999).

[59] M. Buttiker, *J. Phys. Condens. Matter* **5**, 9361 (1993).

[60] J. Tomfohr and O. Sankey, *Phys. Rev. B* **65**, 245105 (2002).

[61] T. P. Smith, B. B. Goldberg, P. J. Stiles, and M. Heiblum, *Phys. Rev. B* **32**, 2696 (1985); T. P. Smith, W. J. Wang, and P. J. Stiles, *Phys. Rev. B* **34**, 2995 (1986).

[62] M. Buttiker, H. Thomas, and A. Petre, *Phys. Lett. A* **100**, 364 (1993); M. Buttiker, *J. Cond. Matter* **5**, 9361 (1993).

[63] T. Christen and M. Buttiker, *Phys. Rev. Lett.* **77**, 143 (1996).

[64] J. G. Hou, B. Wang, J. Yang, X. R. Wang, H. Q. Wang, Q. Zhu, and X. Xiao, *Phys. Rev. Lett.* **86**, 5321 (2001).

[65] R. Tamura, *Phys. Rev. B* **64**, 201404 (2001).

[66] See, for example, G. A. Prinz, *Science* **282**, 1660 (1998).

[67] B. G. Wang, J. Wang, and H. Guo, *Phys. Rev. Lett.* **82**, 398 (1999).

[68] See, http://www.atomistix.com.

Molecular and Nano Electronics: Analysis, Design and Simulation
J. M. Seminario (Editor)

Chapter 8

The *gDFTB* tool for molecular electronics

A. Pecchia[a], L. Latessa[a], A. Gagliardi[a,b], Th. Frauenheim[b]
and A. Di Carlo[a]

[a]*Università di Roma "Tor Vergata", Via del Politecnico 1, 00133, Roma, Italy.*
dicarlo@ing.uniroma2.it
[b]*Dept. of Theoretical Physics, University of Paderborn, D-33098 Paderborn, Germany*

1. Introduction

The idea of molecular electronics now dates back three decades [1]; however, the problems that need to be addressed in order to make a nice principle into a real working device are still quite challenging, due to the difficulty of precisely controlling the molecular building blocks and stability issues related to organic materials. Great progresses have been made in recent years in these directions [2–4], and the field of molecular electronics is steadily progressing as new experimental breakthroughs are achieved. Conduction through single molecules can now be routinely obtained by several groups, particularly in break-junctions configurations.

Fingerprints of molecular species can unequivocally be detected in tunneling experiments by using inelastic tunneling spectroscopy (IETS). Such technique not only provide conclusive evidence that electrons are really tunneling through the molecular bridges [5, 6], but could provide in future a valuable characterization tool providing useful information about the interface structure, morphology and molecular geometry.

Despite many successes, the understanding and controlling of transport mechanisms is far from the understood. From a theoretical point of view, the description of conduction in molecular systems is also quite a challenge, as subtle many-electron correlations may play significant role, changing transport to the coherent regime to incoherent or Coulomb blockades. Highly accurate quantum chemical methods are far too computationally expensive to treat the very large clusters required to simulate realistic electrodes (hundreds of atoms). On the other hand, semi-empirical methods are often unable to catch the real behavior of molecular systems between electrodes.

Certainly, molecular devices require new simulation approaches, since the inherent quantum-mechanical physics involved must be treated properly [7]. The *gDFTB* approach for transport computations is based on the density functional tight-binding (DFTB) method [8], extended to the non-equilibrium Green's functions (NEGF) for the self-consistent computation of charge density and electronic transport [9].

The *gDFTB* method allows a nearly first-principle treatment of systems comprising a large number of atoms. The Green's function technique enables the computation of the tunneling current flowing between two contacts in a manner consistent with the open boundary conditions that naturally arise in transport problems. On the other hand, the NEGF formalism allows to compute the charge density consistently with the non-equilibrium conditions in which a molecular device is driven when biased by an external field.

The key ingredient of the self-consistent loop is the solution of the Hartree potential needed in the density functional Hamiltonian. The Hartree potential is calculated by solving the three-dimensional Poisson's equation (with appropriate boundary conditions), for the corresponding non-equilibrium charge density computed via the NEGF formalism. The Green's function method also allows extensions to include electron–phonon and many-body corrections of the electron–electron interaction.

A full description of our methodology is given in the first sections. Applications to the computation of conductance of various molecular systems are also shown, particularly applications to IETS of octane-thiols and heat dissipation in dithio-benzene. We also present how to include electron–electron interactions beyond DFT, using the well-known *GW* approach. In the last section a detailed study of carbon nanotube field-effect devices is presented.

2. DFTB as a semi ab-initio approach

The density functional–based tight-binding formalism (DFTB) has been described in detail in many articles and reviews [8]. All matrix elements and orbital wavefunctions are derived from density-functional calculations. The advantage of the method relies on the use of a small basis set and the restriction to two center integrals, allowing extensive use of look-up tables. What distinguishes our approach from empirical methods is the explicit calculation of the basis wavefunctions, which allows deeper physical insights and better control of the approximations used. The method solves the Kohn–Sham equations self-consistently using a Mulliken charge projection [10].

In the traditional DFTB code a minimal basis set of atomic orbitals is used in order to reduce the matrix dimensions for diagonalization speed-up. This approach has proved to give transferable and accurate interaction potentials and the numerical efficiency of the method allows molecular dynamic simulations of large super-cells, containing several hundreds of atoms, particularly suitable to study the electronic properties and dynamics of large mesoscopic systems and organic molecules such as CNTs, DNA strands or adsorbates on surfaces, semiconducting heterostructure, etc., see [11].

We briefly describe here the self-consistent DFTB method. The method is a development of the idea first introduced by Foulkes, where the electronic density is expanded as a sum of a reference density, $n^0(r)$, (that can be chosen as the superposition of neutral atomic densities) and a deviation, $\delta n(r)$, such that $n(r) = n^0(r) + \delta n(r)$. The total energy of the system can be described, up to second order in the local density fluctuations, as:

$$E_{\text{tot}}[n] = \sum_k n_k < \Psi_k|\mathbf{H}^0|\Psi_k > + E_{\text{rep}}[\mathrm{n}^0] + E^{(2)}[\delta n], \tag{1}$$

The first term in Equation (1) can be written in terms of the TB Hamiltonian, which is given by

$$\begin{cases} \mathbf{H}^0_{\mu\nu} = \varepsilon^{\text{free-atom}}_{\mu}, & \mu = \nu \\ < \phi_\mu | T + v_{\text{eff}}[n^0_i + n^0_j] | \phi_\nu >, & \mu \in i, \nu \in j \end{cases} \tag{2}$$

where ϕ_μ and ϕ_ν are the atomic orbitals localized around the atomic centers i and j; T is the kinetic energy operator, and v_{eff} is the effective one-particle potential, which depends on the density of the two atomic centers i and j.

The term $E_{\text{rep}}[n^0]$ in Equation (1) is the repulsive energy between the ions, screened by the electronic distribution and the exchange energy. This term is short-ranged because of the neutrality of the density $n^0(r)$ and can be expressed as a summation over atomic pair contributions as follows [12]:

$$E_{\text{rep}}[\mathrm{n}^0] = \frac{1}{2} \sum_{\alpha\beta} U_{\alpha\beta}(n^0_\alpha,\ n^0_\beta). \tag{3}$$

The third term in Equation (1) is the second order correction, which can be written as

$$E^2[\delta n] = \frac{1}{2} \int \int \left[\frac{1}{|r - r'|} + \frac{\delta^2 E_{\text{xc}}}{\delta n(r) \delta n(r')} \right] \delta n(r) \delta n(r') \mathrm{d}r \mathrm{d}r', \tag{4}$$

where the Hartree and exchange-correlation potentials have been separated.

This quantity is greatly simplified by retaining only the monopole term in the radial expansion of the atom-centered density fluctuations, written as [12]

$$\delta n_i(r) \approx \Delta q_i \frac{\tau_i^3}{8\pi} \exp(-\tau_i |r - R_i|). \tag{5}$$

It follows that the second order correction can be written as

$$E^2 = \frac{1}{2} \sum_{i,j} \Delta q_i \Delta q_j \gamma_{ij}, \tag{6}$$

where

$$\gamma_{ij} = \int \int \Gamma(r, r', n^0) n_i(r) n_j(r') \mathrm{d}r \mathrm{d}r' \tag{7}$$

is introduced as a shorthand and Δq_i are the atomic charges. The atomic charges can be easily calculated, using the Mulliken charge analysis. This consists on a simple projection of the eigenstates over the local orbitals,

$$\Delta q_i = \sum_k n_k \sum_{\mu \in i} \sum_\nu \mathrm{Re}[c^*_{k\mu} c_{k\nu} S_{\mu\nu}] - q^0_i. \tag{8}$$

Within the local density approximation (LDA) the exchange contribution vanishes for large atomic distances, hence in Equation (4) the second order correction to E_{xc} can be neglected with respect to the Coulomb interaction. The term $\Gamma(r, r', n^0)$ becomes the usual Green's function of the Coulomb potential, $1/|r - r'|$, with vanishing potential at infinity. The on-site terms γ_{ii} are related to the on-site Hubbard parameters, $U_i \equiv \gamma_{ii}$, which are calculated for any atom type within LDA-DFT as the second derivative of the total energy of the atom with respect to the occupation number of the highest occupied atomic orbital. These values are therefore neither adjustable nor an empirical parameter [10]. Therefore, the expression for γ_{ij} only depends on the distance between the atoms i and j. Applying the variational principle to the energy functional of Equation (1) together with Equation (6) and using Equation (8), it is possible to obtain a modified Hamiltonian for the Kohn–Sham equations,

$$\mathbf{H}_{\mu\nu} = \mathbf{H}^0_{\mu\nu} + \frac{1}{2}\mathbf{S}_{\mu\nu}\sum_k (\gamma_{ik} + \gamma_{jk})\Delta q_k, \quad \forall \mu \in i, \nu \in j. \tag{9}$$

Since the atomic charges depend on the one-particle wavefunctions Ψ_k, a self-consistent procedure is required. The improvement of the self-consistent over the non-selfconsistent procedure is considerable in determining structural and energetic properties of molecular systems [10].

3. The non-equilibrium Green's function technique

Despite its mathematical complexity, the non-equilibrium Green's function (NEGF) method for calculations of quantum transport has gained a great popularity in recent years, mostly because of the versatility and numerical stability of the method, in contrast to wavefunction or transfer matrix approaches. The open boundary conditions can be elegantly included by exactly mapping the contacting leads into a finite and small part of the system [13]; furthermore, the Green's function approach can be generalized to many-body quantum theory, allowing the inclusion of electron–phonon [15] as well as electron–electron interactions [16] within a unified and systematic formalism. Good references in many-body quantum theory can be found in [17, 18] and an exhaustive review on NEGF can be found in [19].

The type of systems under study can be represented as in the graph of Figure 1. The system can be arbitrarily partitioned into three parts: two contacts, C_1 and C_2 and a device region, D. In principle more contacts can be included in a general formalism, but here we restrict to two contacts for sake of simplicity.

The contacts are semi-infinite leads and it is assumed that their properties coincide with those of bulk systems [21]. The device is a collection of atoms linking the two

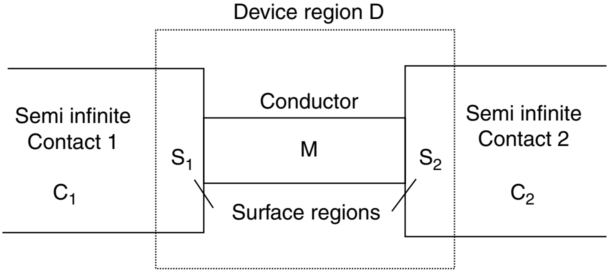


Figure 1 Diagram showing the system comprising the contact regions, C_1 and C_2, the molecular region M, and the surface regions S_1 and S_2 included in the extended-molecule region, D

contacts, comprising the molecular bridge, M, and a portion of the surfaces of the conducting leads, S_1 and S_2. The inclusion of the surfaces within the molecular region is necessary to ensure that the regions C_1 and C_2 can really be considered bulk-like. This assumption can be directly verified by checking that the charge density smoothly joins at the boundaries C_1/S_1 and C_2/S_2. The molecule plus the surfaces can be called extended-molecule.

The two contacts are kept at different electrochemical potentials, driving a current across the molecular bridge. The system can also include external modulating fields, provided by gate plates.

In solving the transport problem it is not possible to make the assumption of local thermodynamical equilibrium, i.e., a global Fermi energy is not defined. The only assumption that can be made is that the connecting leads are kept at different electrochemical potentials, and are considered as large reservoirs where the electrons are effectively in equilibrium. In order to be consistent with this scenario, the conducting bridge must offer the largest source of resistance to the flowing current. Only under such condition is it consistent to assume that the potential drops essentially across the device region, whilst the contacting leads are in equilibrium at two different constant potentials.

The underling assumption of this formalism is that the dominating transport mechanism is by coherent tunneling. Even when inelastic scattering occurs, these events are too few to allow perfect equilibration, and most of the energy relaxation occurs at the final contact. The extended-molecule region is therefore out of equilibrium and a proper treatment must be employed.

In order to compute quantum current, open boundary conditions must be employed. Open boundary conditions are necessary in order to allow the eigenstates to carry any current, which can be derived, for instance, in the context of scattering theory. The total density of states out of equilibrium can be expressed in terms of partial density of states propagating from lead ‘1’ to ‘2’ and vice versa. The key assumption is that the states $|\psi_1>$ originate deep in contact ‘1’ where there is an equilibrium condition with chemical potential μ_1. Similarly the states $|\psi_2>$ originate deep in contact ‘2’ where there is an equilibrium condition with chemical potential μ_2. Hence, the total density of these propagating states is

$$\rho = \int_{-\infty}^{+\infty} \mathrm{d}E[d_1(E)f(E-\mu_1) + d_2(E)f(E-\mu_2)]. \tag{10}$$

This expression can be used to calculate the electronic density needed for the density-functional Hamiltonian. The expression (10) for the density matrix, derived from the scattering states, is equivalent to a direct derivation from the NEGF formalism, first introduced by Keldysh [23] within many-body quantum theory [18] and later adapted to quantum transport problems by several authors [21, 26]. We give a brief overview of the main concepts of the NEGF theory which will serve us to introduce the relevant quantities of the formalism.

We start by introducing the time-ordered (also called ‘causal’) zero-temperature single-particle Green’s function, defined as

$$G(x,t,x',t') = \frac{-\mathrm{i}}{h} \frac{\langle \psi_0 | T\left[\psi_{\mathrm{H}}(x,t)\psi_{\mathrm{H}}^{\dagger}(x',t')\right] | \psi_0 \rangle}{\langle \psi_0 | \psi_0 \rangle} \tag{11}$$

where ψ_0 is the exact many-body ground-state and $\psi_{\mathrm{H}}(t)$ is the Heisenberg representation of the system wavefunction (for details see [18]). The central quantity in constructing the perturbation theory of the Green's function is the S-matrix, which evolves the wavefunction in time. The general assumption is that at time $t = -\infty$ the ground-state, ϕ_0, is for non-interacting particles. The interactions are switched on adiabatically and the wavefunction evolves into the interacting ground-state, $\psi_0 = S(0, -\infty)\phi_0$. The proof of this relationship can be found in [17]. Similarly $\psi_0 = \phi_0 S(\infty, 0)$. However in non-equilibrium conditions there is no guarantee that the system returns to its initial state for asymptotically large times. This is a fundamental assumption for the entire perturbation theory, including the fundamental Wick theorem.

A way around this difficulty is to introduce a contour-ordered Green's function, where the time ordering operator is substituted by a contour ordering operator, T_{C}. The time contour is shown in Figure 2, which, by construction, starts at an infinitely remote time and goes back to that time. The time contour can be divided into two branches (upper and lower) and the turn-point can be placed at any arbitrary time. When the two times t and t' of Equation (11) fall in the same time-branch the quantities of the standard equilibrium theory are recovered. When the two times fall in opposite branches, the time of the lower branch will always be later with respect to the time of the upper branch along the contour, therefore, the correlation function, $G^{<}(r, t, r', t') = i\langle\psi_0|\psi_{\mathrm{H}}^{\dagger}(r, t)\psi_{\mathrm{H}}(r', t')|\psi_0\rangle$, is recovered. (For now on we set $\hbar = 1$.) Any physical quantity of interest can be obtained from $G^{<}(x, t, x', t')$. For example, the charge density, $n(r, t)$, and the current density, $j(r, t)$, given by $n(r, t) = -\mathrm{i}G^{<}(r, t, r', t')$ and

$$j(r, t) = \frac{1}{2}\lim_{r'\to r}(\nabla - \nabla')G^{<}(r, t, r', t'). \tag{12}$$

In steady state the Green's functions only depend on the time difference $t - t'$, which can be Fourier-transformed to energy, E. The quantity $G^{<}(E)$ is proportional to a spectral density of occupied electronic states and the important relationship

$$n(r) = \frac{1}{2\pi\mathrm{i}}\int_{-\infty}^{+\infty} \mathrm{d}E G^{<}(r, E) \tag{13}$$

holds. For the contour ordered Green's function, the formal perturbation theory looks exactly the same as for the equilibrium counterpart. However, the perturbation terms resulting from the series expansion do not have an immediate physical meaning and must be related to the physical quantities G^{r}, G^{a} and $G^{<}$. Physically $G^{<}(E)$ and $G^{>}(E)$ represent respectively the density of occupied and empty states, whereas G^{r}, defined by

$$G^{\mathrm{r}}(E) = \frac{1}{E - H - \Sigma^{\mathrm{r}}(E)}, \tag{14}$$

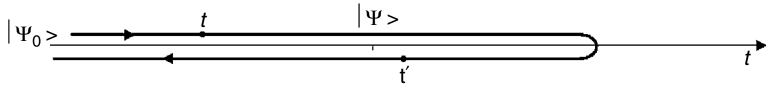


Figure 2 Complex time-contour used to define the non-equilibrium Green's function

is related to the spectral function or, broadly speaking, to the density of states. In Equation (14) we have explicitly included the presence of the retarded self-energy, $\Sigma^{\mathrm{r}}(E)$, which accounts for the presence of the contacts and can account for other perturbing potentials due to phonons, impurities or the other electrons (as shown in Sections 7 and 9). One of the advantages of the Green's function approach is that the contacting leads can be exactly mapped into the extended-molecule by using appropriate self-energies. In local orbital representations, where the interaction among atoms has a finite range, the contact self-energy can be easily calculated by exploiting the fact that the Hamiltonian describing the interaction between the device region and the contacts involves a finite number of atoms close to the junctions. Therefore, the required contact Green's function can be solved just for the matrix block corresponding to atoms close to the extended-molecule region. This so-called surface Green's function can be calculated using powerful recursive algorithms (see [7] and references therein).

Under equilibrium conditions the relationship between $G^{<}(E)$ and $G^{r}(E)$ is simply given by the Fermi distribution, as

$$G^{<}(E) = \mathrm{i}f(E)(G^{\mathrm{r}}(E) - G^{\mathrm{a}}(E)), \tag{15}$$

where $G^{\mathrm{a}}(E)$ is the so-called advanced Green's function, also given by $G^{\mathrm{r}}(E) = G^{\mathrm{a}}(E)$. Under non-equilibrium conditions it is necessary to generalize Equation (15) by taking into account that the occupation depends on the scattering itself, which is controlled by the quantity $\Sigma^{<}(E)$, known as the electron in-scattering function. The same function also controls the in-scattering of electrons from the leads into the device region. The dynamics of the occupation due to inelastic scattering processes from one energy channel to the other is provided by the Keldysh–Kadanoff–Baym (KKB) equation [23, 24], also known as the 'kinetic equation'. Within the TB matrix representation, the KKB equation can be expressed in terms of matrices, as

$$G^{<,>}(E) = G^{\mathrm{r}}(E)\Sigma^{<,>}(E)G^{\mathrm{a}}(E), \tag{16}$$

defined in terms of the retarded, the advanced Green's functions and the non-equilibrium self-energies $\Sigma^{<}$ and $\Sigma^{>}$. In physical terms $\Sigma^{<}(E)$ and $\Sigma^{>}(E)$ respectively represent the rate of injection of electrons and that of holes into the device as a function of energy. In the general case, the self-energies $\Sigma^{<}$ and $\Sigma^{>}$ can include, beside the contact self-energy, terms corresponding to scattering sources. These can be provided by electron–phonon, electron–electron or electron–defect potentials interactions. The total self-energy is a summation of the self-energies due to the leads, $\Sigma_{\alpha}^{<,>}$, and that due to additional phase-breaking interactions, $\Sigma_{\phi}^{<,>}$ [25],

$$\Sigma^{<,>} = \Sigma_{\phi}^{<,>} + \sum_{\alpha} \Sigma_{\alpha}^{<,>}. \tag{17}$$

The rate of injection can be obtained from the assumption of thermodynamical equilibrium of the leads. In fact the self-energies verify an equation similar to Equation (15),

$$\Sigma_{\alpha}^{<,>}(E) = \pm\mathrm{i}f(\pm(E - \mu_{\alpha}))(\Sigma_{\alpha}^{r}(E) - \Sigma_{\alpha}^{a}(E)). \tag{18}$$

Note in the previous equation the identity $f(-E) = 1 - f(E)$.

Putting together Equations (18), (15) and (10) it is possible to obtain general expressions for d_1 and d_2,

$$d_\alpha = \frac{1}{2\pi} G^{\mathrm{r}} \Gamma_\alpha G^{\mathrm{a}}, \tag{19}$$

where Γ_α is defined as:

$$\Gamma_\alpha = \mathrm{i}[\Sigma_\alpha^{\mathrm{r}} - \Sigma_\alpha^{\mathrm{a}}]. \tag{20}$$

In practical computations it is much more convenient to rewrite Equation (10) as

$$\begin{aligned}\rho = & \int_{-\infty}^{+\infty} \mathrm{d}E\,[d_1(E) + d_2(E)]\, f(E-\mu_1) \\ & + \int_{-\infty}^{+\infty} \mathrm{d}E d_2(E)\,[f(E-\mu_2) - f(E-\mu_1)]\end{aligned} \tag{21}$$

Using the definitions given in Equation (19), and invoking time-reversal symmetry ($G^{\mathrm{a}} = G^{\mathrm{r}^*}$), this expression can be written as:

$$\rho = -\frac{1}{\pi} \mathrm{Im}\left[\int_{-\infty}^{+\infty} \mathrm{d}E G^{\mathrm{r}}(E) f(E-\mu_1)\right] + \Delta_2. \tag{22}$$

where Δ_2 is a shorthand defining the second integration in Equation (21).

The integration in Equation (22) can be conveniently evaluated by deforming the integration from the real axis into the path $C + \ell$ shown in Figure 3. The deformation is possible according to the residue theorem:

$$\int \mathrm{d}z G^{\mathrm{r}}(z) f(z-\mu) = -2\pi \mathrm{i} kT \sum_\nu G(z_\nu), \tag{23}$$

where $z_\nu = \mu + \mathrm{i}\pi kT(2\nu + 1)$ are the poles of the Fermi function and the summation in Equation (23) includes the poles enclosed within the contour. The poles of G^{r} are slightly displaced below the real axis. The function is analytic on the upper half complex plane. For $T > 0$ the integration path stays away from the real axis where $G^{\mathrm{r}}(z)$ behaves regularly, with the exception of the point E_{low}, shown in Figure 3, chosen sufficiently below the lowest energy of the electronic spectrum. The integration involved in Δ_2 must be evaluated on the real axis, since the function is analytic on the real axis only (it depends on G^{r} and G^{a}). This integration may be delicate and requires a fine mesh.

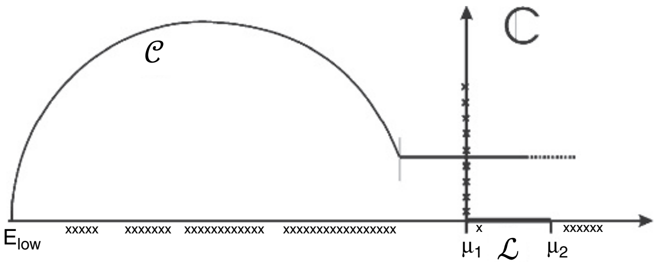


Figure 3 Diagram showing the integration path in the complex plane needed to evaluate the non-equilibrium Green's function. The crosses below the real axis represent the poles of $G^{\mathrm{r}}(z)$, and those on the imaginary axis the poles of the Fermi function

4. Real contact and virtual contact currents

The current flowing in the system can be computed by defining an appropriate current operator [26, 21]

$$I_{\mathrm{op}}(r, r') = \frac{e}{h}\big[G^{<}(r, r')H_{\mathrm{D}}(r) - H_{\mathrm{D}}(r)G^{<}(r, r')\big], \tag{24}$$

whose diagonal elements give the divergence of the total current. An alternative derivation for deriving TB currents involves the definition of a bond-current operator [20], which is very useful for the calculation of local current fluxes. The trace of the current operator gives the net outflow of current per unit energy across an imaginary surface enclosing the molecular region,

$$\mathrm{Tr}[I_{\mathrm{op}}] = \int \nabla \cdot \mathbf{J}(r, E)\mathrm{d}r. \tag{25}$$

Therefore, using Equations (20) and (15) and the relationships $G^{\mathrm{r}} - G^{\mathrm{a}} = G^{>} - G^{<}$, $\Sigma^{\mathrm{r}} - \Sigma^{\mathrm{a}} = \Sigma^{>} - \Sigma^{<}$, the total outflow of current at any specific energy can be written as

$$\mathrm{Tr}[I_{\mathrm{op}}] = \frac{e}{h}\mathrm{Tr}\big[\Sigma^{<}(E)G^{>}(E) - \Sigma^{>}(E)G^{<}(E)\big], \tag{26}$$

Inserting in Equation (26) the expression for the total self-energy (16), it is possible to identify the terminal current contributions to the total current,

$$i_{\alpha}(E) = \frac{e}{h}\mathrm{Tr}\big[\Sigma^{<}_{\alpha}(E)G^{>}(E) - \Sigma^{>}_{\alpha}(E)G^{<}(E)\big]. \tag{27}$$

Expression (27) represents the inflow of current from the α contact into the molecular region per unit energy.

Combining Equations (15) and (27) it is possible to separate the coherent and the incoherent contributions from the total current,

$$i_{\alpha}(E)_{\mathrm{coh}} = \frac{e}{h}\sum_{\beta}\mathrm{Tr}\big[\Sigma^{<}_{\alpha}G^{\mathrm{r}}\Gamma_{\beta}G^{\mathrm{a}} - \Gamma_{\alpha}G^{\mathrm{r}}\Sigma^{<}_{\beta}G^{\mathrm{a}}\big], \tag{28}$$

$$i_{\alpha}(E)_{\mathrm{incoh}} = \frac{e}{h}\sum_{\beta}\mathrm{Tr}\big[\Sigma^{<}_{\alpha}G^{\mathrm{r}}\Gamma_{\phi}G^{\mathrm{a}} - \Gamma_{\alpha}G^{\mathrm{r}}\Sigma^{<}_{\phi}G^{\mathrm{a}}\big]. \tag{29}$$

The incoherent part involves scattering contributions giving non-vanishing $\Sigma^{<}_{\phi}$ and $\Sigma^{>}_{\phi}$. The interactions can be viewed as an exchange of particle with a *virtual* contact (see Figure 4), which adsorbs electrons at a given energy and emits them at another energy. In this respect, in perfect analogy to the real contacts, the virtual contact breakes the phase coherence of the wavefunction. Similarly to Equation (27), the flux of electrons at the virtual contact can be written as:

$$i_{\phi}(E) = \frac{e}{h}\mathrm{Tr}\big[\Sigma^{<}_{\phi}(E)G^{>}(E) - \Sigma^{>}_{\phi}(E)G^{<}(E)\big]. \tag{30}$$

Current conservation must ensure that the net current exchange with the virtual contact is zero, or

$$\int i_{\phi}(E)\mathrm{d}E = 0. \tag{31}$$

This is an important constraint to be checked when developing the interaction self-energy from perturbation theory.

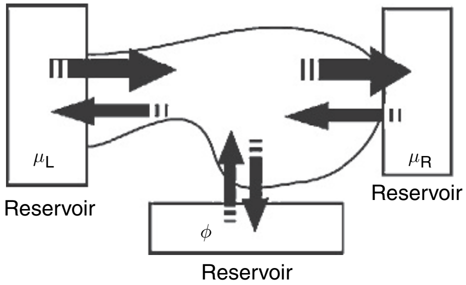


Figure 4 Diagram showing the in-scattering and out-scattering electron in and out of the electrodes and the virtual phase-breaking contact

5. The Poisson equation

As anticipated in Section 2, the Hartree potential needed for the SCC iteration of the Kohn–Sham equations is computed by solving the Poisson's equation with the appropriate boundary conditions imposed by the contacts. The Poisson's equation for the mean field electrostatic potential should be written as:

$$\begin{cases} \nabla^2[V_{\mathrm{el}} + V_{\mathrm{ions}}] = -4\pi \sum_i [n_i^0(r) + \delta n_i(r) + q_i \delta(r - R_i)] \\ \qquad\qquad + \text{boundary conditions} \end{cases} \tag{32}$$

where V_{el} and V_{ions} are respectively the contribution from the electrons (treated in mean field approximation) and the ions to the total electrostatic potentials, and q_i are the ionic charges. The usual boundary condition used to solve the Poisson equation in DFT is that the potential vanishes at infinity. This gives the familiar solution for the electrostatic potential of a point charge, q, as $q/|r - r'|$, and the usual form for the Hartree energy. In transport problem the boundary conditions are more likely imposed by the applied potentials.

In the DFTB implementation the self-consistent potential is related to the density fluctuations $\delta n(r)$. The effective potential for the reference density n^0 is included in $\mathbf{H}^0$ and $E^{\mathrm{rep}}[n^0]$ of Equation (1). Therefore, by linearity, we split the Poisson's Equation (32) into two equations, one for the ionic part plus the reference density:

$$\nabla^2 \left[V_{\mathrm{el}}^{(0)} + V_{\mathrm{ions}}\right] = -4\pi \sum_i \left[n_i^0(r) + q_i \delta(r - R_i)\right], \tag{33}$$

and the other for the self-consistent correction:

$$\nabla^2 V_{\mathrm{el}}^{(2)} = -4\pi \sum_i \delta n_i(r). \tag{34}$$

When solved with the usual boundary condition, Equations (33) and (34) give the electrostatic field included in the usual DFTB calculations. However, Equation (34) is solved using the boundary conditions imposed by the device. These conditions arise from the natural requirement that deep inside the contacts the effective potential for the Kohn–Sham equations must correspond to the bulk electrochemical potentials. Therefore, at the boundaries between the device region and the contacts, the potential must match the intrinsic effective bulk potential (which originates from any equilibrium charge density) shifted by the applied bias. At the device-contacts interfaces, C_α/S_α, the potential must satisfy

$$V^{(2)}_{S_\alpha}(r)|_{C_\alpha/S_\alpha} = V^{(2)}_{C_\alpha,\mathrm{bulk}}(r)|_{C_\alpha/S_\alpha} + \Delta V_\alpha, \tag{35}$$

where ΔV_α is the applied external potential to the α-contact. The decoupling of Equations (33) and (34), which at first may seem arbitrary, is actually a good approximation since the reference density is taken as that of the neutral atoms and therefore cancels with the ionic charges. On the contrary, the excess density produces a long-range Coulomb field that should respect the boundary conditions imposed by the device. For example, the charge which accumulates on the contact surfaces must be consistent with the applied bias.

Within the *gDFTB* approach, the Poisson equation is solved in real space using a three-dimensional multi-grid algorithm applied to a general linear, non-separable, elliptical PDE. The Poisson equation is just a particular case of this kind of equation, which has the general form

$$\sum_i \left[c_{ii}(\mathbf{r}) \frac{\partial^2 V(\mathbf{r})}{\partial x_i^2} + c_i(\mathbf{r}) \frac{\partial V(\mathbf{r})}{\partial x_i} \right] + c(\mathbf{r}) V(\mathbf{r}) = \rho(\mathbf{r}), \tag{36}$$

where $V(\mathbf{r})$ is the unknown solution potential. The previous equation simply reduces to the Poisson equation if the coefficients are taken such that $c_i(\mathbf{r}) = c(\mathbf{r}) = 0$ and $c_{ii}(\mathbf{r}) = -1/4\pi$ at all the points $\mathbf{r}$ of the three-dimensional box in which the equation itself is discretized. The general equation (36) has a great flexibility. Indeed, the possibility of setting the coefficients to different values in different regions of the solution space is the fundamental feature which allows us to impose Dirichlet boundary conditions on arbitrary shaped three-dimensional surfaces, such as planar or cylindrical gates, and handles easily even four terminal geometries.

A Dirichlet boundary condition simply consists of imposing a value for the solution potential on a defined spatial region. Consequently, such a boundary condition can be set by imposing in the spatial region occupied by the metallic gate contact, $c_{ii}(\mathbf{r}) = c_i(\mathbf{r}) = 0, \mathrm{c}(\mathbf{r}) = 1$ and $V(\mathbf{r}) = \rho(\mathbf{r})/c(\mathbf{r})$. The charge density $\rho(\mathbf{r})$ is in turn suitably initialized to the appropriate value, for instance, of an external gate field. The region occupied by the gate can obviously have any arbitrary shape, since the boundary condition is simply reduced to the initialization of a numerical function on a subset of points on which the equation has been discretized.

In the remaining regions of the solution box, where the metallic gate is not present, the coefficient are suitably initialized in order to obtain the effective Poisson Equation 34, and, at the same time, the charge density is evaluated starting from the atomic charge fluctuations projected on the real-space mesh.

Once the computation of $V^{(2)}_{\mathrm{el}}(r)$ is done, it is projected back into the local orbital basis. The mesh is usually chosen as a trade-off between accuracy and computational speed. However, the ansatz (5) for the atomic charge density gives quite smooth functions and usually convergent results are obtained with a mesh spacing of 0.5 atomic units.

6. Applications to molecular conductance

In this section we show results of self-consistent computations of the I-V characteristics, charge density and potential of molecular systems bridging metal contacts. Benzene-dithiol, shown in Figure 5, is a simple and important test system much studied as a test-bed reference for comparisons between theory and experiments.

For such system we find that the potential drops almost linearly across the molecular bridge. The I-V characteristics shown in Figure 5 exhibits an ohmic behavior at small bias which turns into a superlinear behavior beyond 2 V. The ohmic behavior owes to a relatively high conductance also due to the position of the highest occupied level of the molecule (HOMO), lying at 1.0 eV below E_F.

Aviram and Ratner proposed in 1974 to realize a diode with a donor-σ-acceptor molecule connected at either end to metallic leads [1]. The new aspect of this idea was that the combined system of electrons and leads could support a continuous sequence of electron transfer processes. For the particular setup considered, the current–voltage (I-V) characteristic was predicted to be the one of a rectifier. More recently a small rectification ratio was demonstrated in a break-junction experiment [27] for a single molecule. The molecule is a variation of a Tour wire, asymmetrically doped by fluorination of a benzene ring, as shown in Figure 6. The two rings are separated by an insulating bridge of diphenyl, where the insulation is produced by the break of conjugation due to the tilting angle between the two benzene rings of the bridge.

The self-consistent potential for an applied bias of 2.0 V is shown in Figure 6. From the isolines it is possible to see that the potential drops non-linearly across the molecule,

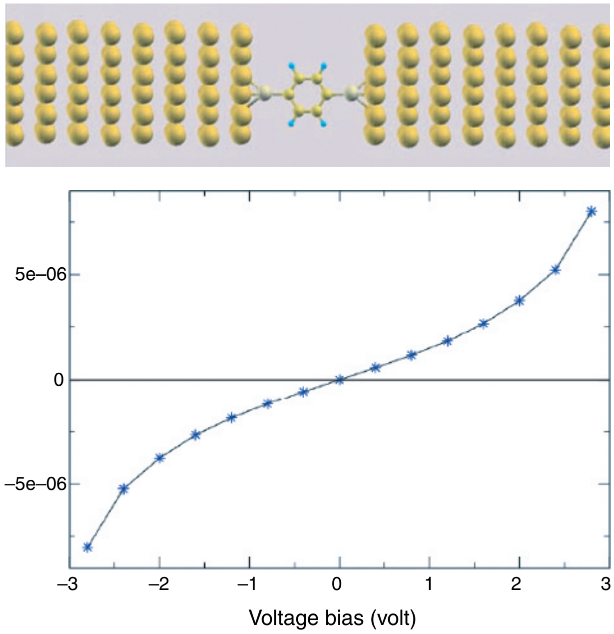


Figure 5 Diagram showing a di-thio-phenyl molecule bridging two Au contacts and the calculated I-V characteristics

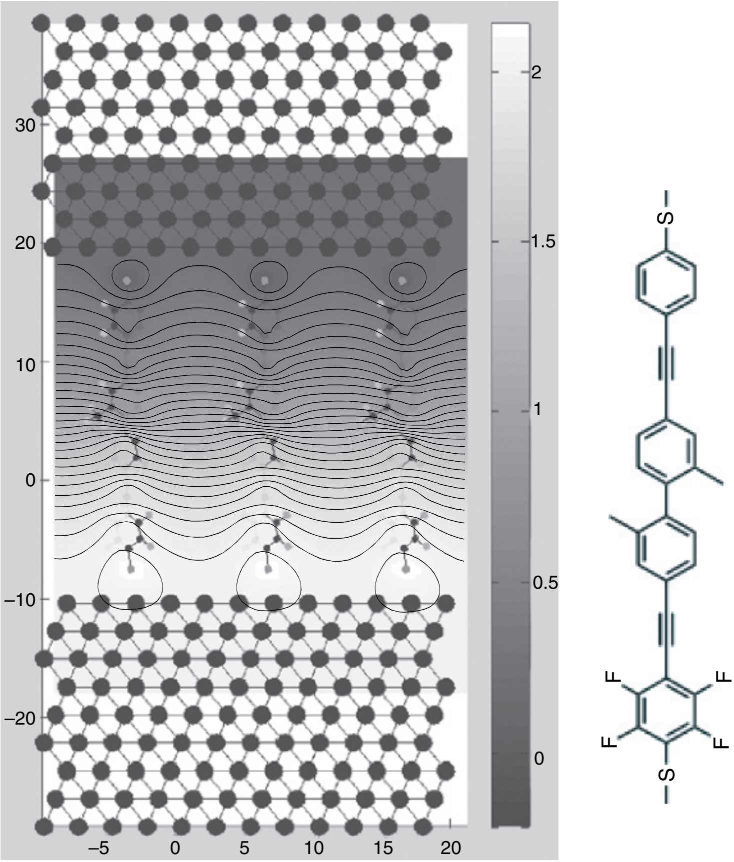


Figure 6 Self-consitent potential of a Tour wire between Au contacts

with a larger drop at the insulating bridge. The computed I-V characteristics do not show relevant rectification within the bias range between −2 and +2 V. The computed DFT eigenstates predict the HOMO level nearly in resonance with the Au Fermi level at $V = 0$. The feature at −4.8 eV is instead related to the local density of states of the contacting Au lead, as it is not affected by the applied bias. As the bias increases more resonances enter in the injection window, as shown in the left panel of Figure 7. This produces the staircase-like I-V characteristics shown in the right panel of Figure 7. Such steps are also observed experimentally, although at different voltages.

The difficulty of making accurate theoretical predictions, in agreement with experimental findings, is due to the high sensitivity on the precise position of the molecular energy levels, which are very difficult to predict using just DFT calculations. As discussed in Section 9, correction for the exchange and correlation energies is necessary for quantitative results. It is worth stressing that in transport calculations it is not just the magnitude of the energy gap that matters, but also the absolute position of the molecular energy levels with respect to the Fermi level of the metal, providing the reference injection energy.

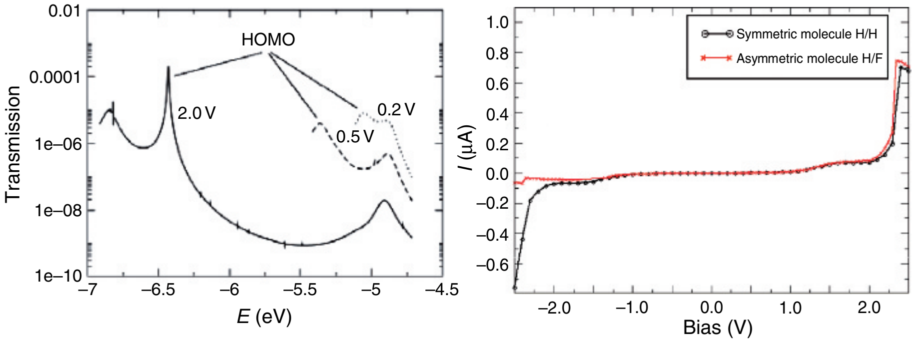


Figure 7 Transmission probability and computed current for the symmetric and asymmetric Tour wire

7. Analysis of IETS spectra

An IETS spectrum is formally defined as the second derivative of I vs. V. As such spectra are obtained at very low temperature (usually 4.2 K) only phonon emission is possible. When the applied bias matches a phonon frequency an additional channel of tunneling assisted by phonon emission opens up. This corresponds to a barely visible kink in the I-V characteristics, but it can be amplified as a peak in the second derivative. The assignment and interpretation of the spectra are not without difficulties. The spectra are usually assigned with the help of IR, Raman and HREELS results for monolayers of the molecule in question, or even isolated molecules. However, the interaction with the substrate and the absence of definite selection rules means that IETS may exhibit markedly different spectra from these other techniques and the characteristics of these spectra are difficult to predict. Theoretical simulations are necessary in order to interpret the measured IETS spectra [15, 28, 29].

To obtain the spectrum we calculate the coherent and incoherent current at $T = 0$, using Equations (28) and (29). The calculation of the incoherent component requires an explicit evaluation of the electron–vibration coupling matrices, $\boldsymbol{\gamma}^q_{\mu\nu}$, obtained by expanding the TB Hamiltonian to first order in the atomic displacements. The couplings are then expressed in terms of derivatives of the Hamiltonian and the overlap matrices, therefore without fitting parameters [30, 31], as

$$\boldsymbol{\gamma}^q_{\mu\nu} = \sqrt{\frac{\hbar}{2\omega_q M_q}} \sum_\alpha \left[\frac{\partial \mathbf{H}_{\mu\nu}}{\partial R_\alpha} - \sum_{\lambda,\sigma} \frac{\partial \mathbf{H}_{\mu\lambda}}{\partial R_\alpha} \mathbf{S}^{-1}_{\lambda\sigma}\mathbf{H}_{\sigma\nu} - \mathbf{H}_{\mu\lambda}\mathbf{S}^{-1}_{\lambda\sigma}\frac{\partial \mathbf{H}_{\sigma\nu}}{\partial R_\alpha}\right]\mathbf{e}^q_\alpha, \tag{37}$$

where R_α are atomic displacements, M_q the atomic masses and $\mathbf{e}^q_\alpha$ the vibrational mode eigenvectors. The relevant self-energy is evaluated within the first order Born approximation, as

$$\Sigma^{<,>}_{\text{el-ph}}(\omega) = \mathrm{i}\sum_q 2\pi \int \mathrm{d}\omega' \gamma^q G^{<,>}(\omega-\omega')\gamma^q D^{<,>}_{0,q}(\omega') \tag{38}$$

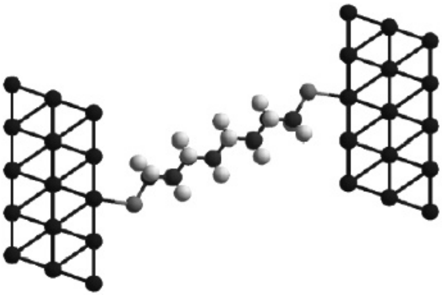


Figure 8 Relaxed Alkanethiol geometry between Au contacts

where the $D_{0,q}^{<,>}(\omega)$ are the correlation functions related to the vibrational modes, assumed Einstein oscillators in thermal equilibrium with a bath,

$$D_{0,q}^{<,>}(\omega) = -2\pi i[(N_q + 1)\delta(\omega \pm \omega_q) + N_q\delta(\omega \mp \omega_q)]. \tag{39}$$

The geometry of the molecule in the junction, shown in Figure 8, is determined in two steps. First, an optimized geometry is obtained for octanethiol chemisorbed through the terminal sulfur to a single Au(111) surface. Periodic boundary conditions are used; however, the chemisorbed molecules were sufficiently far apart to be considered isolated. The geometry for the full electrode–molecule–electrode system was then generated by symmetrizing about a point of inversion between the C4–C5 bond to give octanedithiol bound to two co-facial Au(111).

The computed conductance of such system agrees with a model which assumes that approximately 10,000 molecules are sampled in parallel within the nanopore device used in the experiment [5].

The *gDFTB* code reproduces, to a reasonable extent, experimentally observed vibrational frequencies for octanedithiol chemisorbed on Au [30]. The electrodes produce a significant perturbation in the character of the molecule and as a consequence the vibrational modes associated with the extremities of the molecule differ in frequency from modes of the same character associated with the central region. For example, the C–H symmetric stretching modes occur at different frequencies for the modes associated with the central region (0.368 eV) and the extremities (0.348 eV).

This is reflected in the calculated IETS spectra for octanedithiol, shown in Figure 9 with the peaks assigned as shown in Table 1. Unlike IETS, the spectroscopies used to assign IETS spectra of octanedithiol (IR, Raman, HREELS) all have specific selection rules. It has been observed previously that in IETS spectra both IR-active modes and Raman-active modes can be seen as well as additional modes, although not all IR-active modes and Raman-active modes may be seen. This complex relationship between what is observed in IETS and what is observed in other spectroscopic techniques means that for a system of the complexity of octanedithiol a complete assignment of the IETS spectra from IR, Raman and HREELS is difficult to achieve. For instance, in the experimental IETS spectra for octanedithiol [5] there were a number of unknown peaks attributed to Si_3N_4 impurities.

Indeed, according to our calculations, relevant signal from molecular modes in the high frequency range (above 2000 cm^{-1}) should not be expected. On the other hand,

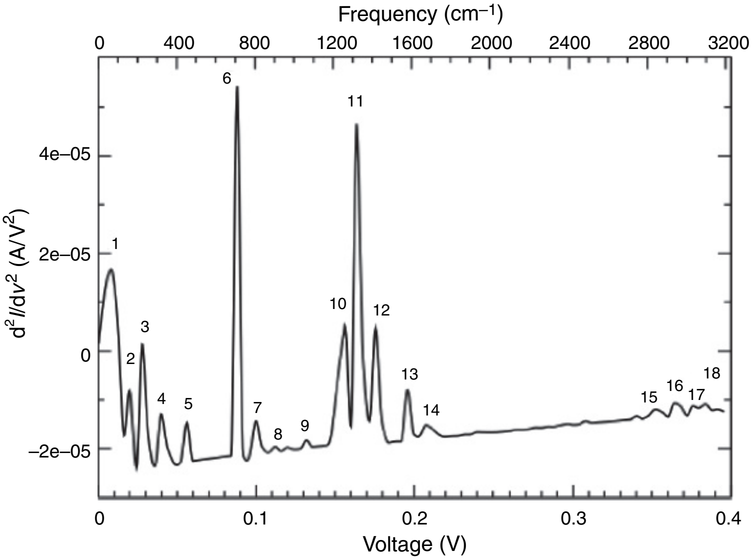


Figure 9 Simulation of the IETS for the octanedithiol (hcp bonding site). Numbers for peaks are related to modes of vibration in Table 1

Table 1 Principal peaks in the calculated IETS spectrum

Peak	Voltage (V)	mode
1	0.008	S–C–C out-of-plane wag
2	0.017	S–C–C out-of-plane wag
3	0.033	Au–S stretch
4	0.044	S–C–C scissor
5	0.060	C–C–C scissor
6	0.083	C–S stretch
7	0.111	CH_2 in-plane rock (extremities)
8	0.121	CH_2 in-plane rock (central)
9	0.135	CH_2 in-plane rock (central)
10	0.150	CH_2 in-plane rock (all)
11	0.157	C–C stretch
12	0.164	C–C stretch
13	0.180	CH_2 scissor (extremities)
14	0.196	CH_2 scissor (central)
15	0.212	CH_2 out-of-plane wag (all)
16	0.217	CH_2 out-of-plane wag (all)
17	0.348	C–H stretch sym (extremities)

peaks 1, 2 and 5 nicely correspond to observed, but not clearly assigned, features [5, 31]. These are low frequency modes (Table 1), the first of which can be described as a rigid out-of-plane oscillation of the four central CH_2-groups, the second as a rigid and symmetric out-of-plane oscillation of the two C–C–C–S backbones. The fifth mode is associated to the longitudinal oscillation of the CH_2–CH_2 subunit pairs. Overall the calculated spectra reflects qualitatively the experimental findings, giving the largest signal from the backbone vibrational modes in the range 1000–1500 cm^{-1}.

8. Power dissipation in molecular junctions

A relevant quantity for technological applications is the amount of power dissipated in the molecule due to inelastic phonon emission, which can be obtained by considering the virtual contact current, as discussed, for instance, in [30]. This is an important quantity since it is related to thermal dissipation and therefore to the stability of the molecule under applied bias. The power dissipated is given by the net rate of energy transferred to the molecule and can be calculated using the virtual contact current as

$$W = \frac{2}{h}\int_{-\infty}^{+\infty} \omega Tr\left[\Sigma_{ph}^{<}(\omega)G^{>}(\omega) - \Sigma_{ph}^{>}(\omega)G^{<}(\omega)\right] d\omega, \tag{40}$$

simply representing the average energy transfer occurring at the virtual contact.

The power dissipated can be used to compute for the rate of phonon emission. This can be done by first observing that Equation (40) can be expressed as a sum over the individual vibrational modes, $W = \sum W_q$, allowing to compute the power dissipated in each mode [30]. In order to take into account for the non-equilibrium phonon population of the molecular modes we set up a phenomenological rate equation. The rate of phonon emission can be defined as $R_q(N_q) = W_q/\hbar\omega_q$ (the energy emitted divided by the phonon energy), which is a function of bias and phonon population. The rate R_q is actually the net rate of phonon emission, also including the absorption rate due to assisted tunneling and electron–hole pair production, very important in metal contacts. In order to compute the phonon population of the vibrational modes a rate equation can be written, including the rate of emission and dissipation into the leads, as

$$\frac{\mathrm{d}N_q}{\mathrm{d}t} = R_q(N_q) - J_q(N_q - N_q^0(T)) = 0, \tag{41}$$

where J_q is the rate of phonon dissipation and $N_q^0(T)$ is the equilibrium thermal distribution of phonons, which is established without applied bias. Under stationary condition, Equation (41) can be used to compute the non-equilibrium phonon population. For the alkanethiol discussed in the previous section most of the power is emitted in the C–C stretching modes, for which we have computed an average of 10 pW per mode. C–S and S–Au modes adsorb approximately 8 pW. The total power emitted in the molecule under this simple stationary model is 0.16 nW at the applied bias of 2.0 V, rising with an approximately linear behavior. The amount of power dissipated in such molecule is actually very small and does not affect considerably the phonon population.

To show the effect, we can consider another molecule, di-thio-phenyl, in which the power dissipation is much larger owing to a larger incoherent current. The analysis is restricted to those vibrational modes which give non-negligible incoherent electron–phonon scattering [32], having frequencies of $\omega_q = 756$, 1147, 1182, and 1754 cm^{-1}, respectively. The molecule and its vibrational modes are represented in Figure 10. The coupling of the vibrational modes with the reservoirs gives a phonon decay rate $J_q \approx 10^{13}$ Hz.

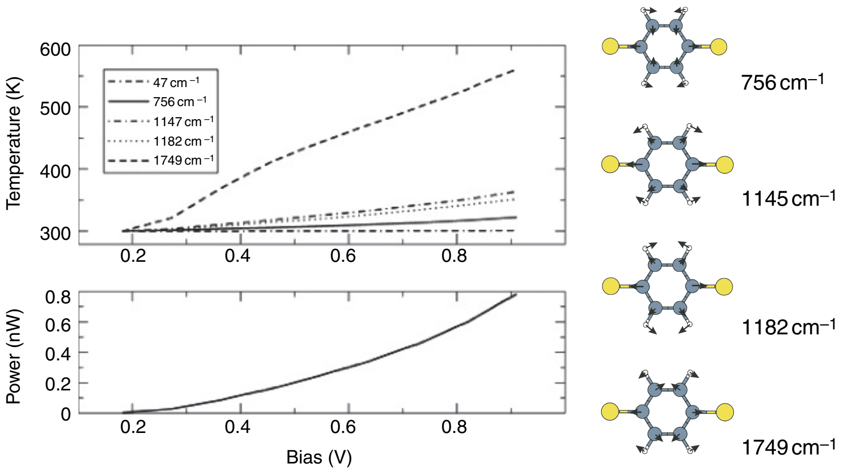


Figure 10 Effective mode temperature and total power dissipated into the molecule as a function of bias for a reservoir temperature of 300 K. Vibrational mode representation of the most important modes for incoherent electron–phonon scattering

Figure 10 shows the power dissipated and the equivalent temperature of each mode as a function of applied bias for a contact temperature of 300 K. The equivalent temperature is obtained from the Bose–Einstein distribution using the self-consistent solution of the steady-state phonon population. It is possible to see that the highest energy modes heat up considerably, reaching a temperature of almost 600 K, while low energy modes are less sensitive. This is related to the larger power emitted in such modes and the fact that the power emitted depends on the population itself. Since the lowest modes have an equilibrium population, $N_q > 1$, the net emission rate is small since emission and absorption probabilities tend to cancel each other.

9. GW corrections and transport

As discussed above, the DFT approach is used to construct the one-particle system Hamiltonian. The advantage is that DFT is a valid method to include several hundreds of atoms which are frequently necessary to include an atomistic description of the contacts. The main problem of DFT is related to the description of the unknown exchange-correlation potential, usually approximated to be locally given by that of a free electron gas of equal density. This tends to overestimate the metallic characteristics of the molecular states, producing among the others, an underestimation of the HOMO–LUMO gap, with relevant consequences to transport. Moreover, DFT is a ground-state theory providing at most an exact electronic density, but it is not meant to compute exact wavefunctions and single particle energy levels, both necessary ingredients for tunneling calculations. In order to obtain quantitative prediction of tunneling currents as well as correct quantitative trends it is necessary to go beyond DFT, even keeping the simple single-particle description. Many attempts have been made to improve the DFT calculations using hybrid functionals, correlated transport [33, 34] and self-interaction corrections [35]. The alternative time-dependent current density-functional scheme [36] could provide a consistent scheme to compute steady-state currents, automatically including excited-state corrections.

Within the NEGF formalism electron–electron interactions contribute both to $\Sigma^<$ and to Σ^r. In a way similar to electron–phonon interactions, the contribution to $\Sigma^<$ takes into account phase-breaking scattering events, whereas Σ^{r} corrects for the single particle propagator,

$$G^{\mathrm{r,a}}(E) = [ES - \mathbf{H} - \Sigma_{\mathrm{R}}^{\mathrm{r,a}}(E) - \Sigma_{\mathrm{L}}^{\mathrm{r,a}}(E) - \Sigma_{\mathrm{ee}}^{\mathrm{r,a}}(E)]^{-1}. \tag{42}$$

Since the Hamiltonian already contains a mean field DFT approximation for the electron–electron interactions, it is important to note that Σ_{ee} must contain terms which subtract such contributions. In fact the general theory for developing the electron–electron interaction starts usually from the free propagator.

The GW approximation is essentially a first order truncation of a general and systematic perturbative expansion of the electron–electron interactions in terms of a screened Coulomb potential. The GW method takes its name from the characteristic form of the self-energy [37],

$$\Sigma_{\mathrm{ee}}(E) = \Sigma_{\mathrm{GW}}(E) = \frac{\mathrm{i}}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}E' e^{\mathrm{i}E'0^+} G_0(E-E')W(E') \tag{43}$$

Where G_0 is the single particle Green's function and $W(E)$ is the screened Coulomb potential, which can be expressed in terms of the bare Coulomb interaction, v, and the generalized dielectric function, $\varepsilon(E)$, as $W(E) = \varepsilon^{-1}(E)v$, or in terms of the polarization function $W = v/(1 - vP(E))$. Σ_{GW} can be split into the sum of two terms so that $\Sigma_{\mathrm{GW}} = \mathrm{i}G_0 v + \mathrm{i}G_0(\varepsilon^{-1} - 1)v = \Sigma_{\mathrm{x}} + \Sigma_{\mathrm{c}}$. The first term reduces to the known formula for the exchange energy contribution of Hartree–Fock, whilst the second term is the correlation part. Such an approximation has been proved to give essentially exact results for a free electron gas [37] and provides very good corrections of bulk semiconductor bandgaps [38, 39]. The method has been also applied with success to molecular systems [40]. Unfortunately the full GW method is very expensive and many approximations are necessary in order to increase the computational speed and make calculations feasible.

The GW correction has been efficiently implemented on the DFTB method [41]. The key approximation of such implementation is to write the wavefunction products into charge monopoles, as

$$\phi_\mu(r)\phi_\nu(r) \approx \frac{1}{2}S_{\mu\nu}(|\phi_\mu(r)|^2 + |\phi_\mu(r))|^2), \tag{44}$$

within the spirit of the Mulliken charge approximation used for the self-consitent charge calculations (see Section 2). The exchange self-energy, Σ_i^{x}, can be computed using

$$\Sigma_i^{\mathrm{x}} = \sum_j^{\mathrm{occ}} \sum_{\mu\nu} q_\mu^{ij}[v]_{\mu\nu} q_\nu^{ij}, \tag{45}$$

where the q_μ^{ij} are generalized Mulliken charges [41], and $v_{\mu\nu}$ are the Coulomb integrals between squared DFTB atomic orbitals. The correlation term is evaluated with a similar approach. First the polarization function of the self-energy is approximated in the RPA framework,

$$P(E) = \sum_{j}^{\mathrm{occ}} \sum_{i}^{\mathrm{virt}} \frac{\phi_i \phi_j \phi_j \phi_i}{E - (\varepsilon_i - \varepsilon_j)}, \tag{46}$$

and then approximated within the plasmon-pole approximation, to reduce the integration in Equation (43) to an analytic form. The approximation consists of assuming a simple form for the inverse dielectric function, $\varepsilon^{-1}(E)$, in terms of its eigenvalues, which are obtained by diagonalization at different test frequencies. With such approximations the difficult integration of equation (43) can be done analytically. Such an approach provides a very fast GW calculation [41], behaving exceptionally well on several test molecules, particularly on π-conjugated systems. The first step requires the explicit calculations of the quasiparticle states, given by

$$\varepsilon_i^{\mathrm{QP}} = \varepsilon_i^{\mathrm{DFTB}} + Z_i \langle \phi_i | \Sigma_{\mathrm{x}} + \Sigma_{\mathrm{c}} - \nu_{\mathrm{xc}} | \phi_i \rangle, \tag{47}$$

where the three terms in the brackets are respectively the exchange, the correlation and the DFTB exchange-correlation contribution to the energy. The Z_i, instead, takes locally into account of the energy dependence of the self-energy [41]. In the common GW approximation the DFT eigenstates are not perturbed. Consistent with such approximation the quasiparticle energies can be seen as improved MO eigenvalues, and the GW self-energy can be projected back into the local orbital representation using the original DFT eigenstates.

In the Plasmon-pole approximation the quasiparticle lifetimes are approximated to be infinite. Since the lifetime is longer than the tunneling time, we can consider this a good approximation and neglect the contribution to the in-scattering function, $\Sigma^<$, in the transport properties. The coherent transmission is modified only by the effect of Σ^{r} on the propagator.

The method has been applied to the computation of the transmission across sulfur-ended molecular chains of thiophenes and phenyl-ethynes of different lengths, from one to five rings, contacted by Cu surfaces. The Cu contacts are considered in the wide-band limit. This approximation, although apparently crude, captures the essential physics of the metal–molecule interaction, without introducing in the present discussion additional complications, for instance, due to surface states and features of the local DOS of the contacts. Such molecules offer an interesting test, because the π-conjugated system they form produces highly delocalized states across the whole molecule, with a consequent highly conducting behavior. In such system the failure of DFT-LDA becomes even more evident, as the HOMO level gets quite close to the Fermi level of the metal contacts, giving rise to a theoretically very high coherent conduction, as opposed to many experiments.

The GW renormalization was tested first on isolated molecules with consistent results. The enlargements of the gap region is very satisfactory, both the HOMO and LUMO states are shifted in energy. The code was already tested for other different classes of organic molecules for which it was showed that the enlargement of the energy gap follows the difference between electron affinity and first ionization energy [41]. In particular for diphenylethyne, we obtain $E_{\mathrm{HOMO}} = -5.17\,\mathrm{eV}$ and $E_{\mathrm{LUMO}} = -2.29\,\mathrm{eV}$ in the DFT calculation, whereas the GW correction leads to quasiparticle energies of $E^{\mathrm{QP}}_{\mathrm{HOMO}} = -6.64\,\mathrm{eV}$ and $E^{\mathrm{QP}}_{\mathrm{LUMO}} = 0.26\,\mathrm{eV}$, comparing much better with the experimental workfunction of $\Phi = -7.94\,\mathrm{eV}$ and affinity of $A < -0.32\,\mathrm{eV}$.

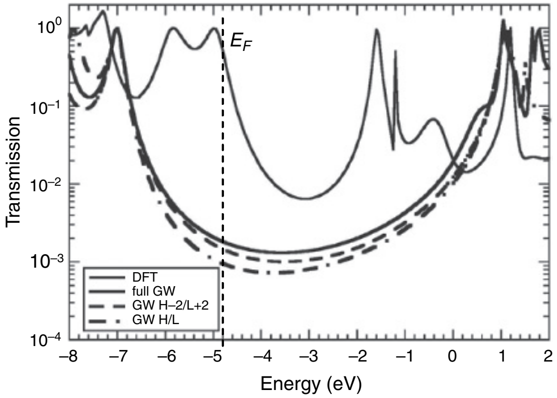


Figure 11 Comparison among different renormalization schemes for a benzene-dithiol. The transmission is fully renormalized, only the levels close to the HOMO and LUMO (from HOMO−2 to LUMO+2)

Figure 11 shows the transmission function across a benzene-dithiol, the shortest of the molecular chains. The black line represents the calculation of the DFTB transmission, the other curves refer to selective renormalization of the molecular energy levels, using GW calculations. Renormalization of all the energy levels is compared to renormalization of the HOMO and LUMO levels only, as well as a selective renormalization of just two states closest to the HOMO–LUMO. The small fluctuation of the transmission in the gap region induced by change in the renormalization scheme supports the idea that the transmission is controlled only by the HOMO–LUMO states and few states closest in energy.

Figure 12 reports calculations of the transmission across the oligomers of increasing length. We observe that associated to the HOMO–LUMO gap renormalization, there

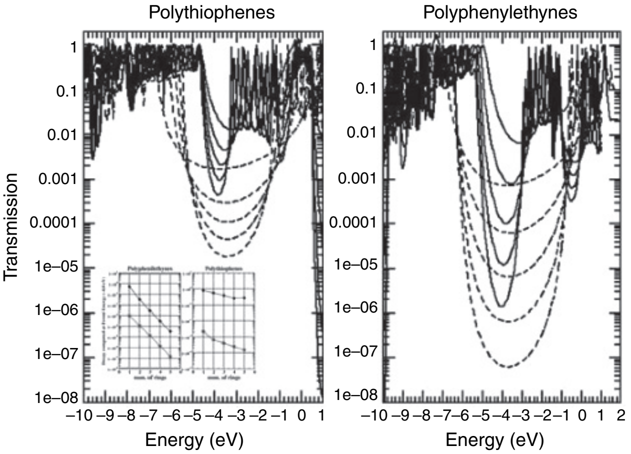


Figure 12 Comparison between the DFTB transmission and the renormalized one with GW for polyphenylethynes and polythiophenes chains up to five rings. In the small box the beta decay of the transmission at Fermi energy (−4.8 eV) is shown

is also a strong reduction on the magnitude of the transmission, especially close to the Fermi energy (around −4.8 eV). The effect is relevant for both chains where the error of DFT in the energy gap produces a near resonance of the HOMO level at the metal Fermi energy. The GW correction preserves the trend of the transmission for increasing molecular lengths, since as the chains become longer the gap shrinks but the transmission reduces. The calculated trends for both molecules is shown in the small box in Figure 12. It is interesting to note that the GW correction, in the poly-phenylethynes, though producing a strong decrease in conduction, does not affect considerably the exponential decay length, with only slight decrease of its value. The effect is instead dramatic for poly-thiophenes where the decay changes qualitatively also.

10. Applications to CNT devices

Carbon nanotubes have been recently acknowledged as promising candidates, among several low-dimensional physical systems, to realize nanoscale electronic devices [42, 43]. In particular, single-wall CNTs are ideal candidates to study general transport properties of quasi-one-dimensional conductors.

Here we focus our attention on the theoretical description which can be given as bulk-modulated, CNT-based field-effect transistors (CNTFETs) by means of the *gDFTB* approach. The physical mechanisms governing transport in these devices, starting from the role played by one-dimensional screening on gate- and drain-induced current modulation, can be correctly predicted at an atomistic level within our approach.

The system we have simulated is shown in Figure 13. It consists of an infinite, semiconducting CNT (7, 0), having a diameter of 5.52 Å and an energy gap $E_g = 1.3$ eV. The nanotube is coaxially gated by a metallic cylindrical contact centered in the middle point of the channel. The gate length is 1.2 nm. An insulator layer with dielectric constant $\varepsilon_r = 3.9$ and a thickness of 0.8 nm surrounds the metallic cylinder.

Details on the interface between the source–drain metallic contacts and the semiconducting nanotube are not necessary to describe bulk-switching mechanisms, which consist of a local gate-field modulation with no electrostatic coupling with the metallic contact structures.

To simulate charge injection in the intrinsic channel, we simply dope the two ending portions of the nanotube by varying the number of valence electrons per carbon atoms [44]. Calculations have been performed retaining only the p_z-orbital part of the

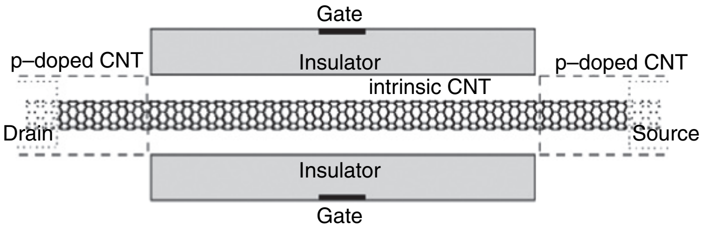


Figure 13 Schematic cross section of the coaxially gated CNTFET we have simulated

Hamiltonian. This is sufficient to accurately describe the first conduction and valence bands of the CNT (7, 0).

10.1. Screening properties of CNTs

In bulk-modulated CNTFETs, current modulation depends on the channel capability to screen the local gate field, that is, on the amount and the distribution of charge that can be locally induced on the CNT by the applied gate bias. Owing to their very small, one-dimensional density of states (DOS), nanotubes are generally not able to completely screen the gate field in the small extension of the electron gas associated to the grapheme sheet.

In this situation, the channel charge response for a given gate bias can no longer be described by using only the geometrical capacitance of the insulator. On the contrary, a correction should be brought to the insulator capacitance to account for the correct value of the induced charge. This correction is obtained by means of the so-called quantum capacitance [45]. The total capacitance of the gated nanotube system, C_Q, is given by the series connection of the geometric and the quantum capacitances [46], respectively C_ins and C_Q as $1/C_G = 1/C_\mathrm{ins} + 1/C_\mathrm{Q}$. A correct evaluation of the quantum capacitance is fundamental in order to correctly describe the charge response (and consequently the barrier height modulation properties) of a bulk-modulated CNTFET. The *gDFTB* approach naturally includes the treatment of the quantum capacitance, since the charge induced on the nanotube is computed self-consistently with the gate electostatics. Furthermore, by using an atomistic, DFT-based approach, the computed charge takes into account the DOS of the one-dimensional system, and retains, at least within a mean-field approximation, some of the electron–electron exchange and correlation effects on the nanotube screening properties. The importance of many-body corrections to the quantum capacitance of a gated nanotube has been recently pointed out in [47] and [48]. In large-diameter tubes, where the contribution of the exchange interaction to the total energy is negligible, the quantum capacitance can be estimated simply from the DOS as $C_\mathrm{Q}^\mathrm{DOS} = e^2\rho_0(\varepsilon_\mathrm{F})L$. For small-diameter tubes, due to the predominance of the exchange interaction over the kinetic energy, the quantum capacitance can instead be very different from the DOS-proportional result, and can even assume negative values. From a physical point of view, this means that the nanotube, owing to the predominance of the attractive exchange interaction among electrons, can accumulate even more charge than what strictly needed to totally screen the gate field, giving rise to a small over-screening of the gate potential in its interior.

In Figure 14 we report the computed inverse quantum capacitance of a uniformly p-doped CNT (7, 0) as a function of the Fermi energy inside the first CNT valence subband. Each value of the Fermi energy univocally corresponds to a given p-doping fraction. These results have been obtained following the methodology presented in [47], and refer to an insulator capacitance $C_\mathrm{ins} = 0.28\,\mathrm{aF}$. Solid line refers to the DFT atomistic computation, while dashed line is representative of the DOS-limited, non-interacting result. The nanotube over-screens the external gate field in the whole range of holes densities we have explored, and no positive values of the quantum capacitance have been obtained.

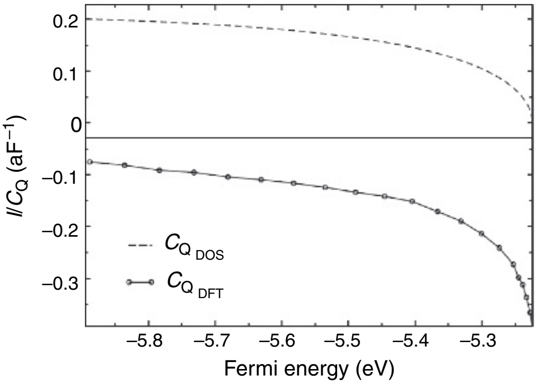


Figure 14 Computed inverse quantum capacitance of the coaxially gated CNT (7, 0) as a function of the Fermi level inside the first valence subband. Solid line refers to DFT calculations. Dashed line refers to $C_Q = e^2 \rho_0(E_F)$

10.2. Output characteristics

In Figure 15 we show the calculated output characteristics for a p-i-p CNTFET. The intrinsic channel of the simulated nanotube is 10.26 nm long. Source and drain contacts are p-doped with a carrier concentration of $2.63 \cdot 10^6\,\text{cm}^{-1}$, corresponding to a doping fraction $f = 0.004$ holes per carbon atom. In realistic devices, this carrier concentration can be easily obtained by an electrostatic doping, controlled by means of a back gate contact [49]. This doping causes the nanotube Fermi level to be shifted inside the first valence subband, at ~16 meV below the subband edge. Charge transfer at the junction between the degenerate p-type and the intrinsic portion of nanotube generates a barrier for holes transmission in valence band. The different curves shown in Figure 15 refer to

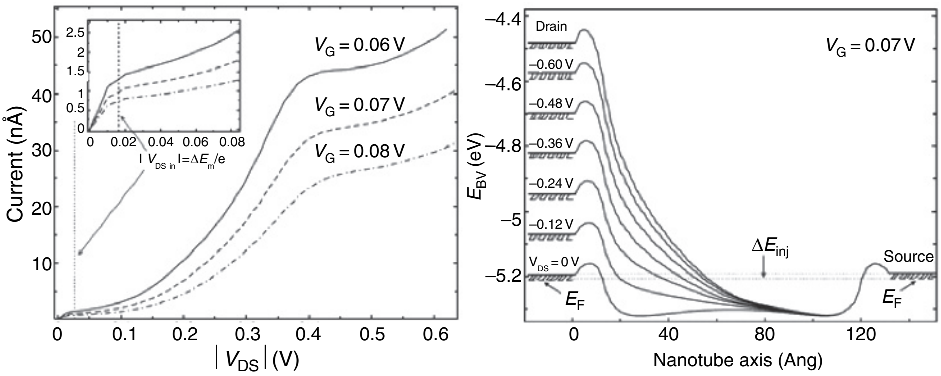


Figure 15 Computed output characteristics of the p-i-p CNTFET. Different curves correspond to different gate potentials. The inset shows details of the first two operative regimes of the device, the linear regime and the carrier injection saturation. The corresponding band profiles are shown in the right panel

different gate biases. Positive potentials applied to the gate contact increase the barrier height for holes transmission, hence reducing the current for a given drain voltage.

The physical mechanisms governing the current dependence on the applied source–drain bias can be understood referring to the band profiles shown in Figure 15 (right). These band profiles have been obtained, for a fixed gate potential, by varying the drain–source applied bias, $V_{\rm DS}$, in a range of negative values (as required for a p-type conduction device).

The current is determined by the holes transmission probability in the energy interval between the Fermi levels of the source and that of the drain contact. Although the applied bias continuously shifts the drain Fermi level with respect to the source, the nanotube bandgap reduces the energy window for hole injection to the energy interval $\Delta E_{\rm inj}$, between the source valence band edge and the source Fermi level (see Figure 15). In other words, the maximum holes injection in the channel is fixed by the source doping rate.

At first, the current linearly increases as a function of the applied bias, simply due to the increase of the energy window for carrier injection between the source and the drain Fermi levels. As the maximum energy window for carrier injection is reached, corresponding to the bias $|V_{\rm DS,inj}| = \Delta E_{\rm inj}/e$ (see also the inset of Figure 15), the current does not completely saturate due to a drain-induced modulation of the barrier width within the fixed injection interval, which considerably increases the holes transmission probability. It should be kept in mind that the fundamental reason for the behavior we observe is that the device operates in such a regime that current is entirely due to tunneling through the potential barrier.

In purely ballistic devices, it is just the barrier height which determines the device behavior. On the contrary, in tunneling regimes, the barrier width and shape are also relevant in determining the current. In the p-i-p structure we have simulated, the contact Fermi level lies closely to the valence subband edge, well below the top of the potential barrier for hole transmission (see Figure 15). This is not totally unexpected for a small-diameter nanotube junction, where charge transfer occurring at the doped-intrinsic interface is distributed over an exiguous number of carbon atoms, owing to the marked quasi-one-dimensionality of the system. A large shift in the atomic energy can also, therefore, result from a relatively small charge transfer, causing a very high potential barrier with respect to carrier injection level. This deeply differentiates the behavior of a CNTFET with respect to the ballistic transport theory developed for silicon nano-MOSFET [50]. The strong dependence of the saturation current on the drain bias we have so far observed is just caused by the exponential dependence of the tunneling current on the width of the barrier. We can refer to the mechanism responsible for the non-saturation of the current as drain-induced barrier thinning (DIBT).

Finally, it should be noted that the short length of the channel enhances the sensitivity of the tunneling current to drain-induced barrier modulation. Longer nanotubes are therefore expected to show a considerably lesser sensitivity of the saturation current to the drain bias. In Figure 16 we show the trans-characteristic obtained for the p-i-p CNTFET for a fixed drain bias of −0.4 V.

Simulation results show the exceptional transport characteristics of bulk-modulated CNTFETs. First, we note that the device shows a perfectly unipolar behavior. Current is carried only by holes which are injected from the p-doped source contact, and is progressively switched off by increasing the barrier for holes transmission with the gate

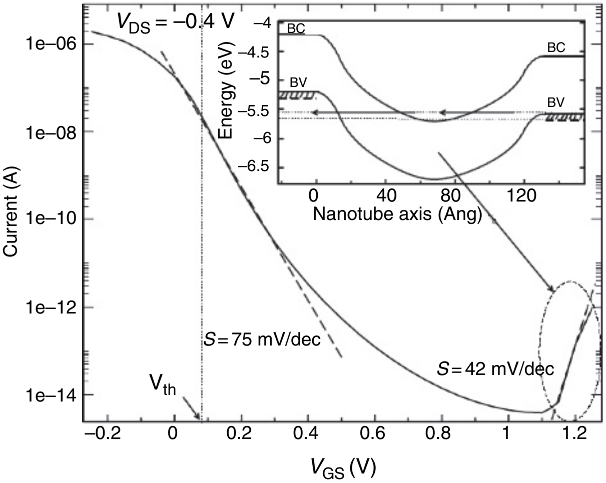


Figure 16 Trans-characteristic of a p-i-p CNTFET consisting of the semiconducting CNT (7, 0) doped at the contacts with 0.004 holes per carbon atom. The non-monotonic behavior of the current for high gate voltages is associated to band-to-band tunneling

bias. The non-monotonic behavior that can be observed in the trans-characteristic when a high gate voltage is applied is associated to band-to-band (BTB) tunneling mechanisms, which allows holes to propagate from the valence subband of source contact into the conduction band in the middle of the channel, and then to be collected into the valence subband of the drain [51] (see the inset in Figure 16). Thanks to their intrinsic unipolar behavior, vertically scaled CNTFETs accomplish one of the fundamental requirements for a high-performance device which can be used in CMOS-like logic circuits.

The computed output-characteristics show also a very high $I_{\mathrm{on}}/I_{\mathrm{off}}$ ratio, $\sim 10^8$, which is an evidence of the exceptional effectiveness of the bulk-switching mechanism in modulating the current. Exceptional performances have also been obtained concerning the sub-threshold swing parameter, $S = \mathrm{d}V_{\mathrm{G}}/\mathrm{d}(\log I_{\mathrm{DS}})$, which measures the effectiveness of the gate bias in switching off the current in the sub-threshold regime. The computed output characteristic shows a very good value for the inverse sub-threshold slope, 75 mV/dec. Note that the BTB branch of the characteristic in Figure 16 shows a sub-threshold swing of 42 mV/dec, which is even lower than the minimum value attainable for any conventional silicon MOSFET at room temperature, which is $K_{\mathrm{B}}T\ln(10)/e = 60\,\mathrm{mV/dec}$.

11. Conclusions

We have presented a detailed description of the *gDFTB* tool for quantum transport calculations in molecular systems and nanostructures. Several applications of the method have been shown, ranging from coherent transmission across molecular junctions and CNT devices to studies of inelastic electron–vibration interactions with applications to IETS. Methodology development to improve the DFT spectra using the GW approach has been also shown.

Acknowledgments

The authors wish to acknowledge the fruitful collaboration with G. Solomon and J. Reimers of the University of Sidney, Australia and with F. Pump and G. Cuniberti of the University of Regensburg, Germany. A. Gagliardi acknowledges PaSCO Graduate School for its funding.

References

[1] A. Aviram, M. A. Ratner, *Chem. Phys. Lett*. **29**, 277 (1974).
[2] M. A. Reed, C. Zhou, C. J. Muller, T. P. Burgin, J. M. Tour, *Science* (Washington, D. C.) **278**, 252 (1997).
[3] B. Xu, N. J. Tao, *Science* (Washington, D. C.) **301**, 1221 (2003).
[4] W. Wang, T. Lee, I. Kretzschmar, M. A. Reed, *Nano Lett.* **4**, 643 (2004).
[5] W. Wang, T. Lee, I. Kretzschmar, M. Reed, *Nano Lett.* **4**, 643 (2004).
[6] J. G. Kushmerick, J. Lazorcik, C. H. Patterson, R. Shashidhar, D. S. Seferos, G. C. Bazan, *Nano Lett*. **4**, 639 (2004).
[7] A. Pecchia, A. Di Carlo, *Rep. Prog. Phys.* **67**, 1497 (2004).
[8] T. Frauenheim, G. Seifert, M. Elstner, T. Niehaus, C. Kohler, M. Amkreutz, M. Sternberg, Z. Hajnal, A. Di Carlo, S. Suhai, *J. Phys.: Condensed Matter* **14**, 3015 (2002).
[9] A. Pecchia, A. Di Carlo, *Introducing Molecular Electronics*, Springer 2005, G. Cuniberti, G. Fagas, K. Richter, eds.
[10] M. Elstner, D. Prezag, G. Jugnickel, J. Elsner, M. Haugk, T. Frauenheim, S. Suhai, G. Seifer, *Phys. Rev. B* **58**, 7260 (1998).
[11] P. Maragakis, R. L. Barnett, E. Kaxiras, M. Elstner, T. Frauenheim, *Phys. Rev. B* **66**, 241 104 (2002).
[12] D. Porezag, M. R. Pederson, T. Frauenheim, T. Köhler, *Phys. Rev. B* **52**, 14 963 (1995).
[13] C. Caroli, R. Combescot, P. Nozieres, D. Saint-James, *J. Phys. C: Solid State Phys.* **5**, 21 (1972).
[14] R. Lake, S. Datta, *Phys. Rev. B* **45**, 6670 (1992).
[15] M. Galperin, M. Ratner, A. Nitzan, *Nano Lett.* **4**, 1605 (2004).
[16] S. V. Faleev, M. I. Stockman, *Phys. Rev. B* **66**, 085 318 (2002).
[17] A. L. Fetter, J. D. Walecka, *Quantum Theory of Many Particle Systems* (Dover Publications, 1971).
[18] D. M. Mahan, *Many Particle Physics* (Plenum Press, 1981).
[19] H. Haung, A. P. Jauho, *Quantum Kinetics in Transport and Optics of Semiconductors*, Vol. 123 (Springer Series in Sol. State Sci., 1993)
[20] T. N. Todorov, *J. Physics: Condens. Matter* **14**, 3049 (2002).
[21] S. Datta, *Electronic Transport in Mesoscopic System* (Cambridge University Press, 1995)
[22] T. N. Todorov, J. Hoekstra, A. P. Sutton, *Phil. Mag. B* **80**, 421 (2000).
[23] L. V. Keldysh, *Sov. Phys. JEPT* **20**, 1018 (1965).
[24] L. P. Kadanoff, G. Baym, *Quantum Statistical Mechanics* (W. A. Banjamin, 1962).
[25] S. Datta, *Superlattices and Microstructures* **28**, 253 (2000).
[26] C. Caroli, R. Combescot, P. Nozieres, D. Saint-James, *J. Phys. C: Solid State Phys.* **4**, 916 (1971).
[27] M. Elbing, R. Ochs, M. Köntopp, M. Fischer, C. von Hänisch, F. Weigend, F. Evers, H. B. Weber, M. Mayor, *PNAS* **102**, 8815 (2005).
[28] M. Galperin, M. Ratner, A. Nitzan, *J. Chem. Phys.* **121**, 11 965 (2004).
[29] Y. C. Chen, M. Zwolak, M. Di Ventra, *Nano Lett.* **4**, 1709 (2004).
[30] A. Pecchia, A. Gagliardi, S. Sanna, T. Frauenheim, A. Di Carlo, *Nano Lett.* **4**, 2109 (2004).

[31] G. C. Solomon, A. Gagliardi, A. Pecchia, T. Frauenheim, A. Di Carlo, J. R. Reimers and N. S. Hush, *J. Chem. Phys.* (2006)

[32] A. Pecchia, M. Gheorghe, L. Latessa, A. Di Carlo, *J. Comp. Elec.* **2**, 251 (2003).

[33] N. Sai, M. Zwolak, G. Vignale, M. Di Ventra, *Phys. Rev. Lett.* **94**, 186 810 (2005).

[34] A. Ferretti, A. Calzolari, R. D. Felice, F. Manghi, M. J. Caldas, M. B. Nardelli, E. Molinar, *Phys. Rev. Lett.* **94**, 116 802 (2005).

[35] C. Toher, A. Filippetti, S. Sanvito, K. Burke, *Phys. Rev. Lett.* **95**, 146 402 (2005).

[36] K. Burke, R. Car, R. Gebaue, *Phys. Rev. Lett.* **94**, 146 803 (2005).

[37] L. Hedin, *Phys. Rev.* **139**, A 796 (1965).

[38] F. Aryasetiawan, O. Gunnarsson, *Rep. Prog. Phys.* **61**, 237–312 (1998).

[39] G. Onida, L. Reining, A. Rubio, *Rev. Mod. Phys.* **74**, 601 (2002).

[40] M. Rohlfing, P. Kruger, J. Pollmann, *Phys. Rev. B* **57**, 6485 (1998).

[41] T. A. Niehaus, M. Rohlfing, F. Della Sala, A. Di Carlo, T. Frauenheim, *Phys. Rev. A* **71**, 022 508 (2004).

[42] P. L. McEuen, M. S. Fuhrer, H. Park, *IEEE Trans. Nanotechnol.* **1**, 78 (2002).

[43] J. Appenzeller, J. Knoch, V. Derycke, R. Martel, S. Wind, P. Avouris, *Rep. Prog. Phys.* **67**, 1 (2004).

[44] F. Leonard, J. Tersoff, *Phys. Rev. Lett.* **83**, 5174 (1999).

[45] S. Luryi, *Appl. Phys. Lett.* **52**, 501 (1988).

[46] M. Büttiker, *J. Phys.: Condens. Matter* **5**, 9361 (1993).

[47] L. Latessa, A. Pecchia, A. Di Carlo, P. Lugli, *Phys. Rev. B* **72**, 035 455 (2005).

[48] M. M. Fogler, *Phys. Rev. Lett.* **94**, 056 405 (2005).

[49] Y. M. Lin, J. Appenzeller, J. Knoch, P. Avouris, cond-mat/0501690 (2005)

[50] K. Natori, *J. Appl. Phys.* **76**, 4879 (1994).

[51] J. Appenzeller, Y. M. Lin, J. Knoch, P. Avouris, *Phys. Rev. Lett.* **93**, 196 805 (2004).

Molecular and Nano Electronics: Analysis, Design and Simulation
J. M. Seminario (Editor)

Chapter 9

Theory of quantum electron transport through molecules as the bases of molecular devices

M. Tsukada[a], K. Mitsutake[b] and K. Tagami[a]

[a]*Graduate School of Science and Engineering, Waseda University, 513 Waseda Tsurumaki-cho, Shinjuku-ku, Tokyo 162-0041, Japan. m-tsukada@fiberbit.net*
[b]*Canon Research Center, Canon Inc., 3-30-2 Simomaruko, Ohta-ku Tokyo, 162-0041, Japan*

1. Introduction

Electron transport properties through molecular bridges connecting nano-scale electrodes have been among the hot topics of molecular electronics [1]. Towards realization of molecular bridge devices, ultra-fine bottom-up fabrication technologies should be developed. In spite of its strong expectation, however, to realize the single molecular devices is not easy and the fabrication and characterization methods have not been well developed so far.

In such a situation, the role of the theoretical studies is significant, for example, to propose a useful contact formation method, to analyze observed data properly, and to explore functionality of molecules as devices [2]. In this chapter, we will report some recent attempts to explore the remarkable properties of quantum transport through molecular bridges.

For a theoretical analysis of molecular bridges, we mainly use the non-equilibrium Green's function method with tight-binding basis [2]. If necessary, a self-consistent calculation in the standard density-functional-theory (DFT) level [3], or its simplified version [4], can be incorporated in this approach.

As will be discussed later, the transport through molecular bridges is strongly influenced by the connecting part to the electrodes, though of course it is also influenced by the internal structure of the molecule. The internal current distribution within the molecule, which is induced by the source-drain current, shows remarkable quantum nature and will be one of the topics discussed in this chapters. When the electron incident energy is close to a degenerate molecular level, a large loop current is often generated inside the molecule [5]. The internal current distribution is similar to that induced by the magnetic field.

According to the principle of the quantum mechanics, since the number of the electrons in the molecular island and the phase of the wavefunction in a bridge region are a pair of conjugate physical quantities, there is an uncertainty relation between the two. If the bottleneck at the connection part is weak, the phase tends to be a good quantum number, and a coherent quantum transport is expected. If the bottleneck is strong enough, the electron number in the molecular island will be more or less well defined. In this case the electron transfer is associated with the energy dissipation to the phonon system or the electro-magnetic environment. The crucial problem is to know which regime, coherent or dissipative, dominates the electron transport. Moreover, it is an interesting problem to know the nature of the transition between the two regimes and to explore the marginal situation. For the study of this problem, effects of the electron coupling with medium degrees of freedom should be seriously considered.

2. Non-equilibrium Green's function with tight-binding bases

In this section, as the basic theoretical approach for the quantum transport, non-equilibrium Green's function method is briefly summarized. Hamiltonian $\mathbf{H}$ of the whole system is written as the sum of the Hamiltonian $\mathbf{H}^0$ of the free molecule and its interaction with the environment Σ,

$$\mathbf{H} = \mathbf{H}^0 + \Sigma \tag{1}$$

The self-energy term Σ takes into account the interaction with electrodes, many-body effects such as electron–phonon and electron–electron interactions. The retarded and advanced Green's functions are defined by

$$\mathbf{G}^{\mathrm{R}}(E) = \left\{E\mathbf{I} - \mathbf{H}^0 - \Sigma^R + i\eta\right\}^{-1} \tag{2}$$

$$\mathbf{G}^{\mathrm{A}}(E) = \mathbf{G}^{\mathrm{R}}(E)^{\dagger}, \tag{3}$$

with a positive infinitesimal η. The lesser Green's function is defined by

$$\mathbf{G}^{<}(E) = \mathbf{G}^{\mathrm{R}}(E)\Sigma^{<}(E)\mathbf{G}^{\mathrm{A}}(E) \tag{4}$$

which is used to obtain the current between the site i to j,

$$J_{ji}(E) = \frac{2e}{h} Re\left[H_{ji} G_{ji}^{<}(E)\right] \tag{5}$$

In the above, H_{ji} is the (j,i) element of the Hamiltonian $\mathbf{H}^0$. All the effects due to the interactions with the environment are included in $\Sigma^{\mathrm{R(A)}}$ and the lesser self-energy $\Sigma^{<}$. When the many-body effect and the electron–phonon coupling are neglected, the self-energy comes only from the coupling with leads. In this case $\Sigma^{<}$ is given by

$$\Sigma^{<}(E) = i\sum_{\alpha} f_{\alpha}(E)\Gamma_{\alpha}(E) \tag{6}$$

$$\Gamma_{\alpha}(E) = i\sum\nolimits_{\alpha}^{\mathrm{R}}(E) - \sum\nolimits_{\alpha}^{\mathrm{A}}(E) \tag{7}$$

where f_{α} and $\Sigma_{\alpha}^{\mathrm{R(A)}}$ denote the Fermi distribution function in the lead α, and the retarded (advanced) self-energies, respectively. The transmission probability from lead α to α' is given by

$$T_{\alpha\alpha'}(E) = \mathrm{Tr}\left[\mathbf{G}^{\mathrm{R}}\mathbf{\Gamma}_{\alpha}\mathbf{G}^{\mathrm{A}}\mathbf{\Gamma}_{\alpha'}\right] \tag{8}$$

The Hamiltonian $\mathbf{H}^{0}$ can be obtained in a self-consistent manner in the framework of the DFT with the localized basis set. For the simplest approximation, we use phenomenological tight-binding model.

3. Effects of the linkage structure on the conductance of molecular bridges

Here we consider how the terminal structures of the molecules affect the conductance of the molecular bridge. The transmission spectra of the phenalenyl molecule [6] and the tape porphyrin molecule [7] are examined as case study. In all the cases, the electronic states of the whole systems are described by the density functional–derived self-consistent tight-binding method [8–10], in which the basis is assumed to be the localized atomic orbitals.

Figure 1 shows the molecular bridge of the phenalenyl molecule attached to the two gold electrodes through the mercapto-vinyl groups [6]. The difference between the two systems in (a) and (b) of Figure 1 is seen in the positions of the connecting sites of the leads (mercapto-vinyl groups), i.e., α and β sites for the corresponding cases.

The transmission spectrum when both leads are attached to the α sites is shown in Figure 2(a), while for the case with the leads attached to the β site it is shown in Figure 2(b). In both figures, the origin of the electron energy is the Fermi level of the gold electrodes. For the case with the leads at the α sites, the transmission peak appears

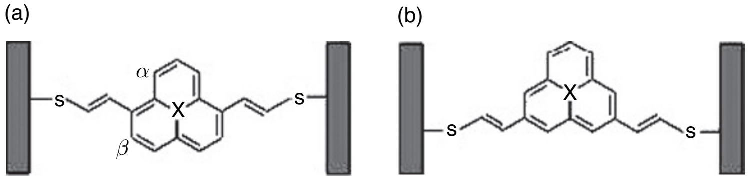


Figure 1 Structure of phenalenyl molecular bridge (X=C atom). The mercapto-vinyl groups are attached to (a) α site and (b) β site

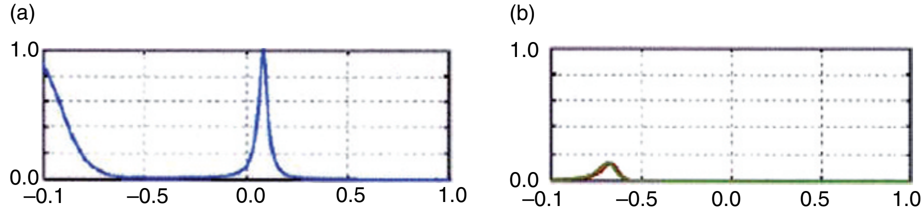


Figure 2 The transmission spectra for the phenalenyl molecular bridges shown in Figure 1. Labels (a) and (b) correspond to those of Figure 1

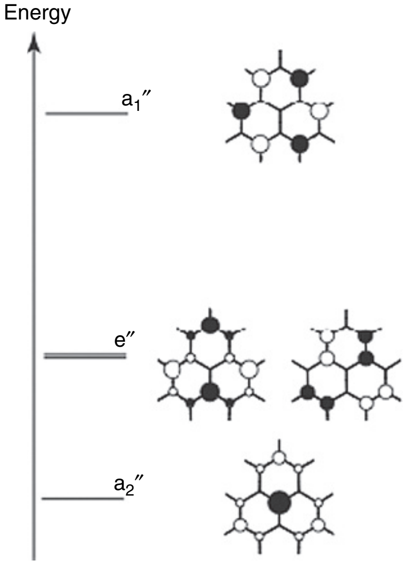


Figure 3 The energy diagram (schematic) and wave-functions of phenalenyl molecule

at the very vicinity of the Fermi level, but not for the case with the leads at the β sites. In the latter case, the prominent peak at around the Fermi level disappears.

Figure 3 shows the energy diagram of the free phenalenyl molecule. The transmission peak in the vicinity of the Fermi level is caused by α_1'' orbital which has no amplitudes at the β sites. This orbital is a single occupied molecular orbital (SOMO) of phenalenyl. It is easily understood from the nature of the SOMO orbital that the orbital α_1'' does not contribute to the transmission, if the leads are connected to β sites. This is the reason for the disappearance of the transmission peak near the Fermi level in Figure 2b. Such a sensitivity of the conductance on the terminal sites has been also found in the tape-porphyrin molecules [11]. Tape porphyrin is a sort of the oligomer of the porphyrin, which shows the vanishing of the HOMO–LUMO gap for a very long chain length. We investigated four different linkage structures of the tape-porphyrin to the gold electrodes as shown in Figure 4. For all the cases, the numbers of the porphyrin molecules in the chain is assumed to be 8 ($n=6$). The calculated results of the transmission spectra

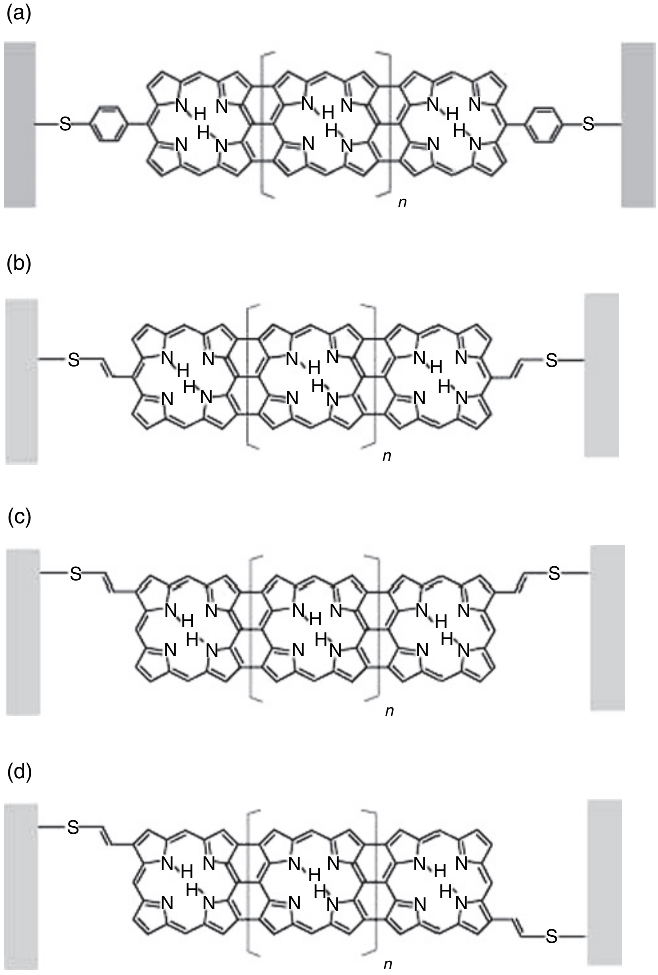


Figure 4 Four kinds of bridge structures using a tape porphyrin

are shown in Figure 5. It is remarkable that only a small difference of the linkage part dramatically influences the conductance of the molecular bridge. Namely, among the four bridges from A through D shown in Figure 4, the case B shows the largest conductance, because there is a prominent transmission peak very close to the electrode Fermi level.

These findings indicate that for designing a whole system of the molecular bridge, special attention should be paid to the linkage parts of the molecule to the electrodes.

4. Internal large loop currents

Figure 6 illustrates a triangular nano-graphene sheet bound with zigzag edges. The protruded atoms along the topmost zigzag edge are numbered from 1 to N. The number of atoms in the nano-graphene is $N^2 + 4N + 1$. The smallest molecule with $N = 2$

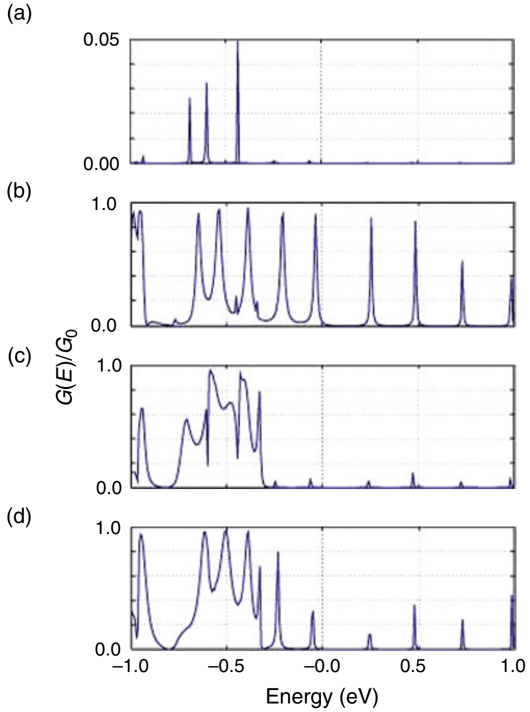


Figure 5 Transmission spectra for the corresponding molecular bridges shown in Figure 4

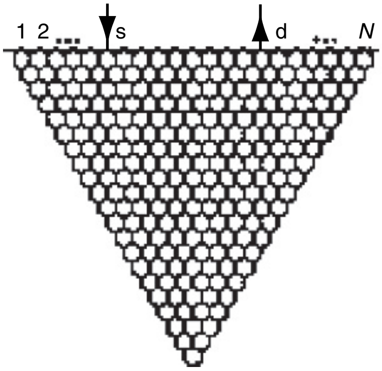


Figure 6 A structure of triangular nano-graphene

corresponds to the phenalenyl molecule discussed in the previous section, but here we focus on the larger system with $N = 56$. The source s and the drain d are connected to the top edge at $(s, d) = (N/4 + 1, 3N/4)$. The connection would be made by the vinyl group. As for the lead-molecule coupling, we set $t' = 0.6t$.

The isolated nano-graphene molecule has many doubly degenerate energy levels in the vicinity of the Fermi level [12], although there are $(N - 1)$-fold degenerate levels at

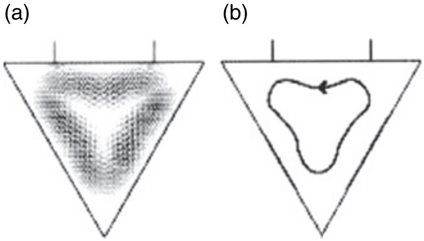


Figure 7 (a) Microscopic loop current distribution in the trianglar nano-graphene and (b) a schematic view of its current direction

$E = 0$, which correspond to the edge states. Due to the localized character of the edge state, they do not contribute to the resonant current [12]. Figure 7(a) shows the internal bond current $J_{ij}(E)$ (from site i to j) at one of the degenerate levels $E = -0.165t$, so that the strongest one is expressed by the darkest color. Figure 7(b) schematically illustrates the orientation of the current flow. Noticeably, a large current loop appears circulating a large area of the graphene molecule. The strength of the current is much larger than the source–drain current. Even when the source and the drain sites are moved to different sites along the upper edge, almost the same current patterns are obtained [12]. The internal current distribution is determined from the nature of the particular degenerate molecular orbitals, rather than being accidentally induced by the current from the leads.

The distribution of the internal bond current is analyzed as follows. The current flowing through the bond connecting the sites i and j is given by

$$\begin{aligned} j_{ij} &= \frac{4e}{h}\mathrm{Im}\left[\Psi_i^* H_{ij} \Psi_j\right] \\ &= \frac{4eH_{ij}}{h}\sum_{\mu\nu}\left(\phi_{\mu i}\phi_{\nu j} - \phi_{\nu i}\phi_{\mu j}\right)\left|a_\mu\right|\left|a_\nu\right|\sin(\theta_\mu - \theta_\nu) \end{aligned} \tag{9}$$

where μ, ν label the molecular states, $\left|a_\mu\right|, \left|a_\nu\right|$are their amplitudes, and θ_μ, θ_ν are the phases of the coefficients expanding the whole scattering electron wave with the energy E. On the other hand $\left\{\phi_{\mu i}\right\}_{i=1,2,\ldots}$ are the LCAO coefficients of the particular molecular orbital μ in the isolated molecule. The envelop factor of $\left|a_\mu\right|\left|a_\nu\right|\sin\left(\theta_\mu - \theta_\nu\right)$ does not depend on the atomic site in the molecule, and resonantly enhanced when the electron energy becomes closer to a degenerate molecular level. In this case, the states μ and ν are certain components of the degenerate state. Because of the sine function factor of the phase difference $\theta_\mu - \theta_\nu$, the energy corresponding to the maximum of the envelop factor is slightly higher or lower than the just on resonance energy. Because of the resonantly enhancing feature of the envelop factor, the magnitude of the loop current becomes larger than the source–drain current often by several tens of times.

The factor $\left(\phi_{\mu i}\phi_{\nu j} - \phi_{\nu i}\phi_{\mu j}\right)$ in Eq. (9) expresses the microscopic current distribution inside the molecule. Interestingly enough, the same factor appears for the current distribution of the isolated molecule under a static magnetic field. Such current causes the diamagnetism of the molecule in general, and for a mesoscopic system, it is the current distribution of the magnetically induced persistent current. One might speculate a close relationship between the source–drain induced loop current and the magnetic field–induced molecular current.

5. Effect of electron–phonon coupling

Though the conductance value of molecular bridges so far theoretically predicted has not been so much reduced from the quantization value G_0 $(= 1/12.9\,\mathrm{k}\Omega)$, experimentally reported values have been usually by some orders of magnitude smaller than this. Unfortunately there have been few conductance measurements, where the contact is characterized in the atomic level. On the theoretical side, approaches ignoring the electron–vibration coupling may not be always satisfactory. Such approaches might result in too much enhancement of the coherent nature and tends to estimate larger values for the conductance. To clarify the reason for the gap between the theory and experiment, we should look into the effect of the molecular vibration on the electron transfer processes.

With the decrease of the transfer integral in the molecular region, the electronic states would change from the extended states to the localized states due to the polaron formation. By considering the transition between these polaron-like states, the carrier transfer in the molecular bridges can be properly analyzed. In this section we propose a unified method to treat the extended and the polaron-like localized states coupled with the molecular vibrations. Second, we will clarify the electron transport processes based on the transition rates between the coupled electronic states with the vibration modes. The case study is made for the one-dimensional molecular bridge made of the thiophene molecules [13].

The model we treat in this chapter is shown in Figure 8. It consists of molecules arranged in a chain in the bridge region and two outer electrodes. The molecules should be regarded to form a large single molecule for the stronger interaction. We consider the case in which one additional electron or hole exists over the otherwise neutral whole molecules. Only a single state, i.e. LUMO for the case of electron or HOMO for the case of hole, is considered as the electronic state on each molecule (or molecular unit for the stronger interaction case).

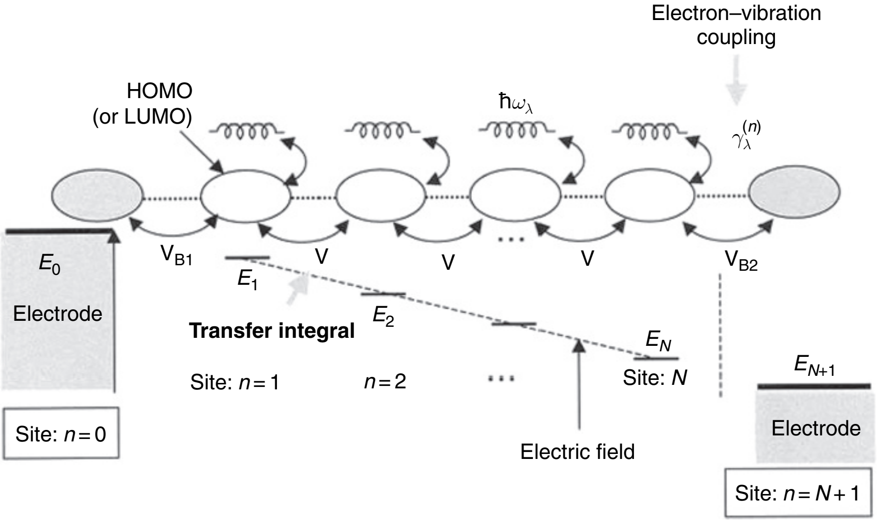


Figure 8 A model of the bridge of linearly arranged thiophene molecules

The Hamiltonian of the total system is given as

$$\mathbf{H}_{\text{total}} = \mathbf{H}_{\text{MB}} + \mathbf{H}_{E} + \mathbf{H}_{\text{MB-E}} \tag{10}$$

where the first two terms of the right hand side describe molecule(s) and electrodes, respectively, and the last term is their coupling term. Each term on the right hand side of the above equation is written as

$$\mathbf{H}_{\text{MB}} = \sum_{n=1}^{N} E_n a_n^+ a_n + \sum_{n=1}^{N-1} V(a_n^+ a_{n+1}^+ a_{n+1}^+ a_n)$$

$$+ \sum_{\lambda} h\omega_\lambda b_\lambda^+ b_\lambda + \sum_{n=1}^{N} \sum_{\lambda} \gamma_\lambda^{(n)} h\omega_\lambda a_n^+ a_n (b_\lambda^+ + b_\lambda) \tag{11a}$$

$$\mathbf{H}_{\text{E}} = E_0 a_0^+ a_0 + E_{N+1} a_{N+1}^+ a_{N+1} \tag{11b}$$

$$\mathbf{H}_{\text{MB-E}} = V_{B1}(a_0^+ a_1 + a_1^+ a_0) + V_{B2}(a_N^+ a_{N+1} + a_{N+1}^+ a_N) \tag{11c}$$

Here $a_n^+(a_n)$ creates (annihilates) an electron or hole in the n-th molecule (hereafter, “molecule” should be interpreted as a molecular unit for a strong V case) or one of the electrodes with energy E_n.

The indexes $n = 0$ and $n = N + 1$ correspond to the electronic state of the electrodes. Since the number of carriers is assumed to one, $\sum_{n=0}^{N+1} a_n^+ a_n = 1 \cdot b_\lambda^+(b_\lambda)$ creates (annihilates) a vibration quanta in the mode λ with energy $\hbar\omega_\lambda$. The electron–vibration coupling between the carrier at the n-th molecular site and λ-th vibration mode is given by the parameter $\gamma_\lambda^{(n)}$.

For the estimation of the electron–phonon coupling parameter, the following relation is useful;

$$\gamma_\lambda^{(n)} = -\sqrt{\frac{m_\lambda \omega_\lambda}{2h}}\, \Delta\vec{d} \cdot \vec{e}_\lambda \tag{12}$$

with $\Delta\vec{d}$ being the $3N_{\text{atom}}$-dimensional relaxation vector from neutral state to ionic state (N_{atom} is the number of atoms in the molecule), and $\vec{e}_\lambda$ being the λ-th normal vibration mode of the molecule. In the case of the thiophene molecule, among 21 normal vibration modes, the 16th mode from the lowest has the largest coupling. This is because the displacement of the 16th mode is nearly the same as the relaxation vector. Thus the value of the coupling constant is calculated by *ab initio* molecular orbital method.

Transfer integral V is also estimated by *ab initio* molecular orbital method. Examples of the transfer integral of the HOMOs of thiophene molecules are shown in Figure 9.

The electronic states in the molecular bridge region can be obtained as follows. First, in the limiting case of $V = 0$, the exact wavefunctions are given by

$$\phi_n^{m_\lambda} = \prod_{\lambda} \exp(-\gamma_\lambda^{(n)}(b_\lambda^+ - b_\lambda))\, |m_\lambda\rangle\, \Phi^{(n)} \tag{13}$$

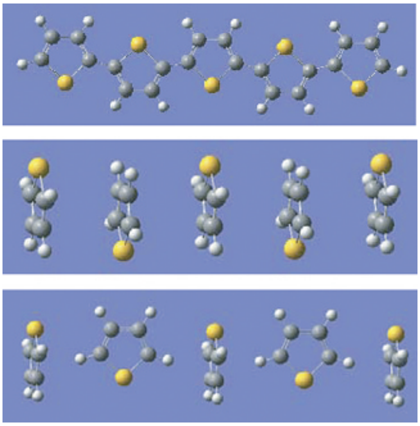


Figure 9 Values of transfer integral for the linear array of thoiphene molecules

which will be called "the polaron-like states", hereafter. Here "$|m_\lambda\rangle$" means the vibration state with m_λ quanta in the mode λ. Second, if the electron–vibration coupling is equal to 0, the eigen states should be in the form of

$$\chi_m^{\{m_\lambda\}} = \left\{ \sum_{n=1}^{N} f_n^m \Phi^{(n)} \right\} |\{m_\lambda\}\rangle \, . \tag{14}$$

which are called the "the undressed states". The first factor of $\left\{ \sum_{n=1}^{N} f_n^m \Phi^{(n)} \right\}$ corresponds to the molecular orbital for the system without vibration degrees of freedom. The state $|\{m_\lambda\}\rangle$ means $|m_1\rangle\,|m_2\rangle \cdots |m_n\rangle \cdots$. Generally, the wavefunction is expressed by

$$\psi_a^{\{m_\lambda^a\}} = \sum_{n=1}^{N} A_n^a \phi_n^{\{m_\lambda^a\}} + \sum_{m=1}^{N} B_m^a \chi_m^{\{m_\lambda^a\}} \tag{15}$$

with A_n^a and B_m^a determined by the minimum energy condition for the ground states:

$$\psi_a^{\{0\}} = \sum_{n=1}^{N} A_n^a \phi_n^{\{0\}} + \sum_{m=1}^{N} B_m^a \chi_m^{\{0\}} = \sum_{n=1}^{N} \left[A_n^a \prod_{\lambda} \exp(-\gamma_\lambda^{(n)}(b_\lambda^+ - b_\lambda)) + C_n^a \right] \Phi^{(n)} \left|0_{\mathrm{ph}}\right\rangle \tag{16}$$

Here we set $C_n^a = \sum_{m=1}^{N} f_n^m B_m^a$.

The electronic states in the molecular bridges with five sites will be shown below. In this model the energy difference between neighboring sites is set to 0.04 eV (see Figure 8). The single vibration mode is assumed for each molecular units, with the value of the coupling being 1.6. Figure 10 shows how the energy levels of the electronic states depend on the transfer integral. There are two kinds of states in this figure: one with almost constant energies and the other showing the linear dependence with transfer integral.

In Figure 11 the coefficients of the wavefunction are shown in the case of (a) $|V| = 0.1$ eV and (b) $|V| = 0.7$ eV. The coefficients of the polaron-like part and the undressed

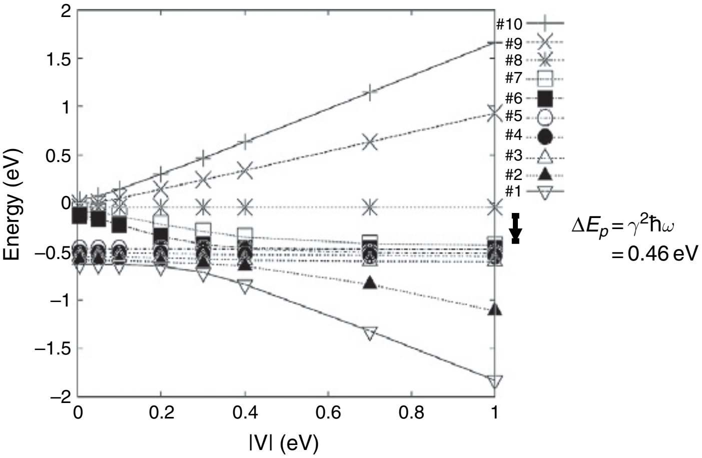


Figure 10 Energy levels of the molecular bridges as the function of the transfer integral V. The parameters are so chosen to satisfy, $\gamma^2\hbar\omega = 0.46\,\mathrm{eV}$

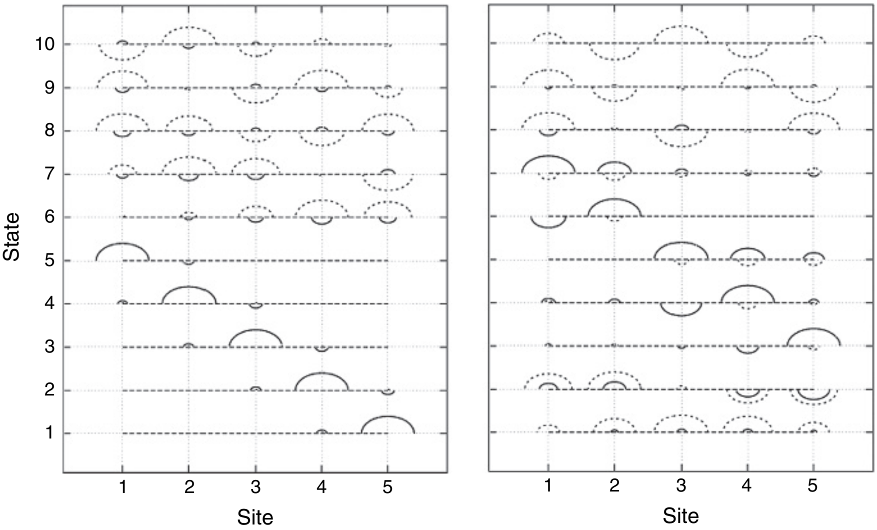


Figure 11 Composition of each eigen-state coupled with vibration modes for (a) $|V| = 0.1\,\mathrm{eV}$ and (b) $|V| = 0.7\,\mathrm{eV}$. Other parameters are $\hbar\omega = 0.18\,\mathrm{eV}$ and $\gamma = 1.6$

part are plotted by full and dotted lines, respectively, with the upper and lower half circles indicating positive and negative values, and the radius being the relative magnitude normalized by the maximum value. In case (a), the lower and upper five states are the polaron and the undressed states, respectively. On the other hand, in case (b), from the third to the seventh, the states are the polaron states, while the other states are the undressed ones. The ground state for lower $|V|$ is polaron, but with the increase of $|V|$, the ground state becomes undressed state.

6. Carrier transport process

In this section, the carrier transport in the molecular bridge is discussed in the model of the previous section. Here the states of the electrodes are expressed as

$$\psi_{\text{left}}^{\{m_\lambda^{\text{L}}\}} = \Phi^{(0)} \left|\{m_\lambda^{\text{L}}\}\right\rangle \text{ (for left electrode)} \tag{17a}$$

$$\psi_{\text{right}}^{\{m_\lambda^{\text{R}}\}} = \Phi^{(N+1)} \left|\{m_\lambda^{\text{R}}\}\right\rangle \text{ (for right electrode)} \tag{17b}$$

Transition rates from a state “a” to another state “b” is calculated by Fermi’s golden rule as

$$W_{\text{a}\to\text{b}} = \frac{2\pi}{h} \sum_{\{m_\lambda^{\text{b}}\}} \left\langle \left|\mathbf{H}_{\text{ab}}\right|^2 \delta\left(E_{\text{a}} - E_{\text{b}}\right) \right\rangle_{\{m_\lambda^{\text{a}}\}} \tag{18}$$

The state “a” or “b” is defined in Eq. (15) or (17a) or (17b), and the matrix element $\mathbf{H}_{\text{ab}}$ and the energy E_{a} are calculated by

$$E_a = \left\langle \Psi_{\text{a}}^{\{m_\lambda^{\text{a}}\}} \middle| \mathbf{H}_{\text{total}} \middle| \Psi_{\text{a}}^{\{m_\lambda^{\text{a}}\}} \right\rangle \tag{19}$$

$$\mathbf{H}_{\text{ab}} = \left\langle \Psi_{\text{a}}^{\{m_\lambda^{\text{a}}\}} \middle| \mathbf{H}_{\text{total}} \middle| \Psi_{\text{b}}^{\{m_\lambda^{\text{b}}\}} \right\rangle \tag{20}$$

In Eq. (18), $\langle \ldots \rangle_{\{m_\lambda^{\text{a}}\}}$ means the thermal average with the vibration state of “a”. In the following, temperature is set to 0 Kelvin.

The carrier transport can be analyzed from the population probability $\{P_{\text{a}}(t)\}$ obtained by the master equation,

$$\frac{\mathrm{d}P_{\text{a}}(t)}{\mathrm{d}t} = \sum_{\text{b}} W_{\text{b}\to\text{a}} P_{\text{b}}(t) - \sum_{\text{b}} W_{\text{a}\to\text{b}} P_{\text{a}}(t) \tag{21}$$

Figures 12(a) and (b) shows the time-dependent occupation of each site and state, respectively. The 0th and 6th sites in Figure 12(a) means the left or right electrode. Before $t = 0$ the carrier is assumed to be in the left electrode. At $t = 0$, we started to solve the master equation of Eq. (12). In other words, the injection from the left electrode to the molecular bridges started at $t = 0$. When time goes by, the occupation at the left electrode rapidly decreases, and the occupation at the sites 1, 2, 4, 5 increases as shown in Figure 12(a). After 0.1 fsec, the occupation of the 6th site at the right electrode increases and finally becomes “1”.

The time-dependent occupation of each state is shown in Figure 12(b). At first, the 8th state is occupied and then the 7th , 5th and 2nd states are occupied. After these processes, the 0th state in right electrode is finally occupied. We analyze these processes by the wavefunction and the energy levels. The wavefunctions in the molecular bridges are shown in Figure 12(c). Notice that because a different value of the coupling parameter γ is assumed ($\gamma = 0.8$), the composition of each state is somewhat different from the case of Figure 11 ($\gamma = 1.6$). The energy level of each state from 0 through 11 in the electrodes and the molecules are shown in Figure 12(d). From Figure 12(c) and (d), the 8th state has a large amplitude at the 1st site, i.e., the site closest to the left electrode and its

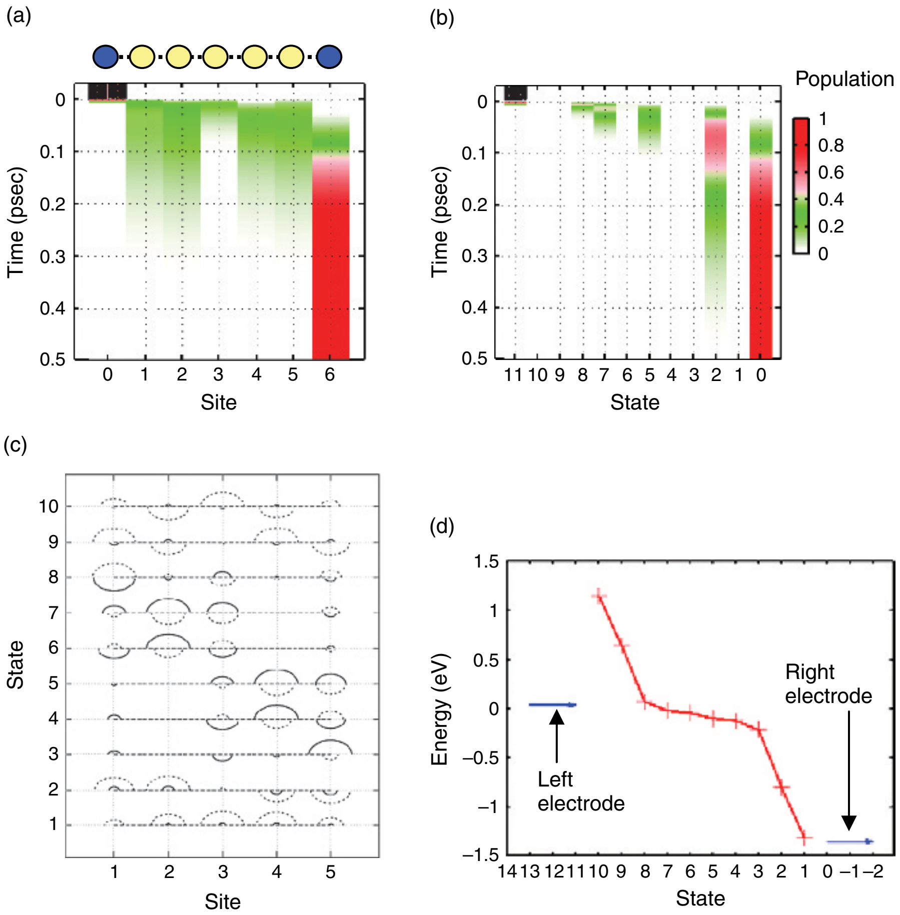


Figure 12 Time-dependent occupation of each (a) site and (b) state, for the case with the state components and the levels are shown in (c) and (d), respectively

energy is near the level of the left electrode. So it is easy to understand why the 8th state becomes occupied at the earliest stage. By analyzing the relative magnitudes of the transition rate $W_{i\to k}$, the most probable path is 11th (left) → 8th → 7th → 5th → 2nd→ 1st→ 0th (right). This is consistent with the result of Figure 12. The rate of $W_{2\to 1}$ has a smaller value and this is a rate-limiting step. The reasons for the small $W_{2\to 1}$ are: first, the 2nd state has a node at the 3rd site while the 1st state does not have a node: second, a large energy gap exists between the two states 1 and 2, which suppresses the transition. The transition rates $W_{7\to 6}$ and $W_{5\to 4}$ are also negligibly small, because of the almost orthogonal character between the states 6 and 7, or between the states 5 and 4. In Figure 13, the center of the charge and its velocity are shown as functions of time. For bulk organic materials, time-of-flight (TOF) experiments are frequently used to estimate the carrier mobility. The result of the velocity in Figure 13 is similar to that observed by TOF. Therefore, this might indicate the method of the estimation for the carrier mobility in the molecular bridges described above is legitimate.

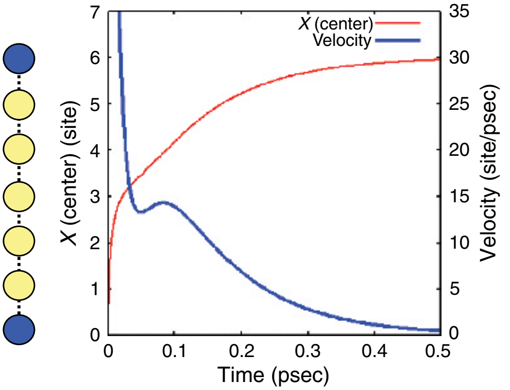


Figure 13 The time dependence of the center of mass position and the velocity of the carrier

References

[1] James M. Tour, *Molecular Electronics*, World Scientific, 2003.
[2] M. Tsukada, K. Tagami, K. Hirose and N. Kobayashi, *J. Phys. Soc. Jpn.*, 74 (2005) 1079.
[3] J. Taylor, H. Guo and J. Wang, *Phys. Rev. B* 63 (2001) 245407.
[4] S. Datta, *Electronic Transport in Mesoscopic System* (Cambridge University Press, Cambridge, U.K., 1995).
[5] K. Tagami and M. Tsukada, *Curr. Appl. Phys.* 3 (2003) 439.
[6] K. Tagami L. Wang and M. Tsukada, *Nano Lett.* 4 (2004) 209.
[7] K. Tagami and M. Tsukada, e-J.Surf. *Sci. Nanotechnol.* 2 (2004) 186.
[8] T. Frauenheim, G. Seifert, M. Elstner, T. Niehaus, C. Köhler, M. Amkreutz, M. Sternberg, Z. Hajnal, A. D. Carlo and S. Suhai, *J. Phys.: Condens. Matter*, 14 (2002) 3015.
[9] K. Tagami and M. Tsukada, *Jpn. J. Appl. Phys.* 42 (2003) 3606.
[10] K. Tagami, M. Tsukada, Y. Wada, T. Iwasaki and H. Nishide, *J. Chem. Phys.* 119 (2003) 7491.
[11] K. Tagami and M. Tsukada, *e-J. Surf. Sci. Nanotechinol.* 1 (2003) 945.
[12] K. Tagami and M. Tsukada, *e-J. Surf. Sci. Nanotechnol.* 2 (2004) 205.
[13] K. Mitsutake and M. Tsukada, *e-J. Surf. Sci. Nanotechnol.* 4 (2006) 311.

Molecular and Nano Electronics: Analysis, Design and Simulation
J. M. Seminario (Editor)

Chapter 10

Time-dependent transport phenomena

G. Stefanucci[a], S. Kurth[a], E. K. U. Gross[a] and A. Rubio[b]

[a] *Institut für Theoretische Physik, Freie Universität Berlin, Arnimallee 14, D-14195 Berlin, Germany and European Theoretical Spectroscopy Facility (ETSF)*
[b] *Departamento de Física de Materiales, Facultad de Ciencias Químicas, UPV/EHU, Unidad de Materiales Centro Mixto CSIC-UPV/EHU and Donostia International Physics Center (DIPC), San Sebastián, Spain, and European Theoretical Spectroscopy Facility (ETSF). hardy@physik.fu-berlin.de*

1. Introduction

The aim of this chapter is to give a pedagogical introduction to our recently proposed *ab initio* theory of quantum transport. It is not intended to be a general overview of the field. For further information we refer the interested reader to [1–3]. The nomenclature *quantum transport* has been coined for the phenomenon of electron motion through constrictions of transverse dimensions smaller than the electron wavelength, e.g., quantum-point contacts, quantum wires, molecules, etc. The typical experimental setup is displayed in Figure 1 where a central region C of meso- or nano-scopic size is coupled to two metallic electrodes L and R which play the role of charge reservoirs. The whole system is initially (at time $t < 0$) in a well-defined equilibrium configuration, described by a *unique* temperature and chemical potential (thermodynamic consistency). The charge density of the electrodes is perfectly balanced and no current flows through the junction.

As originally proposed by Cini [4], we may drive the system out of equilibrium by exposing the electrons to an external time-dependent potential which is local in time and space. For instance, we may switch on an electric field by putting the system between two capacitor plates far away from the system boundaries. The dynamical formation of dipole layers screens the potential drop along the electrodes and the total potential turns out to be uniform in the left and right bulks. Accordingly, the potential drop is entirely limited to the central region. As the system size increases, the remote parts are less disturbed by the junction, and the density inside the electrodes approaches the equilibrium bulk density.

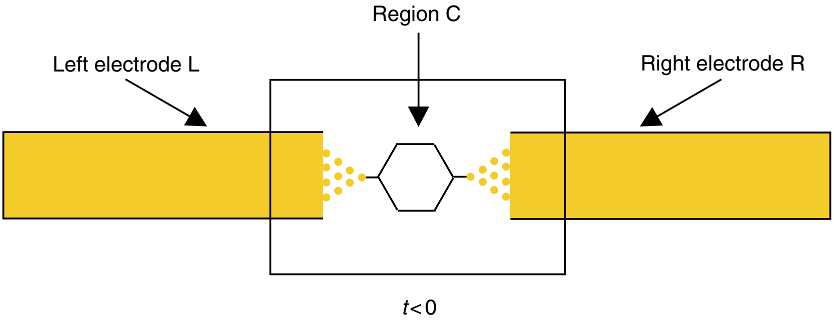


Figure 1 Schematic sketch of the experimental setup described in the text. A central region which also includes few layers of the left and right electrodes is coupled to macroscopically large metallic reservoirs. The system is in equilibrium for negative times

The Cini scheme can be combined with Time-Dependent Density Functional Theory (TDDFT) [5]. In this theory, the time-dependent density of an interacting system moving in an external, time-dependent local potential can be calculated via a fictitious system of noninteracting electrons moving in a local, effective, time-dependent potential. Therefore this theory is in principle well suited for the treatment of nonequilibrium transport problems [6]. However, as far as the leads are treated as *noninteracting*, it is not obvious that in the long time limit a steady-state current can ever develop. The reason behind the uncertainty is that the bias represents a large perturbation and, in the absence of dissipative effects, e.g., electron–electron or electron–phonon scattering, the return of time-translational invariance is not granted. In this chapter we will show that the total current tends to a steady-state value provided the effective potential of TDDFT is independent of time and space in the left and right bulks. Also, the physical mechanism leading to the dynamical formation of a steady state is clarified.

It should be mentioned that there has already been considerable activity in the density functional theory (DFT) community to describe transport phenomena through systems like the one in Figure 1. Most approaches are limited to the steady-state regime and are based on a self-consistency procedure first proposed by Lang [7]. In this steady-state approach based on DFT, exchange and correlation is approximated by the static Kohn–Sham (KS) potential and the charge density is obtained self-consistently in the presence of the steady current. However, the original justification involved subtle points such as different Fermi levels deep inside the left and right electrodes (which is not thermodynamically consistent) and the implicit reference of nonlocal perturbations such as tunneling Hamiltonians within a DFT framework. (For a detailed discussion we refer the reader to [8].) Furthermore, the transmission functions computed from static DFT have resonances at the noninteracting KS excitation energies which in general do not coincide with the true excitation energies.

Our TDDFT formulation, as opposed to the static DFT formulation, is thermodynamically consistent, is not limited to the steady-state regime (we can study transients, AC responses, etc.) and has the extra merit of accessing the true excitation energies of interacting systems [9].

We will first use the nonequilibrium Green's function (NEGF) technique to discuss the implications of our approach. For those readers that are not familiar with the Keldysh formalism and with NEGF, in Section 2 we give an elementary introduction to the Keldysh contour, the Keldysh–Green functions and the Keldysh book-keeping. The aim of this section is to derive some of the identities needed for the discussion (thus providing a self-contained presentation) and to establish the basic notation. In Section 3 we set up the theoretical framework by combining TDDFT and NEGF. An exact expression for the time-dependent total current $I(t)$ is written in terms of Green's functions projected in region C. It is also shown that a steady-state regime develops provided: (1) the KS Hamiltonian *globally* converges to an asymptotic KS Hamiltonian when $t \to \infty$, (2) the electrodes form a continuum of states (thermodynamic limit), and (3) the local density of states is a smooth function in the central region. It is worth noting that the steady-state current results from a pure dephasing mechanism in the fictitious KS system. Also, the resulting steady current only depends on the KS potential at $t = \infty$ and not on its history. However, the KS potential might depend on the history of the external applied potential and the resulting steady-state current might be history dependent. A practical scheme to calculate $I(t)$ is presented in Section 4. The main idea is to propagate the KS orbitals in region C only, without dealing with the infinite and non-periodic system [10]. We first show how to obtain the KS eigenstates ψ_s of the undisturbed system in Section 4.1. Then, in Section 4.2 we describe an algorithm for propagating ψ_s under the influence of a time-dependent disturbance. The numerical approach of Section 4 is completely general and can be applied to any system having the geometry sketched in Figure 1. In order to demonstrate the feasibility of the scheme we implement it for one-dimensional model systems in Section 5. Here we study the dynamical current response of several systems perturbed by DC and AC biases. We verify that for noninteracting electrons the steady-state current does not depend on the history of the applied bias. Also, we present preliminary results on net currents in unbiased systems as obtained by pumping mechanisms. We summarize our findings and draw our conclusions in Section 6.

2. The Keldysh formalism

2.1. The Keldysh contour

In quantum mechanics we associate to any observable quantity O a hermitian operator $\hat{O}$. The expectation $\langle\Psi|\hat{O}|\Psi\rangle$ gives the value of O when the system is described by the state $|\Psi\rangle$. For an isolated system the Hamiltonian $\hat{\mathrm{H}}_0$ does not depend on time, and the expectation value of *any* observable quantity is constant provided $|\Psi\rangle$ is an eigenstate of $\hat{\mathrm{H}}_0$. In this section we discuss how to describe systems which are not isolated but perturbed by external fields. Without loss of generality, we assume that the system is isolated for negative times t and that $\hat{\mathrm{H}}(t<0) = \hat{\mathrm{H}}_0$. The evolution of the state $|\Psi\rangle$ is governed by the Schrödinger equation $i\frac{\mathrm{d}}{\mathrm{d}t}|\Psi(t)\rangle = \hat{\mathrm{H}}(t)|\Psi(t)\rangle$, and, correspondingly, the value of O evolves in time as $O(t) = \langle\Psi(t)|\hat{O}|\Psi(t)\rangle$. The time-evolved state $|\Psi(t)\rangle = \hat{S}(t;0)|\Psi(0)\rangle$, where the evolution operator $\hat{S}(t;t')$ can be formally written as

$$\hat{S}(t;t') = \begin{cases} T\,\mathrm{e}^{-\mathrm{i}\int_{t'}^{t}\mathrm{d}\bar{t}\,\hat{\mathrm{H}}(\bar{t})} & t > t' \\ \overline{T}\,\mathrm{e}^{-\mathrm{i}\int_{t'}^{t}\mathrm{d}\bar{t}\,\hat{\mathrm{H}}(\bar{t})} & t < t' \end{cases}. \tag{1}$$

In Eq. (1), T is the time-ordering operator and rearranges the operators in chronological order with later times to the left; $\overline{T}$ is the anti-chronological time-ordering operator.

The evolution operator is unitary and satisfies the group property $\hat{S}(t;t_1)\hat{S}(t_1;t') = \hat{S}(t;t')$ for any t_1. It follows that $O(t)$ is the average on the initial state $|\Psi(0)\rangle$ of the operator $\hat{O}$ in the Heisenberg representation, $\hat{O}_{\rm H}(t) = \hat{S}(0;t)\hat{O}\hat{S}(t;0)$, i.e.,

$$\begin{aligned} O(t) &= \langle\Psi(0)|\hat{S}(0;t)\hat{O}\hat{S}(t;0)|\Psi(0)\rangle \\ &= \langle\Psi(0)|\overline{T}\mathrm{e}^{-\mathrm{i}\int_t^0 \mathrm{d}\bar{t}\,\hat{\mathrm{H}}(\bar{t})}\,\hat{O}\,T\mathrm{e}^{-\mathrm{i}\int_0^t \mathrm{d}\bar{t}\,\hat{\mathrm{H}}(\bar{t})}|\Psi(0)\rangle. \end{aligned} \tag{2}$$

We can now design an oriented contour γ with a forward branch going from $t = 0$ to t and a backward branch coming back from t and ending in $t = 0$, see Figure 2a. Denoting with $\bar{z}$ the variable running on γ, Eq. (2) can be formally recast as follows

$$O(t) = \langle\Psi(0)|T_{\rm K}\left\{\mathrm{e}^{-\mathrm{i}\int_\gamma \mathrm{d}\bar{z}\,\hat{\mathrm{H}}(\bar{z})}\,\hat{O}(t)\right\}|\Psi(0)\rangle. \tag{3}$$

The contour ordering operator $T_{\rm K}$ moves the operators with "later" contour variable to the left. A point z is later than a point z' if z' is closer to the starting point than z, see Figure 2a. In Eq. (3), $\hat{O}(t)$ is *not* the operator in the Heisenberg representation [the latter is denoted with $\hat{O}_{\rm H}(t)$]. Actually, $\hat{O}(t) = \hat{O}$ for any t. The reason for the time argument stems from the need for specifying the position of the operator $\hat{O}$ in the contour ordering.

Let us now extend the contour γ up to infinity, as shown in Figure 2b. For any physical time t there are two points $z = t_+$ and $z = t_-$ on γ; t_- lies on the forward branch while t_+ lies on the backward branch and it is later than t_- according to the orientation chosen. We have $T_{\rm K}\{\mathrm{e}^{-\mathrm{i}\int_\gamma \mathrm{d}\bar{z}\,\hat{\mathrm{H}}(\bar{z})}\,\hat{O}(t_-)\} = \hat{S}(0;\infty)\hat{S}(\infty;t)\hat{O}(t)\hat{S}(t;0) = \hat{S}(0;t)\hat{O}\hat{S}(t;0)$, and similarly $T_{\rm K}\{\mathrm{e}^{-\mathrm{i}\int_\gamma \mathrm{d}\bar{z}\,\hat{\mathrm{H}}(\bar{z})}\,\hat{O}(t_+)\} = \hat{S}(0;t)\hat{O}(t)\hat{S}(t;\infty)\hat{S}(\infty;0) = \hat{S}(0;t)\hat{O}\hat{S}(t;0)$. Thus, the expectation value $O(t)$ in Eq. (3) is also given by the formula

$$O(z) = \langle\Psi(0)|T_{\rm K}\left\{\mathrm{e}^{-\mathrm{i}\int_\gamma \mathrm{d}\bar{z}\,\hat{\mathrm{H}}(\bar{z})}\,\hat{O}(z)\right\}|\Psi(0)\rangle. \tag{4}$$

where γ is the contour in Figure 2b; γ is called the *Keldysh contour* [11, 12]. In Eq. (4) the variable z can be either t_- or t_+ and $O(t_-) = O(t_+) = O(t)$.

The Keldysh contour can be further extended to account for statistical averages [13]. In statistical physics a system is described by the density matrix $\hat{\rho} = \sum_n w_n|\Psi_n\rangle\langle\Psi_n|$,

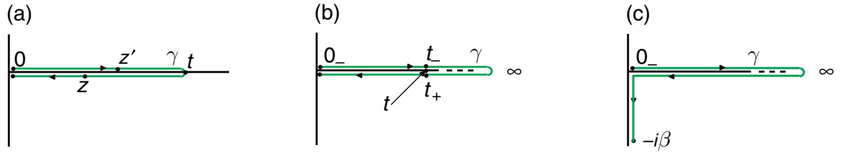


Figure 2 (a) The oriented contour γ described in the main text with a forward and a backward branch between 0 and t. According to the orientation the point z is later than the point z'. (b) The extended oriented contour γ described in the main text with a forward and a backward branch between 0 and ∞. For any physical time t we have two points $t_\pm$ on γ at the same distance from the origin. (c) The generalization of the original Keldysh contour. A vertical track going from 0 to $-i\beta$ has been added and, according to the orientation chosen, any point lying on it is later than a point lying on the forward or backward branch

with w_n being the probability of finding the system in the state $|\Psi_n\rangle$ and $\sum_n w_n = 1$. The states $|\Psi_n\rangle$ may not be orthogonal. We say that the system is in a pure state if $\hat{\rho} = |\Psi\rangle\langle\Psi|$. In a system described by a density matrix $\hat{\rho}(0)$ at $t = 0$, the time-dependent value of the observable O is a generalization of Eq. (4), i.e., $O(z) = \sum_n w_n \langle\Psi_n(0)|T_{\mathrm{K}}\{\mathrm{e}^{-\mathrm{i}\int_\gamma \mathrm{d}\bar{z}\,\hat{\mathrm{H}}(\bar{z})}\,\hat{O}(z)\}|\Psi_n(0)\rangle$. Among all possible density matrices there is one that is very common in physics and describes a system in thermal equilibrium: $\hat{\rho} = \exp[-\beta(\hat{\mathrm{H}}_0 - \mu\hat{\mathrm{N}})]/\mathcal{Z}$, with the inverse temperature β, the chemical potential μ, the operator $\hat{\mathrm{N}}$ corresponding to the total number of particles and the grand-partition function $\mathcal{Z} = \mathrm{Tr}\exp[-\beta(\hat{\mathrm{H}}_0 - \mu\hat{\mathrm{N}})]$. Assuming that $\hat{\mathrm{H}}_0$ and $\hat{\mathrm{N}}$ commute, the statistical average $O(z)$ with the thermal density matrix can be written as $O(z) = \mathrm{Tr}\,[\,\mathrm{e}^{\beta\mu\hat{\mathrm{N}}}\mathrm{e}^{-\beta\hat{\mathrm{H}}_0}T_{\mathrm{K}}\{\mathrm{e}^{-\mathrm{i}\int_\gamma \mathrm{d}\bar{z}\,\hat{\mathrm{H}}(\bar{z})}\,\hat{O}(z)\}\,]/\mathcal{Z}$. We can now extend further the Keldysh contour as shown in Figure 2c and define $\hat{\mathrm{H}}(z) = \hat{\mathrm{H}}_0$ for any z on the vertical track. With this definition $\hat{\mathrm{H}}(z)$ is continuous along the entire contour since $\hat{\mathrm{H}}(0) = \hat{\mathrm{H}}_0$. According to the orientation displayed in the figure, any point lying on the vertical track is later than a point lying on the forward or backward branch. We use this observation to rewrite $O(z)$ as

$$O(z) = \frac{\mathrm{Tr}\left[\mathrm{e}^{\beta\mu\hat{\mathrm{N}}}T_{\mathrm{K}}\left\{\mathrm{e}^{-\mathrm{i}\int_\gamma \mathrm{d}\bar{z}\,\hat{\mathrm{H}}(\bar{z})}\,\hat{O}(z)\right\}\right]}{\mathrm{Tr}\left[\mathrm{e}^{\beta\mu\hat{\mathrm{N}}}T_{\mathrm{K}}\left\{\mathrm{e}^{-\mathrm{i}\int_\gamma \mathrm{d}\bar{z}\,\hat{\mathrm{H}}(\bar{z})}\right\}\right]}, \tag{5}$$

where T_{K} is now the ordering operator on the extended contour. It is worth noting that the denominator in the above expression is simply $\mathcal{Z}$. We have already shown that choosing z on one of the two horizontal branches, Eq. (5) yields the time-dependent statistical average of the observable O. On the other hand, if z lies on the vertical track $O(z) = \mathrm{Tr}\,[\,\mathrm{e}^{\beta\mu\hat{\mathrm{N}}}\mathrm{e}^{-\mathrm{i}\int_z^{-\mathrm{i}\beta}\hat{\mathrm{H}}_0}\,\hat{O}\mathrm{e}^{-\mathrm{i}\int_0^z\hat{\mathrm{H}}_0}\,]/\mathcal{Z} = \mathrm{Tr}\,[\,\mathrm{e}^{-\beta(\hat{\mathrm{H}}_0-\mu\hat{\mathrm{N}})}\hat{O}]/\mathcal{Z}$, where the cyclic property of the trace has been used. The result is independent of z and coincides with the thermal average of the observable O.

To summarize, in Eq. (5) the variable z lies on the contour of Figure 2c; the r.h.s. gives the time-dependent statistical average of the observable quantity O when z lies on the forward or backward branch, and the statistical average before the system is disturbed when z lies on the vertical track.

2.2. *The Keldysh–Green function*

The idea presented in the previous section can be used to define correlators of many operators on the extended Keldysh contour. The Keldysh–Green function $\boldsymbol{G}$ is the correlator of two field operators $\psi(\boldsymbol{r})$ and $\psi^\dagger(\boldsymbol{r})$ which obey the anticommutation relations $\{\psi(\boldsymbol{r}), \psi^\dagger(\boldsymbol{r}')\} = \delta(\boldsymbol{r} - \boldsymbol{r}')$. It is defined as

$$\langle \boldsymbol{r}|\boldsymbol{G}(z;z')|\boldsymbol{r}'\rangle = \frac{1}{\mathrm{i}}\,\frac{\mathrm{Tr}\left[\mathrm{e}^{\beta\mu\hat{\mathrm{N}}}T_{\mathrm{K}}\left\{\mathrm{e}^{-\mathrm{i}\int_\gamma \mathrm{d}\bar{z}\,\hat{\mathrm{H}}(\bar{z})}\,\psi(\boldsymbol{r},z)\psi^\dagger(\boldsymbol{r}',z')\right\}\right]}{\mathrm{Tr}\left[\mathrm{e}^{\beta\mu\hat{\mathrm{N}}}T_{\mathrm{K}}\left\{\mathrm{e}^{-\mathrm{i}\int_\gamma \mathrm{d}\bar{z}\,\hat{\mathrm{H}}(\bar{z})}\right\}\right]}, \tag{6}$$

where the contour variable in the field operators specifies the position in the contour ordering (there is no true dependence on z in ψ and $\psi^\dagger$). Here and in the following we use boldface to indicate matrices in one-electron labels, e.g., $\boldsymbol{G}$ is a matrix and $\langle \boldsymbol{r}|\boldsymbol{G}|\boldsymbol{r}'\rangle$ is the $(\boldsymbol{r}, \boldsymbol{r}')$ matrix element of $\boldsymbol{G}$. Due to the contour ordering operator T_{K}, the Green function $\boldsymbol{G}$ has the following structure

$$\boldsymbol{G}(z; z') = \theta(z, z')\boldsymbol{G}^{>}(z; z') + \theta(z', z)\boldsymbol{G}^{<}(z; z'), \tag{7}$$

where $\theta(z, z') = 1$ if z is later than z' on the contour and zero otherwise. We say that a two-point function on the contour having the above structure belongs to the *Keldysh space*. The Green function $\boldsymbol{G}(z; z')$ obeys an important cyclic relation on the extended Keldysh contour. As we shall see, the relations below play a crucial role since they provide the boundary conditions for solving the Dyson equation. Choosing $z = 0_{-}$ we find

$$\langle \boldsymbol{r}|\boldsymbol{G}(0_{-}; z')|\boldsymbol{r}'\rangle = -\frac{1}{\mathrm{i}} \frac{\mathrm{Tr}\left[\mathrm{e}^{\beta\mu\hat{\mathrm{N}}} T_{\mathrm{K}}\left\{\mathrm{e}^{-\mathrm{i}\int_{\gamma}\mathrm{d}\bar{z}\,\hat{\mathrm{H}}(\bar{z})}\psi^{\dagger}(\boldsymbol{r}', z')\right\}\psi(\boldsymbol{r})\right]}{\mathrm{Tr}\left[\mathrm{e}^{\beta\mu\hat{\mathrm{N}}} T_{\mathrm{K}}\left\{\mathrm{e}^{-\mathrm{i}\int_{\gamma}\mathrm{d}\bar{z}\,\hat{\mathrm{H}}(\bar{z})}\right\}\right]}, \tag{8}$$

where we have taken into account that 0_{-} is the earliest time and therefore $\psi(\boldsymbol{r}, 0_{-})$ is always moved to the right when acted upon by T_{K}. The extra minus sign in the r.h.s. comes from the contour ordering. More generally, rearranging the field operators ψ and $\psi^{\dagger}$ (later arguments to the left), we also have to multiply by $(-1)^{P}$, where P is the parity of the permutation. Inside the trace we can move $\psi(\boldsymbol{r})$ to the left. Furthermore, we can exchange the position of $\psi(\boldsymbol{r})$ and $\mathrm{e}^{\beta\mu\hat{\mathrm{N}}}$ by noting that $\psi(\boldsymbol{r})\mathrm{e}^{\beta\mu\hat{\mathrm{N}}} = \mathrm{e}^{\beta\mu(\hat{\mathrm{N}}+1)}\psi(\boldsymbol{r})$. Using the fact that T_{K} moves operators with later times to the left we have $\psi(\boldsymbol{r})T_{\mathrm{K}}\{\ldots\} = T_{\mathrm{K}}\{\psi(\boldsymbol{r}, -\mathrm{i}\beta)\ldots\}$. Therefore, we conclude that

$$\boldsymbol{G}(0_{-}; z') = -\mathrm{e}^{\beta\mu}\boldsymbol{G}(-\mathrm{i}\beta; z'), \qquad \boldsymbol{G}(z; 0_{-}) = -\mathrm{e}^{-\beta\mu}\boldsymbol{G}(z; -\mathrm{i}\beta), \tag{9}$$

where the second of these relations can be obtained in a similar way. The conditions in Eq. (9) are the so-called Kubo–Martin–Schwinger (KMS) boundary conditions [14, 15].

2.3. The Keldysh book-keeping

In this section we derive some of the identities that we will use for dealing with time-dependent transport phenomena. A systematic and more exhaustive discussion can be found in [16].

Let $k(z; z')$ belong to the Keldysh space: $k(z; z') = \theta(z, z')k^{>}(z; z') + \theta(z', z)k^{<}(z; z')$. For any $k(z; z')$ in the Keldysh space we define the *greater* and *lesser* functions on the physical time axis

$$k^{>}(t; t') \equiv k(t_{+}; t'_{-}), \qquad k^{<}(t, t') \equiv k(t_{-}; t'_{+}). \tag{10}$$

We also define the *left* and *right* functions with one argument t on the physical time axis and the other τ on the vertical track

$$k^{\rceil}(t; \tau) \equiv k(t_{\pm}; \tau), \qquad k^{\lceil}(\tau, t) \equiv k(\tau; t_{\pm}). \tag{11}$$

In the definitions of $k^{\rceil}$ and $k^{\lceil}$ we can arbitrarily choose t_+ or t_- since τ is later than both of them. The symbols “$\rceil$” and “$\lceil$” have been chosen in order to help the visualization of the time arguments. For instance, “$\rceil$” has a horizontal segment followed by a vertical one; correspondingly, $k^{\rceil}$ has a first argument which is real (and thus lies on the horizontal axis) and a second argument which is imaginary (and thus lies on the vertical axis). We will also use the convention of denoting with Latin letters the real time and with Greek letters the imaginary time.

It is straightforward to show that if $a(z;z')$ and $b(z;z')$ belong to the Keldysh space, then $c(z;z') = \int_\gamma \mathrm{d}\bar{z}\, a(z;\bar{z})b(\bar{z};z')$ also belongs to the Keldysh space. From the definitions (10–11) we find

$$\begin{aligned} c^>(t;t') &= \int_{0_-}^{t'_-} \mathrm{d}\bar{z}\, a^>(t_+;\bar{z})b^<(\bar{z};t'_-) + \int_{t'_-}^{t_+} \mathrm{d}\bar{z}\, a^>(t_+;\bar{z})b^>(\bar{z};t'_-) \\ &\quad + \int_{t_+}^{-i\beta} \mathrm{d}\bar{z}\, a^<(t_+;\bar{z})b^>(\bar{z};t'_-) \\ &= \int_0^{t'} \mathrm{d}\bar{t}\, a^>(t;\bar{t})b^<(\bar{t};t') + \int_{t'}^{t} \mathrm{d}\bar{t}\, a^>(t;\bar{t})b^>(\bar{t};t') \\ &\quad + \int_t^0 \mathrm{d}\bar{t}\, a^<(t;\bar{t})b^>(\bar{t};t') + \int_0^{-i\beta} \mathrm{d}\bar{\tau}\, a^{\rceil}(t;\bar{\tau})b^{\lceil}(\bar{\tau};t'). \end{aligned} \tag{12}$$

The second integral on the r.h.s. is an ordinary integral on the real axis of two well-defined functions and may be rewritten as $\int_{t'}^{t} \mathrm{d}\bar{t}\; a^>(t;\bar{t})b^>(\bar{t};t') = \int_{t'}^{0} \mathrm{d}\bar{t}\; a^>(t;\bar{t})b^>(\bar{t};t') + \int_0^t \mathrm{d}\bar{t}\; a^>(t;\bar{t})b^>(\bar{t};t')$. Using this relation, Eq. (12) becomes

$$c^>(t;t') = \int_0^\infty \mathrm{d}\bar{t}\,[a^>(t;\bar{t})b^{\mathrm{A}}(\bar{t};t') + a^{\mathrm{R}}(t;\bar{t})b^>(\bar{t};t')] + \int_0^{-\mathrm{i}\beta} \mathrm{d}\bar{\tau}\, a^{\rceil}(t;\bar{\tau})b^{\lceil}(\bar{\tau};t'), \tag{13}$$

where we have introduced two other functions on the physical time axis

$$\begin{aligned} k^{\mathrm{R}}(t;t') &\equiv \theta(t-t')[k^>(t;t') - k^<(t;t')], \\ k^{\mathrm{A}}(t;t') &\equiv -\theta(t'-t)[k^>(t;t') - k^<(t;t')]. \end{aligned} \tag{14}$$

The *retarded* function $k^{\mathrm{R}}(t;t')$ vanishes for $t < t'$, while the *advanced* function $k^{\mathrm{A}}(t;t')$ vanishes for $t > t'$. A relation similar to Eq. (13) can be obtained for the lesser component $c^<$. It is convenient to introduce a shorthand notation for integrals along the physical time axis and for those between 0 and $-\mathrm{i}\beta$. The symbol “$\cdot$” will be used to write $\int_0^\infty \mathrm{d}\bar{t} f(\bar{t})g(\bar{t})$ as $f \cdot g$, while the symbol “$*$” will be used to write $\int_0^{-\mathrm{i}\beta} \mathrm{d}\bar{\tau} f(\bar{\tau})g(\bar{\tau})$ as $f * g$. Then

$$c^> = a^> \cdot b^{\mathrm{A}} + a^{\mathrm{R}} \cdot b^> + a^{\rceil} * b^{\lceil}, \qquad c^< = a^< \cdot b^{\mathrm{A}} + a^{\mathrm{R}} \cdot b^< + a^{\rceil} * b^{\lceil}. \tag{15}$$

Equation (15) can be used to extract the retarded and advanced component of c. By definition $c^{\mathrm{R}}(t;t') = \theta(t-t')[c^>(t;t') - c^<(t;t')]$ therefore

$$\begin{aligned} c^{\mathrm{R}}(t;t') &= \theta(t-t') \textstyle\int_0^\infty \mathrm{d}\bar{t}\, a^{\mathrm{R}}(t;\bar{t})[b^>(\bar{t};t') - b^<(\bar{t};t')] \\ &\quad + \theta(t-t') \textstyle\int_0^\infty \mathrm{d}\bar{t}\,[a^>(t;\bar{t}) - a^<(t;\bar{t})]b^{\mathrm{A}}(\bar{t};t'). \end{aligned} \tag{16}$$

Due to the θ function, we have $t > t'$ for $c^{\mathrm{R}} \neq 0$. In the second term on the r.h.s. $b^{\mathrm{A}}(\bar{t}; t')$ contains a $\theta(t' - \bar{t})$ and hence it must be $t > \bar{t}$; therefore we can replace the difference in the square bracket with a^{R}. Then we break the first term on the r.h.s. into two pieces by inserting θ functions: one for $\bar{t} < t'$ and the other for $\bar{t} > t'$. In compact notation, we end up with

$$c^{\mathrm{R}} = a^{\mathrm{R}} \cdot b^{\mathrm{R}}, \qquad c^{\mathrm{A}} = a^{\mathrm{A}} \cdot b^{\mathrm{A}}, \tag{17}$$

where the second relation can be proven in a similar way. It is worth noting that in the expressions for c^{R} and c^{A} no integration along the imaginary track is required. For later purposes we also define the *Matsubara* function $k^{\mathrm{M}}(\tau; \tau')$ with both the arguments in the interval $(0, -\mathrm{i}\beta)$:

$$k^{\mathrm{M}}(\tau; \tau') \equiv k(z = \tau; z' = \tau'). \tag{18}$$

It is straightforward to prove the following identities

$$c^{\rceil} = a^{\mathrm{R}} \cdot b^{\rceil} + a^{\rceil} * b^{\mathrm{M}}, \qquad c^{\lceil} = a^{\lceil} \cdot b^{\mathrm{A}} + a^{\mathrm{M}} * b^{\lceil}, \qquad c^{\mathrm{M}} = a^{\mathrm{M}} * b^{\mathrm{M}}. \tag{19}$$

Finally, we consider the case of a Keldysh function $k(z; z')$ multiplied on the left by a scalar function $l(z)$. The function $k_l(z; z') = l(z)k(z; z') = \theta(z, z')l(z)k^{>}(z; z') + \theta(z'; z)l(z)k^{<}(z; z')$ and hence belongs to the Keldysh space. The Keldysh components can be extracted using the definitions (10, 11, 14, 18). Choosing for instance $z = t_+$ and $z' = t'_-$ we find $k_l^{>}(t; t') = l(t)k^{>}(t; t')$ and similarly for $z = t_-$ and $z' = t'_+$ we find $k_l^{<}(t; t') = l(t)k^{<}(t; t')$. More generally, the function l is simply a prefactor: $k_l^{\mathrm{x}} = lk^{\mathrm{x}}$, where x is one of the Keldysh components ($\lessgtr$, R, A, $\rceil$, $\lceil$, M). The same is true for $k_r(z; z') = k(z; z')r(z')$, where $r(z')$ is a scalar function: $k_r^{\mathrm{x}} = k^{\mathrm{x}}r$.

3. Quantum transport using TDDFT and NEGF

3.1. Merging the Keldysh and TDDFT formalisms

The one-particle scheme of TDDFT corresponds to a fictitious system of non-interacting electrons described by the Kohn–Sham (KS) Hamiltonian $\hat{\mathrm{H}}_s(z) = \int \mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{r}' \psi^{\dagger}(\boldsymbol{r})\langle \boldsymbol{r}|\boldsymbol{H}_s(z)|\boldsymbol{r}'\rangle\psi(\boldsymbol{r}')$. The potential $v_s(\mathbf{r}, t)$ experienced by the electrons in the free-electron Hamiltonian $\boldsymbol{H}_s(t)$ is called the KS potential and it is given by the sum of the external potential, the Coulomb potential of the nuclei, the Hartree potential and the exchange-correlation potential v_{xc}. The latter accounts for the complicated many-body effects and is obtained from an exchange-correlation action functional, $v_{\mathrm{xc}}(\mathbf{r}, t) = \delta A_{\mathrm{xc}}[n]/\delta n(\mathbf{r}, t)$ (as pointed out in [17], the causality and symmetry properties require that the action functional $A_{\mathrm{xc}}[n]$ is defined on the Keldysh contour). A_{xc} is a functional of the density and of the initial density matrix. In our case, the initial density matrix is the thermal density matrix which, due to the extension of the Hohenberg–Kohn theorem [18] to finite temperatures [19], also is a functional of the density. We should mention here that an alternative formulation based on TDDFT has been recently proposed by Di Ventra and Todorov [20]. In their approach the system is initially unbalanced and therefore the exchange-correlation functional depends on the initial state and not only on the density.

The fictitious Keldysh–Green function $\mathcal{G}(z;z')$ of the KS system satisfies a one-particle equation of motion

$$\left\{ \mathrm{i}\overrightarrow{\frac{\mathrm{d}}{\mathrm{d}z}}\mathbf{1} - \boldsymbol{H}_s(z) \right\} \mathcal{G}(z;z') = \mathbf{1}\delta(z-z'),$$

$$\mathcal{G}(z;z') \left\{ -\mathrm{i}\overleftarrow{\frac{\mathrm{d}}{\mathrm{d}z'}}\mathbf{1} - \boldsymbol{H}_s(z') \right\} = \mathbf{1}\delta(z-z'), \tag{20}$$

with KMS boundary conditions (9). In Eqs (20) the arrow specifies where the derivative along the contour acts. The left and right equations of motion are equations on the extended Keldysh contour of Figure 2c and $\delta(z-z') = \frac{\mathrm{d}}{\mathrm{d}z}\theta(z,z') = -\frac{\mathrm{d}}{\mathrm{d}z'}\theta(z,z')$. For any $z \neq z'$, the equations of motion are solved by the evolution operator on the contour $\boldsymbol{S}(z;z') = T_{\mathrm{K}}\{\mathrm{e}^{-\mathrm{i}\int_{z'}^{z}\mathrm{d}\bar{z}\boldsymbol{H}_s(\bar{z})}\}$, since $\mathrm{i}\overrightarrow{\frac{\mathrm{d}}{\mathrm{d}z}}\boldsymbol{S}(z;z') = \boldsymbol{H}_s(z)\boldsymbol{S}(z;z')$ and $\boldsymbol{S}(z;z')(-\mathrm{i}\overleftarrow{\frac{\mathrm{d}}{\mathrm{d}z'}}) = \boldsymbol{S}(z;z')\boldsymbol{H}_s(z')$. Therefore, any Green function

$$\mathcal{G}(z;z') = \theta(z,z')\boldsymbol{S}(z;0_-)\boldsymbol{f}^{>}\boldsymbol{S}(0_-;z') + \theta(z',z)\boldsymbol{S}(z;0_-)\boldsymbol{f}^{<}\boldsymbol{S}(0_-;z'), \tag{21}$$

with $\boldsymbol{f}^{\lessgtr}$ constrained by $\boldsymbol{f}^{>} - \boldsymbol{f}^{<} = -\mathrm{i}\mathbf{1}$, is a solution of Eqs (20). In order to fix the matrix $\boldsymbol{f}^{>}$ or $\boldsymbol{f}^{<}$ we impose the KMS boundary conditions. The matrix $\boldsymbol{H}_s(z)$ equals $\boldsymbol{H}_s$ for any z on the vertical track, meaning that $\boldsymbol{S}(-\mathrm{i}\beta;0_-) = \mathrm{e}^{-\beta\boldsymbol{H}_s}$. Equations (9) then implies $\boldsymbol{f}^{<} = -\mathrm{e}^{-\beta(\boldsymbol{H}_s-\mu)}\boldsymbol{f}^{>}$, and taking into account the constraint $\boldsymbol{f}^{>} - \boldsymbol{f}^{<} = -\mathrm{i}\mathbf{1}$ we conclude that $\boldsymbol{f}^{<} = \mathrm{i}f(\boldsymbol{H}_s)$, where $f(\omega) = 1/[\mathrm{e}^{\beta(\omega-\mu)}+1]$ is the Fermi distribution function. The matrix $\boldsymbol{f}^{>}$ takes the form $\boldsymbol{f}^{>} = \mathrm{i}[f(\boldsymbol{H}_s) - \mathbf{1}]$.

The Green function $\mathcal{G}(z;z')$ for a system of noninteracting electrons is now completely fixed. Let us consider some Keldysh–Green functions. For $z = t_+$ and $z' = t_-$ we have the *greater* Green function while for $z = t_-$ and $z' = t_+$ we have the *lesser* Green function

$$\mathcal{G}^{>}(t;t') = \mathrm{i}\,\boldsymbol{S}(t;0)[f(\boldsymbol{H}_s) - \mathbf{1}]\boldsymbol{S}(0;t'), \quad \mathcal{G}^{<}(t;t') = \mathrm{i}\,\boldsymbol{S}(t;0)f(\boldsymbol{H}_s)\boldsymbol{S}(0;t'). \tag{22}$$

Both $\mathcal{G}^{>}$ and $\mathcal{G}^{<}$ depend on the initial distribution function $f(\boldsymbol{H}_s)$. The diagonal matrix element of $-\mathrm{i}\mathcal{G}^{<}$ is nothing but the time-dependent value of the local electron density $n(\boldsymbol{r},t)$, while $\mathrm{i}\mathcal{G}^{>}$ gives the local density of holes. Another way of writing $-\mathrm{i}\mathcal{G}^{<}$ is in terms of the eigenstates $|\psi_s(0)\rangle$ of $\boldsymbol{H}_s$ with eigenvalues ε_s. From the time-evolved eigenstate $|\psi_s(t)\rangle = \boldsymbol{S}(t;0)|\psi_s(0)\rangle$, we can calculate the time-dependent wavefunction $\psi_s(\boldsymbol{r},t) = \langle\boldsymbol{r}|\psi_s(t)\rangle$. Inserting $\sum_s |\psi_s(0)\rangle\langle\psi_s(0)|$ in the expression for $\mathcal{G}^{<}$ we find $-\mathrm{i}\langle\boldsymbol{r}|\mathcal{G}^{<}(t;t')|\boldsymbol{r}'\rangle = \sum_s f(\varepsilon_s)\psi_s(\boldsymbol{r},t)\psi_s^*(\boldsymbol{r}',t')$, which for $t = t'$ reduces to the time-dependent density matrix. Knowing the greater and lesser Green functions we can also calculate $\mathcal{G}^{\mathrm{R,A}}$. Taking into account the definitions (14) we find

$$\mathcal{G}^{\mathrm{R}}(t;t') = -\mathrm{i}\theta(t-t')\boldsymbol{S}(t;t'), \quad \mathcal{G}^{\mathrm{A}}(t;t') = \mathrm{i}\theta(t'-t)\boldsymbol{S}(t;t') = [\mathcal{G}^{\mathrm{R}}(t';t)]^{\dagger}. \tag{23}$$

In the above expressions for $\mathcal{G}^{\mathrm{R,A}}$, the Fermi distribution function has disappeared. The information carried by $\mathcal{G}^{\mathrm{R,A}}$ is the same as contained in the one-particle evolution operator. There is no information on how the system is prepared (how many particles, how they are distributed, etc.). We use this observation to rewrite $\mathcal{G}^{\lessgtr}$ in terms of $\mathcal{G}^{\mathrm{R,A}}$

$$\mathcal{G}^{\lessgtr}(t;t') = \mathcal{G}^{\mathrm{R}}(t;0)\mathcal{G}^{\lessgtr}(0;0)\mathcal{G}^{\mathrm{R}}(0;t'). \tag{24}$$

Thus, $\mathcal{G}^{\lessgtr}$ is completely known once we know how to propagate the one-electron orbitals in time and how they are populated before the system is perturbed [4, 21]. For later purposes, we also observe that an analogous relation holds for $\mathcal{G}^{\rceil,\lceil}$

$$\mathcal{G}^{\rceil}(t;\tau) = \mathrm{i}\mathcal{G}^{\mathrm{R}}(t;0)\mathcal{G}^{\rceil}(0;\tau), \quad \mathcal{G}^{\lceil}(\tau;t) = -\mathrm{i}\mathcal{G}^{\lceil}(\tau;0)\mathcal{G}^{\mathrm{A}}(0;t). \tag{25}$$

3.2. Total current using TDDFT

The fictitious $\mathcal{G}$ of the KS system will not, in general, give correct one-particle properties. However, by definition, $\mathcal{G}^{<}$ gives the correct density $n(\boldsymbol{r},t) = -\mathrm{i}\langle\boldsymbol{r}|\mathcal{G}^{<}(t;t)|\boldsymbol{r}\rangle$. Also total currents are correctly given by TDDFT. If, for instance, I_α is the total current from a particular region α we have

$$I_\alpha(t) = -e\int_\alpha \mathrm{d}\boldsymbol{r}\,\frac{\mathrm{d}}{\mathrm{d}t}n(\boldsymbol{r},t) = e\int_\alpha \mathrm{d}\boldsymbol{r}\,\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}\langle\boldsymbol{r}|\mathcal{G}^{<}(t;t)|\boldsymbol{r}\rangle \tag{26}$$

where the space integral extends over the region α (e is the electron charge). We stress here that I_α is the electronic current (the direction of the current coincides with the direction of the electrons).

At this point, it is convenient to partition the system into three main regions: a central region C consisting of the junction and a few atomic layers of the left and right electrodes and two regions L and R which describe the left and right bulk electrodes. According to this partitioning, the KS Hamiltonian $\boldsymbol{H}_s$ can be written as a 3×3 block matrix, and the left equation of motion in (20) reads

$$\left\{\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}z}\mathbf{1} - \begin{bmatrix} \boldsymbol{H}_{\mathrm{LL}}(z) & \boldsymbol{H}_{\mathrm{LC}} & 0 \\ \boldsymbol{H}_{\mathrm{CL}} & \boldsymbol{H}_{\mathrm{CC}}(z) & \boldsymbol{H}_{\mathrm{CR}} \\ 0 & \boldsymbol{H}_{\mathrm{RC}} & \boldsymbol{H}_{\mathrm{RR}}(z) \end{bmatrix}\right\}\mathcal{G}(z;z') = \delta(z-z')\mathbf{1}, \tag{27}$$

with

$$\mathcal{G}(z;z') = \begin{bmatrix} \mathcal{G}_{\mathrm{LL}}(z;z') & \mathcal{G}_{\mathrm{LC}}(z;z') & \mathcal{G}_{\mathrm{LR}}(z;z') \\ \mathcal{G}_{\mathrm{CL}}(z;z') & \mathcal{G}_{\mathrm{CC}}(z;z') & \mathcal{G}_{\mathrm{CR}}(z;z') \\ \mathcal{G}_{\mathrm{RL}}(z;z') & \mathcal{G}_{\mathrm{RC}}(z;z') & \mathcal{G}_{\mathrm{RR}}(z;z') \end{bmatrix} \tag{28}$$

(a similar equation is easily obtained for the right equation of motion). Choosing z on the forward branch of the Keldysh contour and z' on the backward branch of the same contour, we obtain a left and a right equation for the lesser Green function. These equations can be used to get rid of the time derivative in Eq. (26). We find for $\alpha = \mathrm{L,R}$

$$\begin{aligned} I_\alpha(t) &= e\int \mathrm{d}\boldsymbol{r}\,\langle\boldsymbol{r}|\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{G}^{<}_{\alpha\alpha}(t;t)|\boldsymbol{r}\rangle \\ &= e\int \mathrm{d}\boldsymbol{r}\,\langle\boldsymbol{r}|\boldsymbol{H}_{\alpha\mathrm{C}}\mathcal{G}^{<}_{\mathrm{C}\alpha}(t;t) - \mathcal{G}^{<}_{\alpha\mathrm{C}}(t;t)\boldsymbol{H}_{\mathrm{C}\alpha}|\boldsymbol{r}\rangle = 2e\,\mathrm{Re}\left[\mathrm{Tr}_{\mathrm{C}}\left\{\boldsymbol{Q}_\alpha(t)\right\}\right], \end{aligned} \tag{29}$$

where

$$\begin{aligned}\boldsymbol{Q}_\alpha(t) &\equiv \boldsymbol{\mathcal{G}}^<_{\mathrm{C}\alpha}(t;t)\boldsymbol{H}_{\alpha\mathrm{C}} = \left[\boldsymbol{\mathcal{G}}^{\mathrm{R}}(t,0)\boldsymbol{\mathcal{G}}^<(0,0)\boldsymbol{\mathcal{G}}^{\mathrm{A}}(0,t)\right]_{\mathrm{C}\alpha}\boldsymbol{H}_{\alpha\mathrm{C}}\\ &= \boldsymbol{\mathcal{G}}^{\mathrm{R}}_{\mathrm{CC}}(t;0)\boldsymbol{\mathcal{G}}^<_{\mathrm{CC}}(0;0)\boldsymbol{\mathcal{G}}^{\mathrm{A}}_{\mathrm{C}\alpha}(0;t)\boldsymbol{H}_{\alpha\mathrm{C}}\\ &\quad+\sum_{\beta=\mathrm{L,R}}\boldsymbol{\mathcal{G}}^{\mathrm{R}}_{\mathrm{C}\beta}(t;0)\boldsymbol{\mathcal{G}}^<_{\beta\mathrm{C}}(0;0)\boldsymbol{\mathcal{G}}^{\mathrm{A}}_{\mathrm{C}\alpha}(0;t)\boldsymbol{H}_{\alpha\mathrm{C}}\\ &\quad+\sum_{\gamma=\mathrm{L,R}}\boldsymbol{\mathcal{G}}^{\mathrm{R}}_{\mathrm{CC}}(t;0)\boldsymbol{\mathcal{G}}^<_{\mathrm{C}\gamma}(0;0)\boldsymbol{\mathcal{G}}^{\mathrm{A}}_{\gamma\alpha}(0;t)\boldsymbol{H}_{\alpha\mathrm{C}}\\ &\quad+\sum_{\beta\gamma=\mathrm{L,R}}\boldsymbol{\mathcal{G}}^{\mathrm{R}}_{\mathrm{C}\beta}(t;0)\boldsymbol{\mathcal{G}}^<_{\beta\gamma}(0;0)\boldsymbol{\mathcal{G}}^{\mathrm{A}}_{\gamma\alpha}(0;t)\boldsymbol{H}_{\alpha\mathrm{C}}\end{aligned}\tag{30}$$

is a one-particle operator in the central region C and Tr_{C} denotes the trace over a complete set of one-particle states of C. Let us express the quantity $\boldsymbol{Q}_\alpha$ in terms of the Green function $\boldsymbol{\mathcal{G}}_{\mathrm{CC}}$ projected in the central region. We introduce the uncontacted Green function $\boldsymbol{g}$ which obeys Eqs (20) with $\boldsymbol{H}_{\alpha\mathrm{C}}=\boldsymbol{H}_{\mathrm{C}\alpha}=0$,

$$\left\{\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}z}\mathbf{1}-\begin{bmatrix}\boldsymbol{H}_{\mathrm{LL}}(z) & 0 & 0\\ 0 & \boldsymbol{H}_{\mathrm{CC}}(z) & 0\\ 0 & 0 & \boldsymbol{H}_{\mathrm{RR}}(z)\end{bmatrix}\right\}\boldsymbol{g}(z;z')=\delta(z-z')\mathbf{1},\tag{31}$$

where

$$\boldsymbol{g}(z;z')=\begin{bmatrix}\boldsymbol{g}_{\mathrm{LL}}(z;z') & 0 & 0\\ 0 & \boldsymbol{g}_{\mathrm{CC}}(z;z') & 0\\ 0 & 0 & \boldsymbol{g}_{\mathrm{RR}}(z;z')\end{bmatrix}\tag{32}$$

and the same KMS boundary conditions as $\boldsymbol{\mathcal{G}}$. The uncontacted $\boldsymbol{g}$ allows us to convert Eqs (20) into an integral equation which entails the KMS boundary conditions for $\boldsymbol{\mathcal{G}}$:

$$\begin{aligned}\boldsymbol{\mathcal{G}}(z;z') &= \boldsymbol{g}(z;z')+\int_\gamma \mathrm{d}\bar{z}\,\boldsymbol{g}(z;\bar{z})\boldsymbol{H}_{\mathrm{off}}\boldsymbol{\mathcal{G}}(\bar{z};z')\\ &= \boldsymbol{g}(z;z')+\int_\gamma \mathrm{d}\bar{z}\,\boldsymbol{\mathcal{G}}(z;\bar{z})\boldsymbol{H}_{\mathrm{off}}\boldsymbol{g}(\bar{z};z'),\end{aligned}\tag{33}$$

γ being the extended Keldysh contour of Figure 2c and $\boldsymbol{H}_{\mathrm{off}}$ is the off-diagonal part of $\boldsymbol{H}_s$. Using the relations (17) of Section 2.3 we find

$$\boldsymbol{\mathcal{G}}^{\mathrm{R,A}}_{\mathrm{C}\alpha}=\boldsymbol{\mathcal{G}}^{\mathrm{R,A}}_{\mathrm{CC}}\cdot\boldsymbol{H}_{\mathrm{C}\alpha}\boldsymbol{g}^{\mathrm{R,A}}_{\alpha\alpha},\quad \boldsymbol{\mathcal{G}}^{\mathrm{A}}_{\beta\alpha}=\delta_{\beta\alpha}\boldsymbol{g}^{\mathrm{A}}_{\beta\beta}+\boldsymbol{g}^{\mathrm{A}}_{\beta\beta}\boldsymbol{H}_{\beta\mathrm{C}}\cdot\boldsymbol{\mathcal{G}}^{\mathrm{A}}_{\mathrm{CC}}\cdot\boldsymbol{H}_{\mathrm{C}\alpha}\boldsymbol{g}^{\mathrm{A}}_{\alpha\alpha}.\tag{34}$$

In Eq. (30) all matrix elements of $\boldsymbol{\mathcal{G}}^<$ are evaluated at times (0; 0). From Eq. (15) we see that $c^<(0;0)=\left[a^\rceil * b^\lceil\right](0;0)$, due to the θ functions in the retarded and advanced components. Therefore

$$\boldsymbol{\mathcal{G}}^<_{\beta\mathrm{C}}(0;0)=\left[\boldsymbol{g}^\rceil_{\beta\beta}\boldsymbol{H}_{\beta\mathrm{C}}*\boldsymbol{\mathcal{G}}^\lceil_{\mathrm{CC}}\right](0;0),\quad \boldsymbol{\mathcal{G}}^<_{\mathrm{C}\gamma}(0;0)=\left[\boldsymbol{\mathcal{G}}^\rceil_{\mathrm{CC}}*\boldsymbol{H}_{\mathrm{C}\gamma}\boldsymbol{g}^\lceil_{\gamma\gamma}\right](0;0),\tag{35}$$

and exploiting the first two relations in Eq. (19) we also find that

$$\boldsymbol{\mathcal{G}}^<_{\beta\gamma}(0;0)=\delta_{\beta\gamma}\boldsymbol{g}^<_{\beta\beta}(0;0)+\left[\boldsymbol{g}^\rceil_{\beta\beta}\boldsymbol{H}_{\beta\mathrm{C}}*\boldsymbol{\mathcal{G}}^{\mathrm{M}}_{\mathrm{CC}}*\boldsymbol{H}_{\mathrm{C}\gamma}\boldsymbol{g}^\lceil_{\gamma\gamma}\right](0;0).\tag{36}$$

Substituting Eqs (34–36) into Eq. (30) and using the identities (24 and 25) for the Green function $\boldsymbol{g}$, we obtain the following expression for $\boldsymbol{Q}_\alpha(t)$:

$$\begin{aligned}\boldsymbol{Q}_\alpha(t) = & \sum_{\beta=\mathrm{L,R}} \left[\boldsymbol{G}^{\mathrm{R}} \cdot \boldsymbol{\Sigma}_\beta^< \cdot \left(\delta_{\beta\alpha} + \boldsymbol{G}^{\mathrm{A}} \cdot \boldsymbol{\Sigma}_\alpha^{\mathrm{A}}\right)\right](t;\, t) \\ & + \sum_{\beta=\mathrm{L,R}} \left[\boldsymbol{G}^{\mathrm{R}} \cdot \boldsymbol{\Sigma}^{\rceil} * \boldsymbol{G}^{\mathrm{M}} * \boldsymbol{\Sigma}_\beta^{\lceil} \cdot \left(\delta_{\beta\alpha} + \boldsymbol{G}^{\mathrm{A}} \cdot \boldsymbol{\Sigma}_\alpha^{\mathrm{A}}\right)\right](t;\, t) \\ & + \mathrm{i} \sum_{\beta=\mathrm{L,R}} \boldsymbol{G}^{\mathrm{R}}(t;\, 0) \left[\boldsymbol{G}^{\rceil} * \boldsymbol{\Sigma}_\beta^{\lceil} \cdot \left(\delta_{\beta\alpha} + \boldsymbol{G}^{\mathrm{A}} \cdot \boldsymbol{\Sigma}_\alpha^{\mathrm{A}}\right)\right](0;\, t) \\ & + \left(\boldsymbol{G}^{\mathrm{R}}(t;\, 0)\boldsymbol{G}^<(0;\, 0) - \mathrm{i}\left[\boldsymbol{G}^{\mathrm{R}} \cdot \boldsymbol{\Sigma}^{\rceil} * \boldsymbol{G}^{\lceil}\right](t;\, 0)\right)\left[\boldsymbol{G}^{\mathrm{A}} \cdot \boldsymbol{\Sigma}_\alpha^{\mathrm{A}}\right](0;\, t), \end{aligned} \tag{37}$$

where we have used the short-hand notation $\boldsymbol{G} \equiv \boldsymbol{\mathcal{G}}_{\mathrm{CC}}$ and

$$\boldsymbol{\Sigma}(z;\, z') = \sum_{\alpha=\mathrm{L,R}} \boldsymbol{\Sigma}_\alpha, \quad \boldsymbol{\Sigma}_\alpha(z;\, z') = \boldsymbol{H}_{\mathrm{C}\alpha}\, \boldsymbol{g}_{\alpha\alpha}(z;\, z')\boldsymbol{H}_{\alpha\mathrm{C}} \tag{38}$$

is the so-called embedding self-energy which accounts for hopping in and out of region C.

Having the quantity $\boldsymbol{Q}_\alpha(t)$ we can calculate the exact total current $I_\alpha(t)$ of an interacting system of electrons. Eq. (29) allows for studying transient effects and more generally any kind of time-dependent current responses. In the long time limit

$$\lim_{t\to\infty} \boldsymbol{Q}_\alpha(t) = \left[\boldsymbol{G}^{\mathrm{R}} \cdot \boldsymbol{\Sigma}_\alpha^< + \boldsymbol{G}^{\mathrm{R}} \cdot \boldsymbol{\Sigma}^< \cdot \boldsymbol{G}^{\mathrm{A}} \cdot \boldsymbol{\Sigma}_\alpha^{\mathrm{A}}\right](t;\, t) \tag{39}$$

provided $\boldsymbol{G}$ and $\boldsymbol{\Sigma}$ tend to zero when the separation between their time arguments increases (in this case, it is only the first term on the r.h.s. of Eq. (37) that does not vanish). This condition is not stringent and is fulfilled provided the electrode states form a continuum and the local density of states in the central region C is a smooth function. In the next section we investigate under what circumstances a steady current I_α develops in the long-time limit. We will also discuss the dependence of I_α on the history of the external potential.

3.3. Steady state and history dependence

In this section we show that a steady state develops provided (1) the KS Hamiltonian $\boldsymbol{H}_s(t)$ *globally* converges to an asymptotic KS Hamiltonian $\boldsymbol{H}_s^\infty$ when $t \to \infty$ and (2) the electrodes form a continuum of states (thermodynamic limit), and the local density of states is a smooth function in the central region.

Let us define the asymptotic KS Hamiltonian of electrode α as $\boldsymbol{H}_{\alpha\alpha}^\infty = \lim_{t\to\infty} \boldsymbol{H}_{\alpha\alpha}(t)$. The retarded/advanced component of the uncontacted Green function $\boldsymbol{g}$ behaves like

$$\lim_{t\to\infty} \boldsymbol{g}_{\alpha\alpha}^{\mathrm{R}}(t, 0) = \mathrm{i}\, \mathrm{e}^{-\mathrm{i}\boldsymbol{H}_{\alpha\alpha}^\infty t} \boldsymbol{\mathcal{S}}, \qquad \lim_{t\to\infty} \boldsymbol{g}_{\alpha\alpha}^{\mathrm{A}}(0, t) = -\mathrm{i}\, \boldsymbol{\mathcal{S}}^\dagger \mathrm{e}^{\mathrm{i}\boldsymbol{H}_{\alpha\alpha}^\infty t} \tag{40}$$

where $\boldsymbol{\mathcal{S}}$ is a unitary operator and it is defined according to

$$\boldsymbol{\mathcal{S}} = \lim_{t\to\infty} \frac{T\left\{\mathrm{e}^{-\mathrm{i}\int_0^t \boldsymbol{H}_{\alpha\alpha}(t')\mathrm{d}t'}\right\}}{\mathrm{e}^{-\mathrm{i}\boldsymbol{H}_{\alpha\alpha}^\infty t}}, \tag{41}$$

T being the time-ordering operator. In terms of diagonalizing one-body states $|\psi^{\infty}_{m\alpha}\rangle$ of $\boldsymbol{H}^{\infty}_{\alpha\alpha}$ with eigenvalues $\varepsilon^{\infty}_{m\alpha}$, the lesser component of the embedding self-energy, defined in Eq. (38), can be written as

$$\begin{aligned}\lim_{t,t'\to\infty}\boldsymbol{\Sigma}^{<}_{\alpha}(t;t') &= \lim_{t,t'\to\infty}\boldsymbol{H}_{\mathrm{C}\alpha}\boldsymbol{g}^{\mathrm{R}}_{\alpha\alpha}(t;0)\boldsymbol{g}^{<}_{\alpha\alpha}(0;0)\boldsymbol{g}^{\mathrm{A}}_{\alpha\alpha}(0;t')\boldsymbol{H}_{\alpha\mathrm{C}}\\ &= \mathrm{i}\sum_{m,m'}\mathrm{e}^{-\mathrm{i}[\varepsilon^{\infty}_{m\alpha}t-\varepsilon^{\infty}_{m'\alpha}t']}\\ &\quad\times\boldsymbol{H}_{\mathrm{C}\alpha}|\psi^{\infty}_{m\alpha}\rangle\langle\psi^{\infty}_{m\alpha}|f(\boldsymbol{\mathcal{S}}\boldsymbol{H}_{\alpha\alpha}(0)\boldsymbol{\mathcal{S}}^{\dagger})|\psi^{\infty}_{m'\alpha}\rangle\langle\psi^{\infty}_{m'\alpha}|\boldsymbol{H}_{\alpha\mathrm{C}},\end{aligned}\tag{42}$$

where we have taken into account that $\boldsymbol{g}^{<}_{\alpha\alpha}(0;0)=\mathrm{i}f(\boldsymbol{H}_{\alpha\alpha}(0))$. The left and right contraction with a nonsingular hopping matrix $\boldsymbol{H}_{\alpha\mathrm{C}}$ causes a perfect destructive interference for states with $|\varepsilon^{\infty}_{m\alpha}-\varepsilon^{\infty}_{m'\alpha}|\gtrsim 1/(t+t')$ and hence the restoration of translational invariance in time

$$\lim_{t,t'\to\infty}\boldsymbol{\Sigma}^{<}_{\alpha}(t;t')=\mathrm{i}\sum_{m}f_{m\alpha}\boldsymbol{\Gamma}_{m\alpha}\mathrm{e}^{-\mathrm{i}\varepsilon^{\infty}_{m\alpha}(t-t')},\tag{43}$$

where $f_{m\alpha}=\langle\psi^{\infty}_{m\alpha}|f(\boldsymbol{\mathcal{S}}\boldsymbol{H}_{\alpha\alpha}(0)\boldsymbol{\mathcal{S}}^{\dagger})|\psi^{\infty}_{m\alpha}\rangle$ while $\boldsymbol{\Gamma}_{m\alpha}=H_{\mathrm{C}\alpha}|\psi^{\infty}_{m\alpha}\rangle\langle\psi^{\infty}_{m\alpha}|\boldsymbol{H}_{\alpha\mathrm{C}}$. In principle, there may be degeneracies which require a diagonalization to be performed for states on the energy shell. The above *dephasing mechanism* is the key ingredient for a steady state to develop. Substituting Eq. (43) into Eq. (39), we obtain for the steady state current

$$\begin{aligned}I^{(\mathrm{S})}_{\alpha} &= -2e\sum_{m\beta}f_{m\beta}\mathrm{Tr}_{\mathrm{C}}\left\{\boldsymbol{G}^{\mathrm{R}}(\varepsilon^{\infty}_{m\beta})\boldsymbol{\Gamma}_{m\beta}\boldsymbol{G}^{\mathrm{A}}(\varepsilon^{\infty}_{m\beta})\mathrm{Im}[\boldsymbol{\Sigma}^{\mathrm{A}}_{\alpha}(\varepsilon^{\infty}_{m\beta})]\right\}\\ &\quad -2e\sum_{m}f_{m\alpha}\mathrm{Tr}_{\mathrm{C}}\left\{\boldsymbol{\Gamma}_{m\alpha}\mathrm{Im}[\boldsymbol{G}^{\mathrm{R}}(\varepsilon^{\infty}_{m\alpha})]\right\}\end{aligned}\tag{44}$$

with

$$\boldsymbol{G}^{\mathrm{R,A}}(\varepsilon)=\frac{1}{\varepsilon\mathbf{1}_{\mathrm{C}}-\boldsymbol{H}^{\infty}_{\mathrm{CC}}-\boldsymbol{\Sigma}^{\mathrm{R,A}}(\varepsilon)}.\tag{45}$$

The imaginary part of $\boldsymbol{G}^{\mathrm{R}}$ is simply given by $\boldsymbol{G}^{\mathrm{R}}\mathrm{Im}[\boldsymbol{\Sigma}^{\mathrm{R}}]\boldsymbol{G}^{\mathrm{A}}$. By definition, we have

$$\boldsymbol{\Sigma}^{\mathrm{R,A}}_{\alpha}(\varepsilon)=\boldsymbol{H}_{\mathrm{C}\alpha}\frac{1}{\varepsilon\mathbf{1}_{\alpha}-\boldsymbol{H}^{\infty}_{\alpha\alpha}\pm i\eta}\boldsymbol{H}_{\alpha\mathrm{C}}\tag{46}$$

and hence

$$\mathrm{Im}\left[\boldsymbol{\Sigma}^{\mathrm{R,A}}_{\alpha}(\varepsilon)\right]=\mp\pi\sum_{m}\delta(\varepsilon-\varepsilon^{\infty}_{m\alpha})\boldsymbol{\Gamma}_{m\alpha}.\tag{47}$$

Using the above identity, the steady-state current can be rewritten in a Landauer-like [22] form

$$I^{(\mathrm{S})}_{\mathrm{R}}=-e\sum_{m}[f_{m\mathrm{L}}\mathcal{T}_{m\mathrm{L}}-f_{m\mathrm{R}}\mathcal{T}_{m\mathrm{R}}]=-I^{(\mathrm{S})}_{\mathrm{L}}.\tag{48}$$

In the above formula $\mathcal{T}_{m\mathrm{R}} = \sum_n \mathcal{T}_{m\mathrm{R}}^{n\mathrm{L}}$ and $\mathcal{T}_{m\mathrm{L}} = \sum_n \mathcal{T}_{m\mathrm{L}}^{n\mathrm{R}}$ are the TDDFT transmission coefficients expressed in terms of the quantities

$$\mathcal{T}_{m\alpha}^{n\beta} = 2\pi\delta(\varepsilon_{m\alpha}^{\infty} - \varepsilon_{n\beta}^{\infty})\mathrm{Tr}_{\mathrm{C}}\left\{\boldsymbol{G}^{\mathrm{R}}(\varepsilon_{m\alpha}^{\infty})\boldsymbol{\Gamma}_{m\alpha}\boldsymbol{G}^{\mathrm{A}}(\varepsilon_{n\beta}^{\infty})\boldsymbol{\Gamma}_{n\beta}\right\} = \mathcal{T}_{n\beta}^{m\alpha}. \tag{49}$$

Despite the formal analogy with the Landauer formula, Eq. (48) contains an important conceptual difference since $f_{m\alpha}$ is not simply given by the Fermi distribution function. For example, if the induced change in effective potential varies widely in space deep inside the electrodes, the band structure of the α-electrode Hamiltonian $\boldsymbol{\mathcal{S}}\boldsymbol{H}_{\alpha\alpha}(0)\boldsymbol{\mathcal{S}}^{\dagger}$ might differ from that of $\boldsymbol{H}_{\alpha\alpha}^{\infty}$. However, for metallic electrodes with a macroscopic cross section the switching on of an electric field excites plasmon oscillations, which dynamically screen the external disturbance. Such a metallic screening prevents any rearrangements of the initial equilibrium bulk-density, provided the time-dependent perturbation is slowly varying during a typical plasmon time-scale (which is usually less than a femtosecond). Thus, the KS potential v_s undergoes a uniform time-dependent shift deep inside the left and right electrodes and the KS potential-drop is entirely limited to the central region. Denoting with $\Delta v_{\alpha}(t)$ the difference in electrode α between the KS potential at time t and the KS potential at negative times, $\Delta v_{\alpha}(t) = v_s(\mathbf{r} \in \alpha, t) - v_s(\mathbf{r} \in \alpha, 0)$, to leading order in $1/N$ we then have

$$\boldsymbol{H}_{\alpha\alpha}(t) = \boldsymbol{H}_{\alpha\alpha}(0) + \mathbf{1}_{\alpha}\Delta v_{\alpha}(t), \tag{50}$$

meaning that $\boldsymbol{H}_{\alpha\alpha}^{\infty} = \boldsymbol{H}_{\alpha\alpha}(0) + \mathbf{1}_{\alpha}\Delta v_{\alpha}^{\infty}$. Hence, except for corrections which are of lower order with respect to the system size, $\boldsymbol{\mathcal{S}}\boldsymbol{H}_{\alpha\alpha}(0)\boldsymbol{\mathcal{S}}^{\dagger} = \boldsymbol{H}_{\alpha\alpha}(0)$ and

$$f_{m\alpha} = f(\varepsilon_{m\alpha}^{\infty} - \Delta v_{\alpha}^{\infty}). \tag{51}$$

The formula for the current can be further manipulated when Eq. (51) holds. Let us write the embedding self-energy as the sum of a real and imaginary part $\boldsymbol{\Sigma}_{\alpha}^{\mathrm{R,A}}(\varepsilon) = \boldsymbol{\Lambda}_{\alpha}(\varepsilon) \mp \mathrm{i}\boldsymbol{\Gamma}_{\alpha}(\varepsilon)/2$. Using Eq. (47) we can rewrite the transmission coefficients as

$$\mathcal{T}_{m\mathrm{R}} = \mathrm{Tr}_{\mathrm{C}}\left\{\boldsymbol{G}^{\mathrm{R}}(\varepsilon_{m\mathrm{R}}^{\infty})\boldsymbol{\Gamma}_{m\mathrm{R}}\boldsymbol{G}^{\mathrm{A}}(\varepsilon_{m\mathrm{R}}^{\infty})\boldsymbol{\Gamma}_{\mathrm{L}}(\varepsilon_{m\mathrm{R}}^{\infty})\right\}, \tag{52}$$

$$\mathcal{T}_{m\mathrm{L}} = \mathrm{Tr}_{\mathrm{C}}\left\{\boldsymbol{G}^{\mathrm{R}}(\varepsilon_{m\mathrm{L}}^{\infty})\boldsymbol{\Gamma}_{m\mathrm{L}}\boldsymbol{G}^{\mathrm{A}}(\varepsilon_{m\mathrm{L}}^{\infty})\boldsymbol{\Gamma}_{\mathrm{R}}(\varepsilon_{m\mathrm{L}}^{\infty})\right\}. \tag{53}$$

Substituting these expressions in Eq. (48) and taking into account Eq. (51) we obtain

$$I_{\mathrm{R}}^{(\mathrm{S})} = -e\int \frac{\mathrm{d}\varepsilon}{2\pi}\left[f(\varepsilon - \Delta v_{\mathrm{L}}^{\infty}) - f(\varepsilon - \Delta v_{\mathrm{R}}^{\infty})\right]\mathrm{Tr}_{\mathrm{C}}\left\{\boldsymbol{G}^{\mathrm{R}}(\varepsilon)\boldsymbol{\Gamma}_{\mathrm{L}}(\varepsilon)\boldsymbol{G}^{\mathrm{A}}(\varepsilon)\boldsymbol{\Gamma}_{\mathrm{R}}(\varepsilon)\right\}. \tag{54}$$

In the above equation the Green functions are calculated from Eq. (45). The Hamiltonian $\boldsymbol{H}_{\mathrm{CC}}^{\infty}$ is the KS Hamiltonian $\boldsymbol{H}_s(t \to \infty)$ projected on region C and can be obtained by evolving the system for very long times. According to the Runge–Gross theorem, $\boldsymbol{H}_{\mathrm{CC}}^{\infty}$ depends on how the system was prepared at $t = 0$ (in our case the system is contacted and in thermal equilibrium) and on the full history of the time-dependent density. Therefore, *the use of Eq. (54) in the context of static DFT is generally not correct*. Indeed, static DFT is an equilibrium theory while here we are dealing with a nonequilibrium process. One might argue that in the linear-response regime, the static DFT approach is free from the above criticism. Unfortunately, this is not the case. Denoting with δv_α^∞ the small change Δv_α^∞ of the effective potential in electrode α and with $\delta I_\mathrm{R}^{(\mathrm{S})}$ the corresponding current response, to first order in δv_α^∞ Eq. (54) yields

$$\delta I_\mathrm{R}^{(\mathrm{S})} = -e\int \frac{\mathrm{d}\varepsilon}{2\pi}\frac{\partial f(\varepsilon)}{\partial\varepsilon}\mathrm{Tr}_\mathrm{C}\left\{\boldsymbol{G}_0^\mathrm{R}(\varepsilon)\boldsymbol{\Gamma}_{0,\mathrm{L}}(\varepsilon)\boldsymbol{G}_0^\mathrm{A}(\varepsilon)\boldsymbol{\Gamma}_{0,\mathrm{R}}(\varepsilon)\right\}\left(\delta v_\mathrm{R}^\infty - \delta v_\mathrm{L}^\infty\right). \tag{55}$$

The Green functions and the $\boldsymbol{\Gamma}$'s in Eq. (55) refer to the system in equilibrium and static DFT approaches can be used to evaluate the trace. However, DFT is not enough to calculate the change δv_α^∞. Indeed

$$\delta v_\alpha^\infty = \lim_{t\to\infty}\lim_{x\to\pm\infty}\left[\delta v_\mathrm{ext}(\boldsymbol{r},t) + \delta V_\mathrm{H}(\boldsymbol{r},t) + \delta v_\mathrm{xc}(\boldsymbol{r},t)\right], \tag{56}$$

where x is the longitudinal coordinate, the plus sign applies for $\alpha = \mathrm{R}$ and the minus sign for $\alpha = L$. In the above equation v_ext is the external potential and V_H is the Hartree potential; their sum gives the electrostatic Coulomb potential v_C,

$$\delta v_{\alpha,\mathrm{C}} = \lim_{t\to\infty}\lim_{x\to\pm\infty}\left[\delta v_\mathrm{ext}(\boldsymbol{r},t) + \delta V_\mathrm{H}(\boldsymbol{r},t)\right]. \tag{57}$$

The variation δv_xc of the exchange-correlation potential can be expressed in terms of the exchange-correlation kernel $f_\mathrm{xc}(\boldsymbol{r},t;\boldsymbol{r}',t') = \delta v_\mathrm{xc}(\boldsymbol{r},t)/\delta n(\boldsymbol{r}',t')$

$$\delta v_{\alpha,\mathrm{xc}} = \lim_{t\to\infty}\lim_{x\to\pm\infty}\delta v_\mathrm{xc}(\boldsymbol{r},t) = \lim_{t\to\infty}\lim_{x\to\pm\infty}\int \mathrm{d}\boldsymbol{r}'\int \mathrm{d}t' f_\mathrm{xc}(\boldsymbol{r},t;\boldsymbol{r}',t')\delta n(\boldsymbol{r}',t'). \tag{58}$$

The kernel f_xc depends only on the difference $t-t'$. We denote by $f_{\alpha,\mathrm{xc}}(\boldsymbol{r}',\omega)$ the Fourier transform of f_xc evaluated at $x = \pm\infty$ for $\alpha = \mathrm{R},\mathrm{L}$. Then

$$\delta v_{\alpha,\mathrm{xc}} = \lim_{t\to\infty}\int\frac{\mathrm{d}\omega}{2\pi}\mathrm{e}^{-\mathrm{i}\omega t}\int \mathrm{d}\boldsymbol{r}' f_{\alpha,\mathrm{xc}}(\boldsymbol{r}',\omega)\delta n(\boldsymbol{r}',\omega) \tag{59}$$

with $\delta n(\boldsymbol{r},\omega)$ the Fourier transform of $\delta n(\boldsymbol{r},t)$. Rewriting δv_α^∞ as $\delta v_{\alpha,\mathrm{C}} + \delta v_{\alpha,\mathrm{xc}}$ and taking into account Eq. (59), the current response $\delta I_\mathrm{R}^{(\mathrm{S})}$ in Eq. (55) can also be written as

$$\begin{aligned}\delta I_\mathrm{R}^{(\mathrm{S})} = -e\int\frac{\mathrm{d}\varepsilon}{2\pi}\frac{\partial f(\varepsilon)}{\partial\varepsilon}T(\varepsilon)\Bigg[&\left(\delta v_{\mathrm{R,C}} - \delta v_{\mathrm{L,C}}\right) + \lim_{t\to\infty}\int\frac{\mathrm{d}\omega}{2\pi}\mathrm{e}^{-\mathrm{i}\omega t}\\ &\times\int \mathrm{d}\boldsymbol{r}'\left(f_{\mathrm{R,xc}}(\boldsymbol{r}',\omega) - f_{\mathrm{L,xc}}(\boldsymbol{r}',\omega)\right)\delta n(\boldsymbol{r}',\omega)\Bigg]\end{aligned} \tag{60}$$

with $T(\varepsilon) = \mathrm{Tr}_\mathrm{C}\left\{\boldsymbol{G}_0^\mathrm{R}(\varepsilon)\boldsymbol{\Gamma}_{0,\mathrm{L}}(\varepsilon)\boldsymbol{G}_0^\mathrm{A}(\varepsilon)\boldsymbol{\Gamma}_{0,\mathrm{R}}(\varepsilon)\right\}$. At zero temperature $\partial f(\varepsilon)/\partial\varepsilon = \delta(\varepsilon - \varepsilon_F)$, with ε_F being the Fermi energy, and Eq. (60) becomes

$$\begin{aligned}\delta I_\mathrm{R}^{(\mathrm{S})} = G_\mathrm{KS}(\varepsilon_F)\Bigg[&\left(\delta v_{\mathrm{R,C}} - \delta v_{\mathrm{L,C}}\right) + \lim_{t\to\infty}\int\frac{\mathrm{d}\omega}{2\pi}\mathrm{e}^{-\mathrm{i}\omega t}\\ &\times\int \mathrm{d}\boldsymbol{r}'\left(f_{\mathrm{R,xc}}(\boldsymbol{r}',\omega) - f_{\mathrm{L,xc}}(\boldsymbol{r}',\omega)\right)\delta n(\boldsymbol{r}',\omega)\Bigg]\end{aligned} \tag{61}$$

where $G_{\rm KS}(\varepsilon_F) = -eT(\varepsilon_F)/2\pi$ is the conductance of the KS system. We conclude that *also in the linear-response regime static DFT is not appropriate* for calculating the conductance since dynamical exchange-correlation effects might contribute through the last term in Eq. (61). Equation (61) can also be obtained within the framework of time-dependent current density functional theory as it has been shown in [23].

We emphasize that the steady-state current in Eq. (48) results from a pure dephasing mechanism in the fictitious noninteracting problem. The damping effects of scattering are described by $A_{\rm xc}$ and $v_{\rm xc}$. Furthermore, the current depends only on the asymptotic value of the KS potential, $v_s(\mathbf{r}, t \to \infty)$. However, $v_s(\mathbf{r}, t \to \infty)$ might depend on the history of the external applied potential and the resulting steady-state current might be history dependent. In these cases the full-time evolution cannot be avoided. In the case of Time Dependent Local Density Approximation (TDLDA), the exchange-correlation potential $v_{\rm xc}$ depends only locally on the instantaneous density and has no memory at all. If the density tends to a constant, so does the KS potential v_s, which again implies that the density tends to a constant. Owing to the nonlinearity of the problem there might still be more than one steady-state solution or none at all. We are currently investigating the possibility of having more than one steady-state solution.

4. Quantum transport: A practical scheme based on TDDFT

The theory presented in the previous sections allows us to calculate the time-dependent current in terms of the Green function $\mathcal{G}_{\rm CC} = \boldsymbol{G}$ projected in the central region. In practice, it is computationally very expensive to propagate $\boldsymbol{G}(z; z')$ in time (because it depends on two time variables) and also calculate $\boldsymbol{Q}_\alpha$ from Eq. (37). Here we describe a feasible numerical scheme based on the propagation of KS orbitals. We remind the reader that our electrode–junction–electrode system is infinite and non-periodic. Since one can, in practice, only deal with finite systems we will propagate KS orbitals projected in the central region C by applying the correct boundary conditions [10].

We specialize the discussion to nonmagnetic systems at zero temperature and we denote with $\psi_s(\boldsymbol{r}, 0) \equiv \langle \boldsymbol{r}|\psi_s(0)\rangle$ the eigenstates of $\boldsymbol{H}_s(t < 0)$. The time dependent density can be computed in the usual way by $n(\boldsymbol{r}, t) = \sum_{\rm occ} |\psi_s(\boldsymbol{r}, t)|^2$, where the sum is over the occupied Kohn–Sham orbitals and $|\psi_s(t)\rangle$ is the solution of the KS equation of TDDFT $\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}|\psi_s(t)\rangle = \boldsymbol{H}_s(t)|\psi_s(t)\rangle$. Using the continuity equation, we can write the total current $I_\alpha(t)$ of Eq. (26) as

$$
\begin{aligned}
I_\alpha(t) &= -e \sum_{\rm occ} \int_\alpha \mathrm{d}\boldsymbol{r} \boldsymbol{\nabla} \cdot \mathrm{Im}\,[\psi_s^*(\boldsymbol{r}, t) \boldsymbol{\nabla} \psi_s(\boldsymbol{r}, t)] \\
&= -e \sum_{\rm occ} \int_{S_\alpha} \mathrm{d}\sigma\, \hat{\mathbf{n}} \cdot \mathrm{Im}\,[\psi_s^*(\boldsymbol{r}, t) \boldsymbol{\nabla} \psi_s(\boldsymbol{r}, t)]
\end{aligned}
\tag{62}
$$

where $\hat{\mathbf{n}}$ is the unit vector perpendicular to the surface element $\mathrm{d}\sigma$ and the surface S_α is perpendicular to the longitudinal geometry of our system. From Eq. (62) we conclude that in order to calculate $I_\alpha(t)$ we only need to know the time-evolved KS orbitals in region C. This is possible provided we know the dynamics of the remote parts of the system. As at the end of Section 3.3, we restrict ourselves to metallic electrodes. Then, the external potential and the disturbance introduced by the device region are screened deep inside the electrodes. As the system size increases, the remote parts are less disturbed by the junction and the density inside the electrodes approaches the equilibrium bulk-density. Thus, the macroscopic size of the electrodes leads to an enormous simplification since the initial-state self-consistency is not disturbed far away from the constriction. Partitioning the KS Hamiltonian as in Eq. (27), the time-dependent Schrödinger equation reads

$$\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}|\psi_{\mathrm{L}}\rangle\\|\psi_{\mathrm{C}}\rangle\\|\psi_{\mathrm{R}}\rangle\end{bmatrix}=\begin{bmatrix}\boldsymbol{H}_{\mathrm{LL}} & \boldsymbol{H}_{\mathrm{LC}} & 0\\ \boldsymbol{H}_{\mathrm{CL}} & \boldsymbol{H}_{\mathrm{CC}} & \boldsymbol{H}_{\mathrm{CR}}\\ 0 & \boldsymbol{H}_{\mathrm{RC}} & \boldsymbol{H}_{\mathrm{RR}}\end{bmatrix}\begin{bmatrix}|\psi_{\mathrm{L}}\rangle\\|\psi_{\mathrm{C}}\rangle\\|\psi_{\mathrm{R}}\rangle\end{bmatrix}, \tag{63}$$

where $|\psi_\alpha\rangle$ is the projected wave-function onto the region $\alpha = \mathrm{L}, \mathrm{R}, \mathrm{C}$. We can solve the differential equation for ψ_{L} and ψ_{R} in terms of the retarded Green function $\boldsymbol{g}^{\mathrm{R}}_{\alpha\alpha}$. Then, we have for $\alpha = \mathrm{L}, \mathrm{R}$

$$|\psi_\alpha(t)\rangle = i\boldsymbol{g}^{\mathrm{R}}_{\alpha\alpha}(t,0)|\psi_\alpha(0)\rangle + \int_0^t \mathrm{d}t'\boldsymbol{g}^{\mathrm{R}}_{\alpha\alpha}(t,t')\boldsymbol{H}_{\alpha\mathrm{C}}|\psi_{\mathrm{C}}(t')\rangle. \tag{64}$$

Using Eq. (64), the equation for ψ_{C} can be written as

$$\begin{aligned}\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}|\psi_{\mathrm{C}}(t)\rangle = \boldsymbol{H}_{\mathrm{CC}}(t)|\psi_{\mathrm{C}}(t)\rangle + \textstyle\int_0^t \mathrm{d}t'\boldsymbol{\Sigma}^{\mathrm{R}}(t,t')|\psi_{\mathrm{C}}(t')\rangle\\ + i\sum_{\alpha=\mathrm{L,R}}\boldsymbol{H}_{\mathrm{C}\alpha}\boldsymbol{g}^{\mathrm{R}}_{\alpha\alpha}(t,0)|\psi_\alpha(0)\rangle,\end{aligned} \tag{65}$$

where $\boldsymbol{\Sigma}^{\mathrm{R}} = \sum_{\alpha=\mathrm{L,R}}\boldsymbol{H}_{\mathrm{C}\alpha}\boldsymbol{g}^{\mathrm{R}}_{\alpha\alpha}\boldsymbol{H}_{\alpha\mathrm{C}}$, in accordance with Eq. (38). Thus, for any given KS orbital we can evolve its projection onto the central region by solving Eq. (65) in region C. Equation (65) has also been derived elsewhere (for static Hamiltonians) [24]. To summarize, all the complexity of the infinite electrode–junction–electrode system has been reduced to the solution of an open quantum-mechanical system (the central region C) with proper time-dependent boundary conditions.

Equation (65) is the central equation of our numerical approach to time-dependent transport. It is a reformulation of the original time-dependent Schrödinger equation (63) of the full system in terms of an equation for the central (device) region only. The coupling to the leads is taken into account by the lead Green functions $\boldsymbol{g}^{\mathrm{R}}_{\alpha\alpha}$, $\alpha = \mathrm{L}, \mathrm{R}$. Equation (65) has the structure of a time-dependent Schrödinger equation with two extra terms. The first term describes the injection of particles induced by a nonvanishing projection of the initial wave-function onto the leads. The second term involves the self-energy $\boldsymbol{\Sigma}^{\mathrm{R}}$ and the wavefunction in the central region at previous times during the propagation. We will denote it as the memory integral. We should remark here that these memory effects are of different origin than those which are usually discussed in the context of TDDFT [25, 26]. The latter ones arise from the dependence of the exchange-correlation functional on the full history of the time-dependent density. Most density-based functionals used at present rely on the adiabatic approximation therefore ignoring the functional dependence on past time-dependent densities [27].

Equation (65) is first order in time, therefore we need to specify an initial state which is to be propagated. We want to study the time evolution of systems perturbed out of their equilibrium ground state. Of course, the ground state of our noninteracting system is the Slater determinant of the occupied eigenstates of the full, extended Hamiltonian in equilibrium, $\boldsymbol{H}_s(t<0)$. The practical question then arises as to how one can obtain these eigenstates and how one can propagate them in time without having to deal explicitly with the extended Hamiltonian. Below, we show how we have coped with these problems.

4.1. Computation of KS eigenstates

Let us consider our electrode–junction–electrode system in equilibrium ($t<0$) and let $\psi_s(\boldsymbol{r})=\psi_{Ej}(\boldsymbol{r})$ be the j-th degenerate eigenstate of energy E of the KS Hamiltonian $\boldsymbol{H}_s$. The Green functions $\boldsymbol{\mathcal{G}}^{\mathrm{R,A}}(t;t')$ and $\boldsymbol{\mathcal{G}}^{<}(t;t')$ of the undisturbed system depend only on the difference $t-t'$. In absence of magnetic fields, $\boldsymbol{H}_s$ is invariant under time-reversal and the imaginary part of the Fourier transformed $\boldsymbol{\mathcal{G}}^{\mathrm{R}}$ is simply given by

$$-\frac{1}{\pi}\mathrm{Im}\,[\langle\boldsymbol{r}|\boldsymbol{\mathcal{G}}(E)|\boldsymbol{r}'\rangle]=\sum_{E'}\delta(E-E')\sum_{j=1}^{d_{E'}}\psi_{E'j}(\boldsymbol{r})\psi^*_{E'j}(\boldsymbol{r}')\,. \tag{66}$$

Multiplying Eq. (66) by $\psi^*_{Em}(\boldsymbol{r})\psi_{En}(\boldsymbol{r}')$ and integrating over $\boldsymbol{r}$ and $\boldsymbol{r}'$ in region C we obtain

$$\begin{aligned}&-\frac{1}{\pi}\int_{\mathrm{C}}\mathrm{d}\boldsymbol{r}\int_{\mathrm{C}}\mathrm{d}\boldsymbol{r}'\psi^*_{Em}(\boldsymbol{r})\mathrm{Im}\,[\langle\boldsymbol{r}|\boldsymbol{\mathcal{G}}(E)|\boldsymbol{r}'\rangle]\,\psi_{En}(\boldsymbol{r}')\\&=\sum_{E'}\delta(E-E')\sum_{j=1}^{d_{E'}}S_{mj}(E')S_{jn}(E'),\end{aligned} \tag{67}$$

where

$$S_{mj}(E)\equiv\int_{\mathrm{C}}\mathrm{d}\boldsymbol{r}\;\psi^*_{Em}(\boldsymbol{r})\psi_{Ej}(\boldsymbol{r})=S^*_{jm}(E) \tag{68}$$

is the overlap matrix in region C between degenerate states. This matrix is Hermitian and can be diagonalized, i.e.,

$$\sum_{j=1}^{d_E}S_{mj}(E)a_j^{(l)}(E)=\lambda_l(E)a_m^{(l)}(E). \tag{69}$$

Next, we multiply Eq. (67) by $a_m^{(l)*}(E)a_n^{(l')}(E)$ and sum over m and n. The result can be written in terms of the new eigenfunctions $a_{El}(\boldsymbol{r})=\sum_{n=1}^{d_E}a_n^{(l)}(E)\psi_{En}(\boldsymbol{r})$ as

$$-\frac{1}{\pi}\int_{\mathrm{C}}\mathrm{d}\boldsymbol{r}\int_{\mathrm{C}}\mathrm{d}\boldsymbol{r}'a^*_{El}(\boldsymbol{r})\mathrm{Im}\,[\langle\boldsymbol{r}|\boldsymbol{\mathcal{G}}(E)|\boldsymbol{r}'\rangle]\,a_{El'}(\boldsymbol{r}')=\delta_{ll'}\lambda_l^2(E)\sum_{E'}\delta(E-E'), \tag{70}$$

where we have used Eq. (69) and the orthonormality of the S-matrix eigenvectors: $\sum_{j=1}^{d_E}a_j^{(l)*}(E)a_j^{(l')}(E)=\delta_{ll'}$. Equation (70) shows explicitly that the functions $a_{Ej}(\boldsymbol{r})$ diagonalize $\mathrm{Im}\,[\boldsymbol{\mathcal{G}}_{\mathrm{CC}}(E)]$ in the central region and that the eigenvalues are positive. Since any linear combination of degenerate eigenstates is again an eigenstate, diagonalizing

$\mathrm{Im}\,[\mathcal{G}_{\mathrm{CC}}(E)]$ gives us one set of linearly independent, degenerate eigenstates of energy E. In our practical implementation described in more detail in Section 5, we diagonalize

$$-\frac{1}{\pi D_{\mathrm{C}}(E)}\mathrm{Im}\,[\mathcal{G}_{\mathrm{CC}}(E)] \tag{71}$$

where $D_{\mathrm{C}}(E) = -\frac{1}{\pi}\mathrm{Tr}\ \{\mathrm{Im}\,[\mathcal{G}_{\mathrm{CC}}(E)]\}$ is the total density of states in the central region. If we use N_g grid points to describe the central region, the diagonalization, in principle, gives N_g eigenvectors but only a few have the physical meaning of extended eigenstates at this energy. It is, however, very easy to identify the physical states by looking at the eigenvalues: at a given energy E only d_E eigenvalues are nonvanishing and they always add up to unity. The corresponding states are the physical ones. All the other eigenvalues are zero (or numerically close to zero) and the corresponding states have no physical meaning.

The procedure described above gives the correct extended eigenstates only up to a normalization factor. When diagonalizing Eq. (71) with typical library routines one obtains eigenvectors which are normalized to the central region. Physically this might be incorrect. It is possible to fix the normalization by matching the wavefunction for the central region to the known form (and normalization) of the wavefunction in the macroscopic leads.

It should be emphasized that the procedure described here for the extraction of eigenstates of the extended system from $\mathcal{G}_{\mathrm{CC}}(E)$ only works in practice if E is in the continuous part of the spectrum due to the sharp peak of the delta function in the discrete part of the spectrum. Eigenstates in the discrete part of the spectrum can be found by considering the original Schrödinger equation for the full system: $\boldsymbol{H}_s\psi = E\psi$. Using again the block structure of the Hamiltonian this can be transformed into an effective Schrödinger equation for an *energy-dependent* Hamiltonian for the central region only:

$$\left(\boldsymbol{H}_{\mathrm{CC}} + \sum_{\alpha=\mathrm{L,R}} \boldsymbol{H}_{\mathrm{C}\alpha}\frac{1}{E\mathbf{1}_\alpha - \boldsymbol{H}_{\alpha\alpha}}\boldsymbol{H}_{\alpha\mathrm{C}}\right)|\psi_{\mathrm{C}}\rangle = E|\psi_{\mathrm{C}}\rangle. \tag{72}$$

This equation has solutions only for certain values of E which are the discrete eigenenergies of the full Hamiltonian $\boldsymbol{H}_s$. Since the left and right electrodes form a continuum, the dimension of the kernel of $(E - \boldsymbol{H}_{\alpha\alpha})$ is zero for those energies E in the discrete part of the spectrum. We also notice that the second term in parenthesis in Eq. (72) is nothing but the real part of the retarded/advanced self-energy in equilibrium, see Eq. (47). Bound states as well as fully reflected waves will contribute to the density but not to the current and might play a role in the description of charge-accumulation in molecular transport, as, e.g., in Coulomb blockade phenomenon. In our TDDFT formulation, bound states and fully reflected waves also play an extra role, since they are needed for calculating the effective potential v_s (which is a functional of the density) which is in turn used for extracting all extended states.

4.2. Algorithm for the time evolution

In order to calculate the longitudinal current in an electrode–junction–electrode system we need to propagate the KS orbitals. The main difficulty stems from the macroscopic size of the electrodes whose remote parts, ultimately, are taken infinitely far away from the central, explicitly treated, scattering region C.

The problem can be solved by imposing transparent boundary conditions [28] at the electrode–junction interfaces. Efficient algorithms have been proposed for wave-packets initially *localized* in the scattering region and for Hamiltonians constant in time. In this section we describe an algorithm well suited for delocalized initial states, as well as for localized ones, evolving with a time-dependent Hamiltonian.

Let $\boldsymbol{H}_s(t)$ be the time-dependent KS Hamiltonian. We partition $\boldsymbol{H}_s(t)$ as in Section 3.2. The explicitly treated region C includes the first few atomic layers of the left and right electrodes. The boundaries of this region are chosen in such a way that the density outside C is accurately described by an equilibrium bulk density. It is convenient to write $\boldsymbol{H}_{\alpha\alpha}(t)$, with $\alpha = \mathrm{L,R}$, as the sum of a term $\boldsymbol{H}^0_{\alpha\alpha} = \boldsymbol{H}_{\alpha\alpha}(0)$ which is constant in time and another term $\boldsymbol{U}_\alpha(t)$ which is explicitly time-dependent, $\boldsymbol{H}_{\alpha\alpha}(t) = \boldsymbol{H}^0_{\alpha\alpha} + \boldsymbol{U}_\alpha(t)$. In configuration space, $\boldsymbol{U}_\alpha(t)$ is diagonal at any time t since the KS potential is local in space. Furthermore, the diagonal elements $U_\alpha(\boldsymbol{r}, t)$ are spatially constant for metallic electrodes. Thus, $\boldsymbol{U}_\alpha(t) = U_\alpha(t)\mathbf{1}_\alpha$ and $U_\mathrm{L}(t) - U_\mathrm{R}(t)$ is the total potential drop across the central region. We write $\boldsymbol{H}_s(t) = \tilde{\boldsymbol{H}}(t) + \boldsymbol{U}(t)$ with

$$\tilde{\boldsymbol{H}}(t) = \begin{bmatrix} \boldsymbol{H}^0_\mathrm{LL} & \boldsymbol{H}_\mathrm{LC} & 0 \\ \boldsymbol{H}_\mathrm{CL} & \boldsymbol{H}_\mathrm{CC}(t) & \boldsymbol{H}_\mathrm{CR} \\ 0 & \boldsymbol{H}_\mathrm{RC} & \boldsymbol{H}^0_\mathrm{RR} \end{bmatrix}, \quad \text{and} \quad \boldsymbol{U}(t) = \begin{bmatrix} U_\mathrm{L}(t)\mathbf{1}_\mathrm{L} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & U_\mathrm{R}(t)\mathbf{1}_\mathrm{R} \end{bmatrix}. \tag{73}$$

In this way, the only term in $\tilde{\boldsymbol{H}}(t)$ that depends on t is $\boldsymbol{H}_\mathrm{CC}(t)$. For any given initial state $|\psi(0)\rangle = |\psi^{(0)}\rangle$ we calculate $|\psi(t_m = m\Delta t)\rangle = |\psi^{(m)}\rangle$ by using a generalized form of the Cayley method

$$\left(\mathbf{1} + \mathrm{i}\delta\tilde{\boldsymbol{H}}^{(m)}\right)\frac{\mathbf{1} + \mathrm{i}\frac{\delta}{2}\boldsymbol{U}^{(m)}}{\mathbf{1} - \mathrm{i}\frac{\delta}{2}\boldsymbol{U}^{(m)}}|\psi^{(m+1)}\rangle = \left(\mathbf{1} - \mathrm{i}\delta\tilde{\boldsymbol{H}}^{(m)}\right)\frac{\mathbf{1} - \mathrm{i}\frac{\delta}{2}\boldsymbol{U}^{(m)}}{\mathbf{1} + \mathrm{i}\frac{\delta}{2}\boldsymbol{U}^{(m)}}|\psi^{(m)}\rangle, \tag{74}$$

with $\tilde{\boldsymbol{H}}^{(m)} = \frac{1}{2}[\tilde{\boldsymbol{H}}(t_{m+1}) + \tilde{\boldsymbol{H}}(t_m)]$, $\boldsymbol{U}^{(m)} = \frac{1}{2}[\boldsymbol{U}(t_{m+1}) + \boldsymbol{U}(t_m)]$ and $\delta = \Delta t/2$. It should be noted that our propagator is norm conserving (unitary) and accurate to second order in δ, as is the Cayley propagator [29]. Denoting by $|\psi_\alpha\rangle$ the projected wave function onto the region $\alpha = \mathrm{R, L, C}$, we find from Eq. (74)

$$|\psi_\mathrm{C}^{(m+1)}\rangle = \frac{\mathbf{1}_\mathrm{C} - \mathrm{i}\delta\boldsymbol{H}^{(\mathrm{m})}_\mathrm{eff}}{\mathbf{1}_\mathrm{C} + \mathrm{i}\delta\boldsymbol{H}^{(m)}_\mathrm{eff}}|\psi_\mathrm{C}^{(m)}\rangle + |S^{(m)}\rangle - |M^{(m)}\rangle. \tag{75}$$

Here, $\boldsymbol{H}^{(m)}_\mathrm{eff}$ is the effective Hamiltonian of the central region:

$$\boldsymbol{H}^{(m)}_\mathrm{eff} = \boldsymbol{H}^{(m)}_\mathrm{CC} - \mathrm{i}\delta\boldsymbol{H}_\mathrm{CL}\frac{1}{\mathbf{1}_\mathrm{L} + \mathrm{i}\delta\boldsymbol{H}^0_\mathrm{LL}}\boldsymbol{H}_\mathrm{LC} - \mathrm{i}\delta\boldsymbol{H}_\mathrm{CR}\frac{1}{\mathbf{1}_\mathrm{R} + \mathrm{i}\delta\boldsymbol{H}^0_\mathrm{RR}}\boldsymbol{H}_\mathrm{RC} \tag{76}$$

with $\boldsymbol{H}^{(m)}_\mathrm{CC} = \frac{1}{2}[\boldsymbol{H}_\mathrm{CC}(t_{m+1}) + \boldsymbol{H}_\mathrm{CC}(t_m)]$. The source term $|S^{(m)}\rangle$ describes the injection of density into the region C, while the memory term $|M^{(m)}\rangle$ is responsible for the hopping in and out of the region C. In terms of the propagator for the uncontacted and undisturbed α electrode

$$\boldsymbol{g}_\alpha = \frac{\mathbf{1}_\alpha - \mathrm{i}\delta\boldsymbol{H}^0_{\alpha\alpha}}{\mathbf{1}_\alpha + \mathrm{i}\delta\boldsymbol{H}^0_{\alpha\alpha}}, \tag{77}$$

the source term can be written as

$$|S^{(m)}\rangle = -\frac{2\mathrm{i}\delta}{\mathbf{1}_{\mathrm{C}} + \mathrm{i}\delta \boldsymbol{H}^{(m)}_{\mathrm{eff}}} \sum_{\alpha=\mathrm{L,R}} \frac{\Lambda^{(m,0)}_{\alpha}}{u^{(m)}_{\alpha}} \boldsymbol{H}_{\mathrm{C}\alpha} \frac{[\boldsymbol{g}_{\alpha}]^{m}}{\mathbf{1}_{\alpha} + \mathrm{i}\delta \boldsymbol{H}^{s}_{\alpha\alpha}} |\psi^{(0)}_{\alpha}\rangle, \tag{78}$$

with

$$u^{(m)}_{\alpha} = \frac{1 - \mathrm{i}\frac{\delta}{2} U^{(m)}_{\alpha}}{1 + \mathrm{i}\frac{\delta}{2} U^{(m)}_{\alpha}} \quad \text{and} \quad \Lambda^{(m,k)}_{\alpha} = \prod_{j=k}^{m} [u^{(j)}_{\alpha}]^{2}. \tag{79}$$

For a wave packet initially localized in C the projection onto the left and right electrode $|\psi^{(0)}_{\alpha}\rangle$ vanishes and $|S^{(m)}\rangle = 0$ for any m, as it should be. The memory term is more complicated and reads

$$|M^{(m)}\rangle = -\frac{\delta^{2}}{\mathbf{1}_{\mathrm{C}} + \mathrm{i}\delta \boldsymbol{H}^{(m)}_{\mathrm{eff}}} \sum_{\alpha=\mathrm{L,R}} \sum_{k=0}^{m-1} \frac{\Lambda^{(m,k)}_{\alpha}}{u^{(m)}_{\alpha} u^{(k)}_{\alpha}} \left[\boldsymbol{Q}^{(m-k)}_{\alpha} + \boldsymbol{Q}^{(m-k-1)}_{\alpha}\right] \\ \times \left(|\psi^{(k+1)}_{\mathrm{C}}\rangle + |\psi^{(k)}_{\mathrm{C}}\rangle\right) \tag{80}$$

where

$$\boldsymbol{Q}^{(m)}_{\alpha} = \boldsymbol{H}_{\mathrm{C}\alpha} \frac{[\boldsymbol{g}_{\alpha}]^{m}}{\mathbf{1}_{\alpha} + \mathrm{i}\delta \boldsymbol{H}^{s}_{\alpha\alpha}} \boldsymbol{H}_{\alpha\mathrm{C}}. \tag{81}$$

The quantities $\boldsymbol{Q}^{(m)}_{\alpha}$ depend on the geometry of the system and are independent of the initial state $\psi^{(0)}$.

Below, we propose a recursive scheme to calculate the $\boldsymbol{Q}^{(m)}_{\alpha}$'s for those system geometries having semiperiodic electrodes along the longitudinal direction, see Figure 3. In this case $\boldsymbol{H}^{0}_{\alpha\alpha}$ has a tridiagonal block form

$$\boldsymbol{H}^{0}_{\alpha\alpha} = \begin{bmatrix} \boldsymbol{h}_{\alpha} & \boldsymbol{V}_{\alpha} & 0 & \dots \\ \boldsymbol{V}_{\alpha} & \boldsymbol{h}_{\alpha} & \boldsymbol{V}_{\alpha} & \dots \\ 0 & \boldsymbol{V}_{\alpha} & \boldsymbol{h}_{\alpha} & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix}, \tag{82}$$

where $\boldsymbol{h}_{\alpha}$ describes a convenient cell and $\boldsymbol{V}_{\alpha}$ is the hopping Hamiltonian between two nearest neighbor cells. Without loss of generality we assume that both $\boldsymbol{h}_{\alpha}$ and $\boldsymbol{V}_{\alpha}$

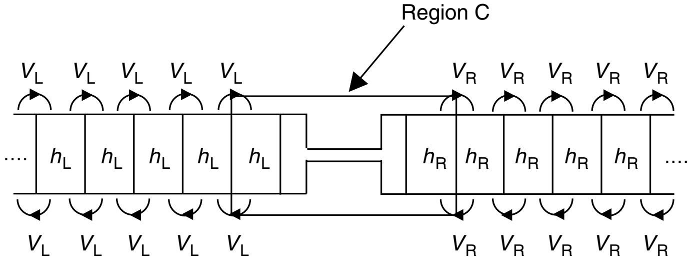


Figure 3 Schematic sketch of an electrode–junction–electrode system with semiperiodic electrodes.

are square matrices of dimension $N_\alpha \times N_\alpha$. Taking into account that the central region contains the first few cells of the left and right electrodes, the matrix $\boldsymbol{Q}_\alpha^{(m)}$ has the following structure

$$\boldsymbol{Q}_{\mathrm{L}}^{(m)} = \begin{bmatrix} \boldsymbol{q}_{\mathrm{L}}^{(m)} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \boldsymbol{Q}_{\mathrm{R}}^{(m)} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \boldsymbol{q}_{\mathrm{R}}^{(m)} \end{bmatrix}. \tag{83}$$

The $\boldsymbol{q}_\alpha^{(m)}$'s are square matrices of dimension $N_\alpha \times N_\alpha$ and are given by

$$\boldsymbol{q}_\alpha^{(m)} = \boldsymbol{V}_\alpha \left[\frac{[\boldsymbol{g}_\alpha]^m}{\mathbf{1}_\alpha + \mathrm{i}\delta \boldsymbol{H}_{\alpha\alpha}} \right]_{1,1} \boldsymbol{V}_\alpha, \tag{84}$$

where the subscript (1, 1) denotes the first diagonal block of the matrix in the square brackets. We introduce the generating matrix function

$$\boldsymbol{q}_\alpha(x, y) \equiv \boldsymbol{V}_\alpha \left[\frac{1}{x\mathbf{1}_\alpha + \mathrm{i}y\delta \boldsymbol{H}_{\alpha\alpha}} \right]_{1,1} \boldsymbol{V}_\alpha, \tag{85}$$

which can also be expressed in terms of continued matrix fractions

$$\begin{aligned} \boldsymbol{q}_\alpha(x, y) &= \boldsymbol{V}_\alpha \frac{1}{x + \mathrm{i}y\delta \boldsymbol{h}_\alpha + y^2\delta^2 \boldsymbol{V}_\alpha \dfrac{1}{x + \mathrm{i}y\delta \boldsymbol{h}_\alpha + y^2\delta^2 \boldsymbol{V}_\alpha \frac{1}{\cdots\cdots} \boldsymbol{V}_\alpha} \boldsymbol{V}_\alpha} \boldsymbol{V}_\alpha \\ &= \boldsymbol{V}_\alpha \frac{1}{x + \mathrm{i}y\delta \boldsymbol{h}_\alpha + y^2\delta^2 \boldsymbol{q}_\alpha(x, y)} \boldsymbol{V}_\alpha. \end{aligned} \tag{86}$$

The $\boldsymbol{q}_\alpha^{(m)}$'s can be obtained from

$$\boldsymbol{q}_\alpha^{(m)} = \frac{1}{m!} \left[-\frac{\partial}{\partial x} + \frac{\partial}{\partial y} \right]^m \boldsymbol{q}_\alpha(x, y) \Big|_{x=y=1}. \tag{87}$$

From Eqs (87) and (86) one can build up a recursive scheme. Let us define

$$\boldsymbol{p}_\alpha^{-1}(x, y) = x + \mathrm{i}y\delta \boldsymbol{h}_\alpha + y^2\delta^2 \boldsymbol{q}_\alpha(x, y)$$

and

$$\boldsymbol{p}_\alpha^{(m)} = \frac{1}{m!} \left[-\frac{\partial}{\partial x} + \frac{\partial}{\partial y} \right]^m \boldsymbol{p}_\alpha(x, y) \Big|_{x=y=1}.$$

Then, by definition, $\boldsymbol{q}_\alpha^{(m)} = \boldsymbol{V}_\alpha \boldsymbol{p}_\alpha^{(m)} \boldsymbol{V}_\alpha$. Using the identity

$$\frac{1}{m!} \left[-\frac{\partial}{\partial x} + \frac{\partial}{\partial y} \right]^m \boldsymbol{p}_\alpha(x, y) \boldsymbol{p}_\alpha^{-1}(x, y) = 0,$$

one finds

$$(1 + \mathrm{i}\delta \boldsymbol{h}_\alpha) \boldsymbol{p}_\alpha^{(m)} = (1 - \mathrm{i}\delta \boldsymbol{h}_\alpha) \boldsymbol{p}_\alpha^{(m-1)} - \delta^2 \sum_{k=0}^{m} (\boldsymbol{q}_\alpha^{(k)} + 2\boldsymbol{q}_\alpha^{(k-1)} + \boldsymbol{q}_\alpha^{(k-2)}) \boldsymbol{p}_\alpha^{(m-k)} \tag{88}$$

with $\boldsymbol{p}_\alpha^{(m)} = \boldsymbol{q}_\alpha^{(m)} = 0$ for $m < 0$. Once $\boldsymbol{q}_\alpha^{(0)}$ has been obtained by solving Eq. (86) with $x = y = 1$, we can calculate $\boldsymbol{p}_\alpha^{(0)} = [1 + i\delta \boldsymbol{h}_\alpha + \delta^2 \boldsymbol{q}_\alpha^{(0)}]^{-1}$. Afterwards, we can use Eq. (88) with $\boldsymbol{q}_\alpha^{(1)} = \boldsymbol{V}_\alpha \boldsymbol{p}_\alpha^{(1)} \boldsymbol{V}_\alpha$ to calculate $\boldsymbol{p}_\alpha^{(1)}$ and hence $\boldsymbol{q}_\alpha^{(1)}$ and so on and so forth.

This concludes the description of our algorithm for the propagation of the time-dependent Schrödinger equation for extended systems. It is worth mentioning an additional complication here which arises for the propagation of a time-dependent Kohn–Sham equation. This complication stems from the fact that in order to compute $|\psi_{\mathrm{C}}^{(m+1)}\rangle$ at time step $m+1$, one needs to know the time-dependent KS potential at the same time step which, via the Hartree and exchange-correlation potentials, depends on the yet unknown orbitals $|\psi_{\mathrm{C}}^{(m+1)}\rangle$. Of course, the solution is to use a predictor-corrector approach: in the first step one approximates $\boldsymbol{H}_{\mathrm{CC}}^{(m)}$ by $\boldsymbol{H}_{CC}(t_m)$, computes new orbitals $|\tilde{\psi}_{\mathrm{C}}^{(m+1)}\rangle$ and from those obtains an improved approximation for $\boldsymbol{H}_{\mathrm{CC}}^{(m)}$.

5. Implementation details for one-dimensional systems and numerical results

All the methodological discussion of Section 4 is general and can be applied to all systems having a longitudinal geometry like the one in Figure 3. In this section we show that the proposed scheme is feasible by testing it against one-dimensional model systems. The extension to real molecular-device configurations is presently under development [30]. We consider systems described by the Hamiltonian

$$\langle x|\boldsymbol{H}|x'\rangle = \delta(x - x')\left[-\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}x^2} + V(x)\right]. \tag{89}$$

We have used a simple three-point discretization for the second derivative

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}\psi(x)|_{x=x_i} \approx \frac{1}{(\Delta x)^2}\left[\psi(x_{i+1}) - 2\psi(x_i) + \psi(x_{i-1})\right] \tag{90}$$

with equidistant grid points x_i, $i = 1, \ldots, N_g$ and spacing Δx. Within this approximation, matrices of the form $\boldsymbol{H}_{\mathrm{C}\alpha}\boldsymbol{M}\boldsymbol{H}_{\alpha\mathrm{C}}$, which are $N_g \times N_g$ matrices and appear, e.g., in Eq. (38) or (81), have only one nonvanishing matrix element. For $\alpha = \mathrm{L}$ this is the $(1, 1)$ element, for $\alpha = \mathrm{R}$ it is the (N_g, N_g) element.

In order to proceed, we have to specify the nature of the leads and therefore the lead Green function. Here we choose the simplest case of semi-infinite, uniform leads at constant potential $U_{\alpha 0}$. In this case, the retarded Green function $\boldsymbol{g}_{\alpha\alpha}^{\mathrm{R}}$ in the energy domain can be given in closed form:

$$\begin{aligned}[\boldsymbol{g}_{\alpha\alpha}^{\mathrm{R}}(E)]_{kl} = &-\frac{\mathrm{i}\Delta x}{\sqrt{2\tilde{E}_\alpha}}\exp\left\{\mathrm{i}\sqrt{2\tilde{E}_\alpha}|x_k - x_l|\right\} \\ &+ \frac{\mathrm{i}\Delta x}{\sqrt{2\tilde{E}_\alpha}}\exp\left\{\mathrm{i}\sqrt{2\tilde{E}_\alpha}(|x_k - x_{\alpha 0}| + |x_l - x_{\alpha 0}|)\right\}\end{aligned} \tag{91}$$

with $\tilde{E}_\alpha = E - U_{\alpha 0}$. The abscissa $x_{\alpha 0}$ is the position of the interface between the lead and the device region; in our implementation $x_{\mathrm{L}0}$ is the first grid point of region C while x_{R0} is the N_g-th grid point of region C. According to the notation in Eq. (63) the one-particle state of region C describing an electron localized in x_{L0} is denoted by $|x_{\mathrm{C1}}\rangle$ while the one-particle state of region C describing an electron localized in x_{R0} is denoted by $|x_{\mathrm{C}N_g}\rangle$. The coordinate $x_k = x_{\alpha 0} \pm k\Delta x$, $k > 0$, where the plus sign applies for $\alpha = \mathrm{R}$ and the minus sign for $\alpha = \mathrm{L}$.

The results of the procedure for calculating extended eigenstates as described in Section 4.1 are illustrated in Figure 4 for a square potential barrier with zero potential in both leads. In the left panel we have the square modulus of eigenstates at an energy below the barrier height, while in the right panel eigenstates with energy higher than the barrier are shown. The states result from diagonalization of Eq. (71). In order to obtain the normalization constant we compute the logarithmic derivative at the boundary of the central region numerically and match it to the analytic form in the lead to obtain the phase shift δ_α:

$$\frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2}\ln(|\psi(x)|^2)\bigg|_{x=x_{\alpha 0}} = q\cot(\delta_\alpha) \tag{92}$$

where $q = \sqrt{2\tilde{E}_\alpha}$. Knowing the phase shift we can rescale the wavefunction such that it matches with the analytic form $\sin(q(x - x_{\alpha 0}) + \delta_\alpha)$ at the interface. Of course, this form of the extended states only applies for $\tilde{E}_\alpha > 0$ but as long as E is in the continuous part of the spectrum, it is correct at least for one of the leads. This is sufficient to determine the normalization. The states obtained numerically with this procedure coincide with the known analytical results.

We then implemented the propagation scheme presented in the previous section. Within our three-point approximation, $\boldsymbol{h}_\alpha$, $\boldsymbol{V}_\alpha$ and $\boldsymbol{q}_\alpha$ are 1×1 matrices. The equation for $q_\alpha^{(0)}$ [see Eqs (86) and (87)] becomes a simple quadratic equation which can be solved explicitly

$$q_\alpha^{(0)} = \frac{-(1 + \mathrm{i}\delta h_\alpha) + \sqrt{(1 + \mathrm{i}\delta h_\alpha)^2 + 4(\delta V_\alpha)^2}}{2\delta^2}. \tag{93}$$

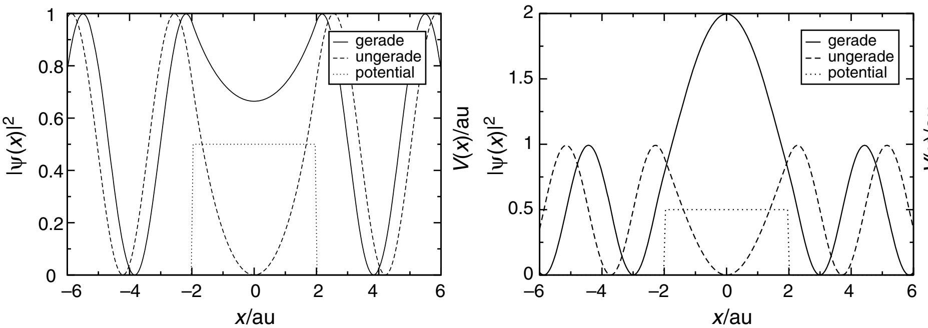


Figure 4 Continuum states of square potential barrier at different energies with leads at zero potential. Left panel: eigenstates for $\varepsilon = 0.45$ au, just below the barrier height of 0.5 au. Right panel: eigenstates at $\varepsilon = 0.6$ au

Although the quadratic equation has two solutions, the above choice for $q_\alpha^{(0)}$ is dictated by the fact that the Taylor expansions for small δ of Eqs (93) and (86) have to coincide. Using this result we then solved the iterative scheme to obtain the $q_\alpha^{(m)}$ for $m \geq 1$.

As a first check on the propagation method we propagated a Gaussian wavepacket which, at initial time $t = 0$, is completely localized in the central device region. (The source terms $|S^{(m)}\rangle$ then vanish identically.) As can be seen in Figure 5, the wavepacket correctly propagates through the boundaries without any spurious reflections.

For the propagation of the extended initial states (the eigenstates of the unperturbed system) we also need to implement the source terms $|S^{(m)}\rangle$. In the following we assume that the left and right leads are at the same potential initially so that the analytic form of the initial states is in both leads given by $\sin(q(x - x_{\alpha 0}) + \delta_\alpha) = [\exp(\mathrm{i}\delta_\alpha - \mathrm{i}qx_{\alpha 0})\exp(\mathrm{i}qx) - \mathrm{c.c.}]/2\mathrm{i}$. Let us specialize the discussion to the case $\alpha = \mathrm{R}$ and define the state $|q_\mathrm{R}\rangle$ according to $\langle x_{\mathrm{R}k}|q_\mathrm{R}\rangle = \exp(\mathrm{i}qk\Delta x)$, where $|x_{\mathrm{R}k}\rangle$ is the one-particle state of electrode R describing an electron localized in $x_k = x_{\mathrm{R}0} + k\Delta x$, $k > 0$. Then, the projection of the initial state onto lead R reads $|\psi_\mathrm{R}^{(0)}\rangle = \frac{1}{2\mathrm{i}}[\exp(\mathrm{i}\delta_\alpha)|q_R\rangle - \exp(-\mathrm{i}\delta_\alpha)| - q_\mathrm{R}\rangle]$. From Eq. (78) the contribution to the source term for $\alpha = \mathrm{R}$ is completely known once we know how $\boldsymbol{H}_{\mathrm{CR}}[\boldsymbol{g}_\mathrm{R}]^m/(\mathbf{1}_\mathrm{R} + \mathrm{i}\delta\boldsymbol{H}_{\mathrm{RR}})$ acts on the state $|q_\mathrm{R}\rangle$. We have

$$\boldsymbol{H}_{\mathrm{CR}}\frac{[\boldsymbol{g}_\mathrm{R}]^m}{(\mathbf{1}_R + \mathrm{i}\delta\boldsymbol{H}_{\mathrm{RR}})}|q_\mathrm{R}\rangle = V_\mathrm{R}|x_{CN_g}\rangle\langle x_{\mathrm{R}1}|\frac{[\boldsymbol{g}_\mathrm{R}]^m}{(\mathbf{1}_\mathrm{R} + \mathrm{i}\delta\boldsymbol{H}_{\mathrm{RR}})}|q_\mathrm{R}\rangle \tag{94}$$

where x_{CN_g} corresponds the N_g-th discretization point of region C (the last point on the right before electrode R starts). We rewrite the unknown quantity as follows

$$\left\langle x_{\mathrm{R}1}\left|\frac{[\boldsymbol{g}_\mathrm{R}]^m}{\mathbf{1}_\mathrm{R} + \mathrm{i}\delta\boldsymbol{H}_{\mathrm{RR}}}\right|q_\mathrm{R}\right\rangle = \left.\frac{[D(x, y)]^m}{m!}\rho(x, y)\right|_{x=y=1}, \tag{95}$$

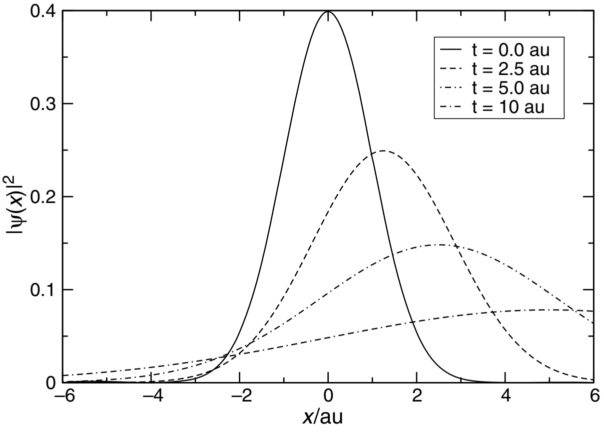


Figure 5 Time evolution of a Gaussian wavepacket with initial width 1.0 au and initial momentum 0.5 au for various propagation times. The transparent boundary conditions allow the wavepacket to pass the propagation region without spurious reflections at the boundaries

with

$$D(x,y)=\left(-\frac{\partial}{\partial x}+\frac{\partial}{\partial y}\right),\quad \rho(x,y)=\langle x_{\mathrm{R}1}|\frac{1}{x\mathbf{1}_{\mathrm{R}}+\mathrm{i}y\delta\boldsymbol{H}_{\mathrm{RR}}}|q_{\mathrm{R}}\rangle. \tag{96}$$

Next, we use the Dyson equation to find an explicit expression for $\rho(x,y)$. We have

$$\frac{1}{x\mathbf{1}_{\mathrm{R}}+\mathrm{i}y\delta\boldsymbol{H}_{\mathrm{RR}}}|q_{\mathrm{R}}\rangle=\frac{1}{x}|q_{\mathrm{R}}\rangle-\frac{1}{x}\frac{\mathrm{i}y\delta}{x\mathbf{1}_{\mathrm{R}}+\mathrm{i}y\delta\boldsymbol{H}_{\mathrm{RR}}}\boldsymbol{H}_{\mathrm{RR}}|q_{\mathrm{R}}\rangle. \tag{97}$$

It is straightforward to realize that the action of $\boldsymbol{H}_{\mathrm{RR}}$ on $|q_{\mathrm{R}}\rangle$ yields

$$\boldsymbol{H}_{\mathrm{RR}}|q_{\mathrm{R}}\rangle=2V_{\mathrm{R}}\cos(q\Delta x)+h_{\mathrm{R}}|q_{\mathrm{R}}\rangle-V_{\mathrm{R}}\mathrm{e}^{-\mathrm{i}q\Delta x}|x_{\mathrm{R}1}\rangle, \tag{98}$$

so that Eq. (97) can be rewritten as

$$\left[1+\frac{2\mathrm{i}y\delta V_{\mathrm{R}}\cos(q\Delta x)}{x}\right]\frac{1}{x\mathbf{1}_{\mathrm{R}}+\mathrm{i}y\delta\boldsymbol{H}_{\mathrm{RR}}}|q_{\mathrm{R}}\rangle=\frac{1}{x}|q_{\mathrm{R}}\rangle+\frac{1}{x}\frac{\mathrm{i}y\delta V_{\mathrm{R}}\mathrm{e}^{-\mathrm{i}q\Delta x}}{x\mathbf{1}_{\mathrm{R}}+\mathrm{i}y\delta\boldsymbol{H}_{\mathrm{RR}}}|x_{\mathrm{R}1}\rangle. \tag{99}$$

Projecting Eq. (99) on $\langle x_{\mathrm{R}0}|$ we find

$$\left[1+\frac{2\mathrm{i}y\delta V_{\mathrm{R}}\cos(q\Delta x)}{x}\right]\rho(x,y)=\frac{1}{x}+\frac{\mathrm{i}y\delta\mathrm{e}^{-\mathrm{i}q\Delta x}}{xV_{\mathrm{R}}}q_{\mathrm{R}}(x,y), \tag{100}$$

where $q_{\mathrm{R}}(x,y)$ is the generating function defined in Eq. (85). Solving Eq. (100) for $\rho(x,y)$ we conclude that

$$V_{\mathrm{R}}\rho(x,y)=\frac{V_{\mathrm{R}}+\mathrm{i}y\delta\mathrm{e}^{-\mathrm{i}q\Delta x}q_{\mathrm{R}}(x,y)}{x+2\mathrm{i}y\delta V_{\mathrm{R}}\cos(q\Delta x)+h_{\mathrm{R}}}. \tag{101}$$

Using the relation in Eq. (87) for the coefficients $q_{\alpha}^{(m)}$ we find

$$\begin{aligned}\frac{[D(x,y)]^m}{m!}\rho(x,y)\bigg|_{x=y=1}&=\frac{(1-2\mathrm{i}\delta V_{\mathrm{R}}\cos(q\Delta x)+\mathrm{i}\delta h_{\mathrm{R}})^m}{(1+2\mathrm{i}\delta V_{\mathrm{R}}\cos(q\Delta x)+\mathrm{i}\delta h_{\mathrm{R}})^{m+1}}+\frac{\mathrm{i}\delta}{V_{\mathrm{R}}}\mathrm{e}^{-\mathrm{i}q\Delta x}\\&\quad\times\sum_{j=0}^{m}\frac{(1-2\mathrm{i}\delta V_{\mathrm{R}}\cos(q\Delta x)+\mathrm{i}\delta h_{\mathrm{R}})^{m-j}}{(1+2\mathrm{i}\delta V_{\mathrm{R}}\cos(q\Delta x)+\mathrm{i}\delta h_{\mathrm{R}})^{m+1-j}}\left(q_{\mathrm{R}}^{(j)}+q_{\mathrm{R}}^{(j-1)}\right).\end{aligned} \tag{102}$$

One may proceed along the same lines for extracting the left component of the source term.

To test our implementation we have propagated eigenstates of the extended system. As expected, these states just pick up an exponential phase factor $\exp(-\mathrm{i}Et)$ during the propagation.

We are now in a position to apply our algorithm to the calculation of time-dependent currents in one-dimensional model systems. The systems are initially in thermodynamic equilibrium. At time $t=0$, a time-dependent perturbation is switched on. In all the examples below the current is calculated according to Eq. (62):

$$\begin{aligned}I(x,t)&=2\int_{-k_{\mathrm{F}}}^{k_{\mathrm{F}}}\frac{\mathrm{d}k}{2\pi}\,\mathrm{Im}\left(\psi_k^*(x,t)\frac{\mathrm{d}}{\mathrm{d}x}\psi_k(x,t)\right)\\&=2\int_0^{k_{\mathrm{F}}}\frac{\mathrm{d}k}{2\pi}\,\mathrm{Im}\left(\psi_k^*\frac{\mathrm{d}}{\mathrm{d}x}\psi_k+\psi_{-k}^*\frac{\mathrm{d}}{\mathrm{d}x}\psi_{-k}\right)\end{aligned} \tag{103}$$

where the prefactor 2 comes from spin and $k_{\mathrm{F}} = \sqrt{2\varepsilon_{\mathrm{F}}}$ is the Fermi wavevector of a system with Fermi energy ε_{F}.

5.1. DC bias

As a first example we considered a system where the electrostatic potential vanishes identically both in the left and right leads as well as in the central region which is explicitly propagated. Initially, all single particle levels are occupied up to the Fermi energy ε_{F}. At $t = 0$ a constant bias is switched on in the leads and the time-evolution of the system is calculated. We chose the bias in the right lead as the negative of the bias in the left lead, $U_{\mathrm{R}} = -U_{\mathrm{L}}$.

The numerical parameters are as follows: the Fermi energy is $\varepsilon_{\mathrm{F}} = 0.3\,\mathrm{au}$, the bias is $U_{\mathrm{L}} = -U_{\mathrm{R}} = 0.05,\ 0.15,\ 0.25\,\mathrm{au}$, the central region extends from $x = -6$ to $x = +6\,\mathrm{au}$ with equidistant grid points with spacing $\Delta x = 0.03\,\mathrm{au}$. The k-integral in Eq. (103) is discretized with 100 k-points, which amounts to a propagation of 200 states. The time step for the propagation was $\Delta t = 10^{-2}\,\mathrm{au}$.

In Figure 6 we have plotted the current densities at $x = 0$ as a function of time for different values of the applied bias. As a first feature we notice that a steady state is achieved, in agreement with the discussion of Section 3.3. The corresponding steady-state current $I^{(\mathrm{S})}$ can be calculated from the Landauer formula. For the present geometry this leads to the steady current

$$
\begin{aligned}
I^{(\mathrm{S})} = 8e \int_{\max(U_{\mathrm{L}},U_{\mathrm{R}})} & \frac{\mathrm{d}\omega}{2\pi} [f(\omega - U_{\mathrm{L}}) - f(\omega - U_{\mathrm{R}})] \\
\times & \frac{\sqrt{\omega - U_{\mathrm{L}}}\sqrt{\omega - U_{\mathrm{R}}}}{\left[\sqrt{\omega - U_{\mathrm{L}}} + \sqrt{\omega - U_{\mathrm{R}}}\right]^2 + U_{\mathrm{L}} U_{\mathrm{R}} \left[\dfrac{\sin(l\sqrt{2\omega})}{\sqrt{\omega}}\right]^2},
\end{aligned}
\tag{104}
$$

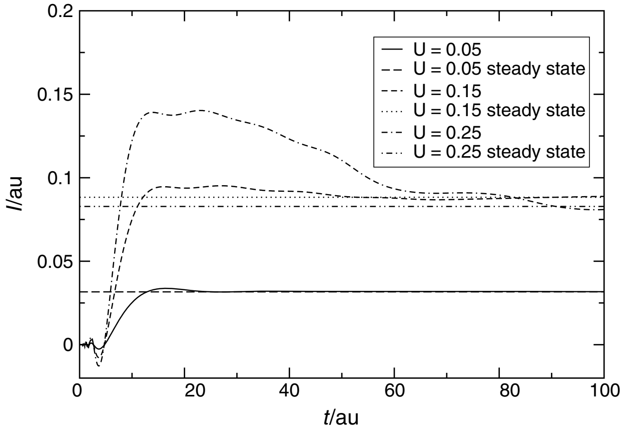


Figure 6 Time evolution of the current for a system where initially the potential is zero in the leads and the propagation region. At $t = 0$, a constant bias with opposite sign in the left and the right leads is switched on, $U = U_{\mathrm{L}} = -U_{\mathrm{R}}$ (values in atomic units). The propagation region extends from $x = -6$ to $x = +6\,\mathrm{au}$. The Fermi energy of the initial state is $\varepsilon_{\mathrm{F}} = 0.3\,\mathrm{au}$. The current in the center of the propagation region is shown

where l is the width of the central region. From Eq. (104) with $l = 12$ au and $U_{\mathrm{L}} = -U_{\mathrm{R}}$, the numerical values for the steady-state currents are 0.0316 au ($U_L = 0.05$ au), 0.0883 au ($U_L = 0.15$ au) and 0.0828 au ($U_L = 0.25$ au). We see that our algorithm yields the same answers. Second, we notice that the onset of the current is delayed in relation to the switching time $t = 0$. This is easily explained by the fact that the perturbation at $t = 0$ happens in the leads only, e.g., for $|x| > 6$ au, while we plot the current at $x = 0$. In other words, we see the delay time needed for the perturbation to propagate from the leads to the center of our device region. We also note that the higher the bias the more the current overshoots its steady-state value for small times after switching on the bias. Finally it is worth mentioning that increasing the bias not necessarily leads to a larger steady-state current.

In the second example we studied a double square potential barrier with electrostatic potential $V(x) = 0.5$ au for $5\,\mathrm{au} \le |x| \le 6$ au and zero otherwise. This time we switch on a constant bias in the left lead only, i.e., $U_R = 0$. The Fermi energy for the initial state is $\varepsilon_{\mathrm{F}} = 0.3$ au. The central region extends from $x = -6$ to $x = +6$ au with a lattice spacing of $\Delta x = 0.03$ au. Again, we use 100 different k-values to compute the current and a time step of $\Delta t = 10^{-2}$ au.

In Figure 7 (Left panel) we plot the current at $x = 0$ as a function of time for several values of the applied bias $U = U_{\mathrm{L}}$. Again, a steady state is achieved for all values of U. As discussed in Figure 6 the transient current can exceed the steady current; the higher the applied voltage the larger is this excess current and the shorter is the time when it reaches its maximum. Furthermore, the oscillatory evolution towards the steady current solution depends on the bias. For high bias, the frequency of the transient oscillations increases. For small bias, the electrons at the bottom of the band are not disturbed and the transient process is exponentially short. On the other hand, for a bias close to the Fermi energy the transient process decays as a power law, due to the band edge singularity. As pointed out in Section 3.3, for non-interacting electrons the steady-state current develops by means of a pure dephasing mechanism. In our examples the transient

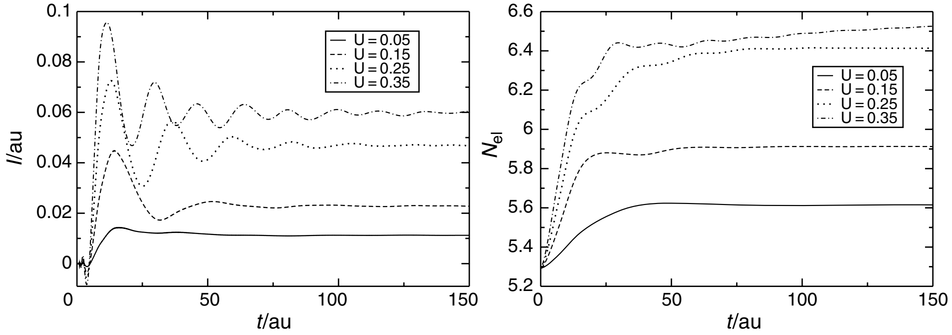


Figure 7 Left panel: Time evolution of the current through a double square potential barrier in response to an applied constant bias (given in atomic units) in the left lead. The potential is given by $V(x) = 0.5$ au for $5 \le |x| \le 6$ au and zero otherwise, the propagation region extends from $x = -6$ to $x = +6$ au. The Fermi energy of the initial state is $\varepsilon_{\mathrm{F}} = 0.3$ au. The current in the center of the structure is shown. Right panel: Time evolution of the total number of electrons in the region $|x| \le 6$ for the same double square potential barrier

process occurs in a femtosecond time-scale, which is much shorter than the relaxation time due to electron–phonon interactions.

In Figure 7 (right panel) we plot the time evolution of the total number of electrons in the device region for the same values of U_L. We see that as a result of the bias a quite substantial amount of charge is added to the device region. This result has important implications when simulating the transport through an interacting system as the effective (dynamical) electronic screening is modified not only due to the external field but also due to the accumulation of charge state in the molecular device. This illustrates that linear response might not be an appropriate tool to tackle the dynamical response and that we will need to resort to a full time-propagation approach as the one presented in this review. Here we emphasize that all our calculations are done without taking into account the electron–electron interaction. If we had done a similar calculation with the interaction incorporated in a time-dependent Hartree or time-dependent DFT framework we would expect the amount of excess charge to be reduced significantly as compared to Figure 7.

5.2. Time-dependent biases

In the previous section we have shown how a steady current develops after the switching on of a constant bias and discussed the transient regime. Here we exploit the versatility of our proposed algorithm for studying different kinds of time-dependent biases.

As a first example we consider how the current responds to a sudden switching off of the bias. For comparison we have considered the same double square potential barrier of Figure 7 subject to the same suddenly switched on bias, but we have turned off the bias at $t = 75$ au. The results (obtained with the same parameters of Figure 7) are displayed in Figure 8. We observe that the current shows a rather well pronounced peak shortly after switching off the perturbation. The amplitude of the peak is proportional to the originally applied bias. This peak always overshoots the value of the current at

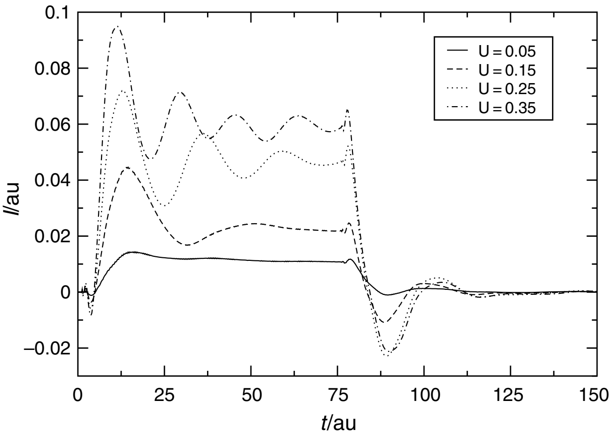


Figure 8 Same system of Figure 7 exposed to a suddenly switched on bias at $t = 0$. The bias is then turned off at $t = 75$ au. The current is measured in the middle of the central region

the steady state. Another interesting feature is the fact that after turning off the bias the transient currents show only two oscillations around the zero current limit and the transient time for switching off is much shorter than for switching on a high bias.

We have also addressed the simulation of AC-transport. We computed the current for a single square potential barrier with $V(x) = 0.6$ for $|x| < 6$ and zero otherwise. Here we applied a time-dependent bias of the form $U_{\mathrm{L}}(t) = U_0 \sin(\omega t)$ to the left lead. The right lead remains at zero bias. The numerical parameters are: Fermi energy $\varepsilon_{\mathrm{F}} = 0.5$ au, device region from $x = -6$ to $x = +6$ au with lattice spacing $\Delta x = 0.03$ au. The number of k-points is 100 and the time step is $\Delta t = 10^{-2}$ au. In Figure 9 we plot the current at $x = 0$ as a function of time for different values of the parameter $U_0 = 0.1, 0.2, 0.3$ au. The frequency was chosen as $\omega = 1.0$ au in both cases. Again, as for the DC-calculation discussed above, we get a transient that overshoots the average current flowing through the constriction; again, this excess current is larger the higher the applied voltage. Also, after the transient we obtain a current through the system with the same period as the applied bias. Note, however, that (especially for the large bias), the current is not a simple harmonic as the applied AC bias.

Exposing the system to an AC bias also allows us to acquire information about the excitation energies of the molecular device. In Figure 10 (left panel) we plot the time-dependent current for a symmetric double square potential barrier in response to a harmonic bias in the left lead, $U_{\mathrm{L}}(t) = U_0 \sin(\omega t)$, with $U_0 = 0.15$ au and $\omega = 0.03$ au. The Fermi energy of the initial state is $\varepsilon_{\mathrm{F}} = 0.3$ au and the current at $x = 0$ is shown. The central region extends from $x = -6$ to $x = 6$ au with lattice spacing $\Delta x = 0.03$ au and the potential $V(x)$ in region C is given by $V(x) = 0$ for $|x|/\mathrm{au} < (6 - d)$ and $V(x) = 0.5$ au for $(6 - d) < |x|/\mathrm{au} < 6$. The number of k-points is 100 and the time step is $\Delta t = 10^{-2}$ au. We have studied barriers of different thickness $d = 1$ au and $d = 2$ au. For $d = 2$ au we observe small oscillations superimposed to the oscillations of frequency $\omega = 0.03$ au driven by the external AC field. Such small oscillations have frequency $\simeq 0.23$ and

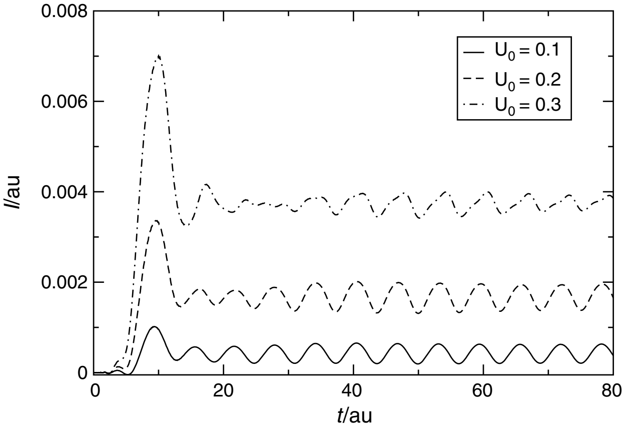


Figure 9 Time evolution of the current for a square potential barrier in response to a time-dependent, harmonic bias in the left lead, $U_{\mathrm{L}}(t) = U_0 \sin(\omega t)$ for different amplitudes U_0 (values in au) and frequency $\omega = 1.0$ au. The potential is given by $V(x) = 0.6$ au for $|x| \leq 6.0$ au and zero otherwise. The propagation region extends from $x = -6$ to $x = +6$ au. The Fermi energy of the initial state is $\varepsilon_{\mathrm{F}} = 0.5$ au. The current at $x = 0$ is shown

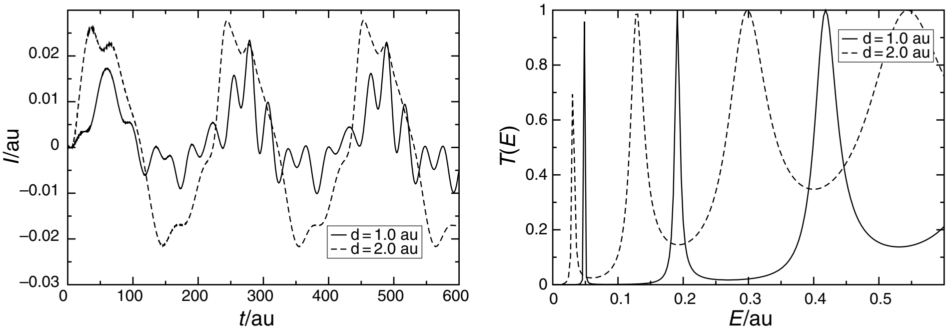


Figure 10 Left panel: Time evolution of the current for a symmetric double square potential barrier in response to a time-dependent, harmonic bias in the left lead, $U_L(t) = U_0 \sin(\omega t)$ with $U_0 = 0.15\,\mathrm{au}$ and $\omega = 0.03\,\mathrm{au}$ for different thickness $d = 1$ and $d = 2\,\mathrm{au}$ of the barriers. Right panel: Transmission function of the same double square potential barrier for $d = 1$ and $d = 2\,\mathrm{au}$

can be understood by looking at the transmission function $T(E)$ in the right panel of Figure 10. For $d = 2\,\mathrm{au}$ both the second and third peaks of $T(E)$ are in the energy window $(\varepsilon_F - U_0, \varepsilon_F + U_0) = (0.15, 0.45)\,\mathrm{au}$. The energy difference between these two peaks corresponds to a good extent to the frequency of the superimposed oscillations. On the contrary, for $d = 1\,\mathrm{au}$ only one peak of the transmission function $T(E)$ is contained in the aforementioned energy window and no superimposed oscillations are clearly visible. This example shows the AC quantum transport can be used also for probing molecular devices.

5.3. History dependence

In Figure 11 we show time-dependent currents for the same double barrier as in Figure 7 for two different ways of applying the bias in the left lead: in one case the constant bias U_0 is switched on suddenly at $t = 0$ (as in Figure 7), in the other case the constant U_0 is achieved with a smooth switching $U(t) = U_0 \sin^2(\omega t)$ for $0 < t < \pi/(2\omega)$. As expected and in agreement with the results of Section 3.3, the same steady state is achieved after the initial transient time. However, the transient current clearly depends on how the bias is switched on.

According to the result in Eq. (39), for noninteracting electrons the independence of the history is not limited to steady-state regimes. The long-time behaviour of currents $I(t)$ and $I'(t)$ induced by biases $U_\alpha(t)$ and $U'_\alpha(t)$ does not change provided $U_\alpha - U'_\alpha \to 0$ as $t \to \infty$. For instance, the current response to an AC bias has the same periodic modulation and the *same phase* independently of how the AC bias is switched on. In Figure 12 we plot the time-dependent current for the same system (and using the same parameters) of Figure 9. The bias remains on zero in the right lead. In the left lead we applied a time-dependent bias of the form $U_L(t) = U_0 f(t) \sin(\omega t)$, with $U_0 = 0.2\,\mathrm{au}$, $\omega = 1.0\,\mathrm{au}$, and we considered two different "switching on" functions $f(t)$. The first is $f(t) = 1$ (as in Figure 9) while the second is a ramp-like switching on $f(t) = \theta(T - t)t/T + \theta(t - T)$ with $T = 30\,\mathrm{au}$. As expected, and in agreement with Eq. (39), the current has the same behaviour in the long-time limit.

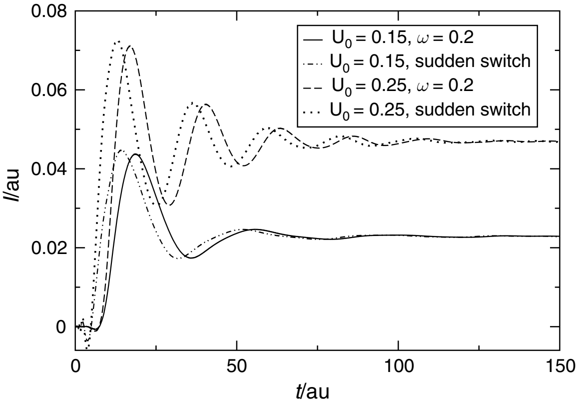


Figure 11 Time evolution of the current for a double square potential barrier when the bias is switched on in two different manners: in one case, the bias U_0 is suddenly switched on at $t = 0$ while in the other case the same bias is achieved with a smooth switching $U(t) = U_0 \sin^2(\omega t)$ for $0 < t < \pi/(2\omega)$. The parameters for the double barrier and the other numerical parameters are the same as the ones used in Figure 7. U_0 and ω given in atomic units

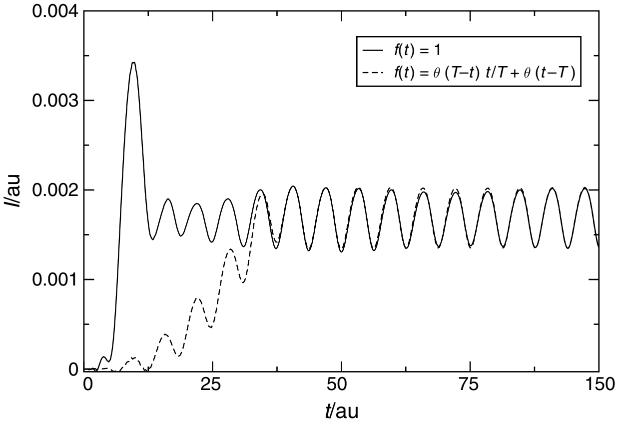


Figure 12 Time evolution of the current for a square potential barrier in response to a time-dependent, harmonic bias in the left lead, $U_L(t) = U_0 f(t) \sin(\omega t)$ with $U_0 = 0.2$ au and frequency $\omega = 1.0$ au. The system and the parameters used are the same as in Figure 9. The current at $x = 0$ is shown for two different "switching on" functions $f(t)$

5.4. Pumping current: Preliminary results

Our algorithm is also well suited to study pumping of electrons. An electron pump is a device which generates a DC current between two electrodes kept at the same bias. The pumping is achieved by applying a periodic gate voltage depending on two or more parameters. Electron pumps have been realized experimentally, e.g., for an open semiconductor quantum dot [31] where pumping was achieved by applying two harmonic gate voltages with a phase shift.

In the literature, different techniques have been used to discuss electron pumping theoretically. Brouwer [32] suggested a scattering approach to describe pumping of noninteracting electrons which has been used, e.g., to study pumping through a double barrier [33]. Nonequilibrium Green's function techniques have been used to study pumping in tight-binding models of coupled quantum dots [34]. Alternatively, Floquet theory which describes evolution of a quantum system under the influence of time-periodic fields is also well suited to describe pumping [35].

As a first example of electron pumping we have calculated the time evolution of the density for a single square barrier exposed to a travelling potential wave $U(t) = U_0 \sin(qx - \omega t)$. The wave is spatially restricted to the explicitly treated device region which in our case also coincides with the static potential barrier. Some snapshots of the density and the potential wave are shown in Figure 13. The density in the device region clearly exhibits local maxima in the potential minima and is transported in pockets by the wave. This is also evident in Figure 14 where we show the time-dependent density as function of both position and time throughout the propagation. The density contour lines show transport of electrons from the left lead at $x = -8$ to the right lead at $x = +8$ au. The pumping mechanism in this example resembles pumping of water with the Archimedean screw.

As a second example we have calculated pumping through a double square barrier by applying two harmonic gate voltages with a phase difference to the barrier potentials, i.e., $U(x, t) = U_0 \sin(\omega t)$ for the left barrier and $U(x, t) = U_0 \sin(\omega t + \phi)$ for the right barrier. Figure 15 shows the DC component of the pump current as a function of the phase ϕ which has a sinusoidal dependence for our parameter values. This is in agreement with similar results of [33] for small amplitudes of the AC bias which were obtained using Brouwer's approach. In addition, this example may be interpreted as a very simple model to describe the experiment of [31].

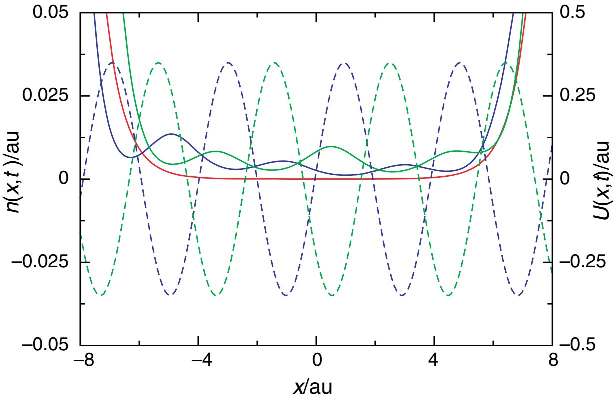


Figure 13 Snapshots of the density for and the travelling potential wave at various times for pumping through a single square barrier by a travelling wave. The barrier with height 0.5 au extends throughout the propagation window from $x = -8$ to $x = +8$. The leads are on zero potential and the Fermi level is at 0.3 au. The travelling potential wave is restricted to the propagation window $|x| < 8$ and has the form $U(t) = U_0 \sin(qx - \omega t)$ with amplitude $U_0 = 0.35$ au, wave number $q = 1.6$ au and frequency $\omega = 0.2$ au. The initial density is given by the red line

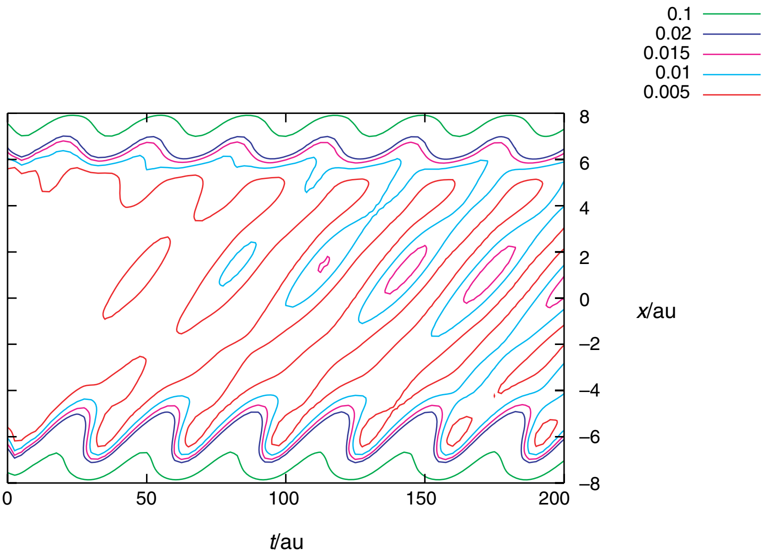


Figure 14 Contour plot of the time-dependent density for pumping through a single square barrier by a traveling potential wave. The parameters are the same as for Figure 13

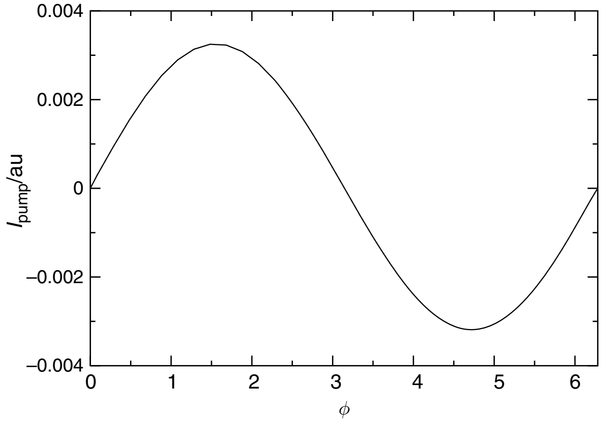


Figure 15 Parametric pumping through a double square barrier. The device region extends from $x = -6$ to $x = +6\,\mathrm{au}$, the static potential has the value 0.525 a.u. for $5 < |x| < 6\,\mathrm{au}$ and zero elsewhere in the device. Pumping is achieved by harmonic variation of the barriers, i.e., $U(x,t) = U_0 \sin(\omega t)$ for the left barrier ($-6\,\mathrm{au} < x < -5\,\mathrm{au}$) and $U(x,t) = U_0 \sin(\omega t + \phi)$ for the right barrier ($5\,\mathrm{au} < x < 6\,\mathrm{au}$). The DC component of the pump current is displayed as a function of the phase ϕ. The parameters are: $U_0 = 0.25\,\mathrm{au}$, $\omega = 0.25\,\mathrm{au}$ and the Fermi energy is $\varepsilon_{\mathrm{F}} = 0.5\,\mathrm{au}$

6. Conclusions and perspectives

In this chapter we have given a self-contained introduction to our recent approach to quantum transport. In essence our approach combines two well-established theories for the description of nonequilibrium phenomena of interacting many-electron systems.

On the one hand there is the formalism of non-equilibrium Keldysh–Green functions. Although this approach, in principle, can be used to study interaction effects, here we only used it in the context of noninteracting electrons. The reason for this is that the self-energy of interacting electrons (which is not to be confused with the embedding self-energy) is long-range and nonlocal. In our scheme which partitions space into left and right leads as well as the central device region, this nonlocality is extremely difficult to deal with in a rigorous manner.

On the other hand, the NEGF formalism for (effectively) noninteracting electrons can easily be combined with the second approach for time-dependent many-particle systems, namely time-dependent density functional theory. Just as the NEGF formalism, TDDFT, in principle, gives the correct time-dependent density of the interacting system (if the exact exchange-correlation potential is used). Moreover, the time-dependent effective single-particle potential of TDDFT is a local and multiplicative potential which is crucial for practical use within the partitioning scheme for transport.

In combining the NEGF and TDDFT approaches we have presented a formally rigorous approach towards the description of charge transport using an open-boundary scheme within TDDFT. We have implemented a specific time-propagation scheme that incorporates transparent boundaries at the device/lead interface in a natural way. In order to have a clear definition of a device region, in Figure 1 we assumed that an applied bias can be described by adding a spatially constant potential shift in the macroscopic part of the leads. This implies an effective “metallic screening” of the constriction. The screening limits the spatial extent of the induced density created by the bias potential or the external field to the central region. Our time-dependent scheme allows to extract both AC and DC I/V device characteristics and it is ideally suited to describe external field (photon) assisted processes.

In order to illustrate the performance and potential of the method we have implemented it for one-dimensional model systems and applied it to a variety of transport situations: we have shown that a steady-state current is always reached upon application of a DC bias. For a harmonic AC bias, the resulting AC current need not be harmonic. In the case of systems at DC bias without any source of dissipation it is known that the steady current is independent of the history of the process [8]. We have explicitly demonstrated this history independence for two different switching processes of the external bias. The history independence for noninteracting electrons not only applies for DC but also for AC bias, which we have also demonstrated in a numerical example. Since we can compute current densities locally, we are not restricted to currents deep inside the leads. In one example we have analyzed the time evolution of the density for localized states which are only weakly coupled to the reservoirs. Finally, we have shown a few simple applications of our algorithm to electron pumping.

The list of the example calculations presented here already demonstrates the versatility and flexibility of our algorithm. It includes the Landauer formalism as the long-time limit for systems under DC bias and allows to study transients. Moreover, it can deal with periodic time-dependent fields (which are usually treated with the Floquet formalism) but is applicable to nonperiodic conditions as well [36].

Most theoretical approaches to transport adopt open boundary conditions and assume that transport is ballistic, i.e., under steady-state conditions inelastic collisions are absent and dissipation occurs only in the idealized macroscopic reservoirs. This might be an unrealistic assumption for transport through single molecules, in particular when the device is not operated in the regime of small bias and linear response. When inelastic scattering dominates, this picture is not applicable. In particular, experiments [37–39] indicate that inelastic scattering with lattice vibrations is present at sufficiently large bias, causing local heating of contacts and molecular devices. In addition, current-induced forces might lead to bond-breaking and electromigrations.

In a joint collaboration with Verdozzi and Almbladh, one of us has included the nuclear degrees of freedom at a classical level [40]. The initial ground state (consisting of bound, resonant and scattering states) has been calculated self-consistently. Also, the time-propagation algorithm of Section 4.2 has been generalized to evolve the system electrons + nuclei in the Ehrenfest approximation. Several aspects of the electron–ion interaction in quantum transport have been investigated.

Electron correlations are also important in molecular conductors, for example, Coulomb blockade effects dominate the transport in quantum dots. Short-range electron correlations seem to be relevant in order to get quantitative description of I/V characteristics in molecular constrictions [41–43]. In particular it is commonly assumed that the energy scales for electron–electron and electron–phonon interactions are different and could be treated separately. However, the metallic screening of the electrodes considerably reduces the Coulomb-charging energy (from eV to meV scale). In this regime the energy scales for the two interactions merge and they need to be treated on the same footing. We would like to emphasize that our scheme allows for a consistent treatment of electronic and ionic degrees of freedom.

It is clear that the quality of the TDDFT functionals is of crucial importance. In particular, exchange and correlation functionals for the nonequilibrium situation are required. Time-dependent linear response theory for DC-steady state has been implemented in [44] within TDLDA assuming jellium-like electrodes (mimicked by complex absorbing/emitting potentials). It has been shown that the DC-conductance changes considerably from the standard Landauer value. Therefore, a systematic study of the TDDFT functionals themselves is needed. A step beyond standard adiabatic approximations and exchange-only potentials is to resort to many-body schemes based on perturbative expansions [45, 46], iterative schemes [47], or variational-functional formulations [48]. Another path is to explore exchange-correlation functionals that depend implicitly [25, 49] or explicitly [50, 51] on the current density.

Acknowledgments

This work was supported in part by the Deutsche Forschungsgemeinschaft, by the EU Research and Training Network EXCITING and by the NANOQUANTA Network of Excellence.

References

[1] S. Datta, *Electronic Transport in Mesoscopic Systems* (Cambridge University Press, Cambridge, 1995).

[2] H. Haug and A.-P. Jauho, *Quantum Kinetics in Transport and Optics of Semiconductors* (Springer, Berlin, 1998).

[3] G. Cuniberti, G. Fagas, K. Richter (eds.), *Introducing Molecular Electronics*, Lecture Notes in Physics, Vol. 680 (Springer, Berlin, 2005).

[4] M. Cini, *Phys. Rev. B* **22**, 5887 (1980).

[5] E. Runge and E. K. U. Gross, *Phys. Rev. Lett.* **52**, 997 (1984).

[6] G. Stefanucci and C.-O. Almbladh, *Europhys. Lett.* **67**, 14 (2004).

[7] N. D. Lang, *Phys. Rev. B* **52**, 5335 (1995).

[8] G. Stefanucci and C.-O. Almbladh, *Phys. Rev. B* **69**, 195318 (2004).

[9] M. Petersilka, U. J. Gossmann and E. K. U. Gross, *Phys. Rev. Lett.* **76**, 1212 (1996).

[10] S. Kurth, G. Stefanucci, C.-O. Almbladh, A. Rubio and E. K. U. Gross, *Phys. Rev. B* **72**, 035308 (2005).

[11] L. P. Kadanoff and G. Baym, *Quantum Statistical Mechanics* (Benjamin, New York, 1962).

[12] L. V. Keldysh, *JETP* **20**, 1018 (1965).

[13] P. Danielewicz, *Ann. Physics* **152**, 239 (1984) and references therein.

[14] R. Kubo, *J. Phys. Soc. Jpn.* **12**, 570 (1957).

[15] P. C. Martin and J. Schwinger, *Phys. Rev.* **155**, 1342 (1959).

[16] R. van Leeuwen, N. E. Dahlen, G. Stefanucci, C.-O. Almbladh, and U. von Barth, *Introduction to the Keldysh Formalism and Applications to Time-Dependent Density Functional Theory*, Lectures Notes in Physics, Springer Verlag, 2006; cond-mat/0506130.

[17] R. van Leeuwen, *Phys. Rev. Lett.* **80**, 1280 (1998).

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

[19] N. D. Mermin, *Phys. Rev.* **137**, A1441 (1965).

[20] M. Di Ventra and T. N. Todorov, *J. Phys. C* **16**, 8025 (2004).

[21] A. Blandin, A. Nourtier, and D. W. Hone, *J. Phys. (Paris)* **37**, 369 (1976).

[22] R. Landauer, IBM *J. Res. Develop.* **1**, 233 (1957).

[23] M. Koentopp, K. Burke, and F. Evers, *Phys. Rev. B* **73**, 121403(R) (2006).

[24] J. R. Hellums and W. R. Frensley, *Phys. Rev. B* **49**, 2904 (1994).

[25] J. F. Dobson, M. J. Bünner and E. K. U. Gross, *Phys. Rev. Lett.* **79**, 1905 (1997).

[26] R. van Leeuwen, *Phys. Rev. Lett.* **82**, 3863 (1999).

[27] N. T. Maitra, K. Burke and C. Woodward, *Phys. Rev. Lett.* **89**, 023002 (2002).

[28] C. Moyer, *Am. J. Phys.* **72**, 351 (2004); and references therein.

[29] Higher order propagation schemes could be developed following the ideas discussed here for the modification of the Cayley method. Such schemes could also handle self-consistent and time-dependent Hamiltonians as the ones appearing in standard TDDFT (see A. Castro, M. A. L. Marques and A. Rubio, *J. Chem. Phys.* **121**, 3425 (2004) and references therein).

[30] H. Appel, L. Wirtz, G. Stefanucci, C.-O. Almbladh, S. Kurth, E. K. U. Gross and A. Rubio (work in progress).

[31] M. Switkes, C. M. Marcus, and A. C. Gossard, *Science* **283**, 1905 (1999).

[32] P. W. Brouwer, *Phys. Rev. B* **58**, R10135 (1998).

[33] Y. Wei, J. Wang, and H. Guo, *Phys. Rev. B* **62**, 9947 (2000).

[34] C. A. Stafford and N. S. Wingreen, *Phys. Rev. Lett.* **76**, 1916 (1996).

[35] S. Kohler, J. Lehmann, and P. Hänggi, *Phys. Rep.* **406**, 379 (2005).

[36] X. Oriols, A. Alarcón, and E. Fernàndez-Díaz, *Phys. Rev. B* **71**, 245322 (2005).

[37] J. G. Kushmerick, J. Lazorcik, C. H. Patterson, R. Shashidhar, D. S. Seferos and G. C. Bazan, *Nano Lett.* **4**, 639 (2004).

[38] R. H. M. Smit, C. Untiedt and J. M. van Ruitenbeek, *Nanotechnology* **15**, S472 (2004).

[39] B. J. LeRoy, S. G. Lemay, J. Kong and C. Dekker, *Nature* **432**, 371 (2004).
[40] C. Verdozzi, G. Stefanucci, and C.-O. Almbladh, *Phys. Rev. Lett.* **97**, 046603 (2006).
[41] P. Delaney and J.C. Greer, *Phys. Rev. Lett*. **93**, 036805 (2004).
[42] F. Evers, F. Weigend and M. Koentopp, *Phys. Rev. B* **69**, 235411 (2004).
[43] A. Ferretti, A. Calzolari, R. Di Felice, F. Manghi, M. J. Caldas, M. B. Nardelli and E. Molinari, *Phys. Rev. Lett*. **94**, 116802 (2005).
[44] R. Baer, T. Seideman, S. Ilani and D. Neuhauser, *J. Chem. Phys*. **120**, 3387 (2004).
[45] A. Marini, R. Del Sole and A. Rubio, *Phys. Rev. Lett*. **91**, 256402 (2003); L. Reining, V. Olevano, A. Rubio and G. Onida, Phys. Rev. Lett. **88**, 066404 (2002).
[46] I. V. Tokatly and O. Pankratov, *Phys. Rev. Lett*. **86**, 2078 (2001).
[47] F. Bruneval, F. Sottile, V. Olevano, R. Del Sole, and L. Reining, *Phys. Rev. Lett*. **94**, 186402 (2005).
[48] U. von Barth, N. E. Dahlen, R. van Leeuwen and G. Stefanucci, *Phys. Rev. B* **72**, 235109 (2005).
[49] Y. Kurzweil and R. Baer, *J. Chem. Phys*. **121**, 8731 (2004).
[50] C. A. Ullrich and G. Vignale, *Phys. Rev*. B **65**, 245102 (2002) and references therein.
[51] I. V. Tokatly and O. Pankratov, *Phys. Rev*. B **67**, 201103(R) (2003).

Index