

# Numerical Study of Quantum Transport in Carbon Nanotube–Based Transistors

M. Pourfath, H. Kosina

*Institute for Microelectronics, TU Wien, Gußhausstraße 27–29/E360, A-1040 Wien, Austria*

## CONTENTS

- 1. Introduction
- 2. Fundamentals of CNTs
- 3. Quantum Transport Models
- 4. Implementation
- 5. Application
- 6. Summary and Conclusions
- Appendices
- References

## 1. INTRODUCTION

The increasing demand for higher computing power, smaller dimensions, and lower power consumption of integrated circuits leads to a pressing need to downscale semiconductor components. However, downscaling of conventional MOSFETs leads to many problems, such as short-channel effects and increased gate-leakage current. Therefore, novel structures and materials such as multiple-gate MOSFETs, CNT-FETs, and molecular-based transistors are expected to be introduced to meet the requirements for scaling.

Since the discovery of carbon nanotubes (CNTs) by Iijima in 1991 [1], significant progress has been achieved in both understanding the fundamental properties and exploring possible engineering applications. The possible application for nanoelectronic devices has been extensively explored since the demonstration of the first CNT transistors (CNT-FETs) [2, 3].

CNTs are attractive for nanoelectronic applications owing to their excellent electrical properties. The phase space for scattering is severely reduced owing to the one-dimensional nature of the density of states. The low-scattering probability is responsible for high on-current in semiconducting CNT transistors. Owing to the chemical stability and perfection of the CNT structure, carrier mobility is not affected by processing and roughness scattering as it is in the conventional semiconductor channel. The fact that there are no

dangling bond states at the surface of CNTs allows for a wide choice of gate insulators. This improves gate control while meeting gate leakage constraints. The purely one-dimensional transport properties of single-wall CNTs (SWNTs) should lead to a suppression of short-channel effects in transistor devices [4]. Furthermore, the conduction and valence bands are symmetric, which is advantageous for complementary applications, and finally, the combined impact of transport and electrostatic benefits together with the fact that semiconducting CNTs are, unlike silicon, direct-gap materials and suggests applications in optoelectronics as well [5, 6].

Section 2 describes the fundamentals of CNTs. It presents a comprehensive overview of electron and phonon properties along with electron–phonon interaction parameters, which determine transport phenomena in CNTs. The operation of these devices can be explained in terms of SBs which are formed at the metal–CNT interfaces. CNT-FETs can operate by modulating the transmission coefficient through these barriers, which results in device characteristics different from that of conventional MOSFETs.

Section 3 outlines the theory of the nonequilibrium Green's function (NEGF) formalism. Knowledge of the single-particle Green's function provides both the complete equilibrium or nonequilibrium properties of the system and the excitation energies of the systems containing one more or one less particle. The many-particle information about the system is cast into self-energies, parts of the equations of motion for Green's functions. Green's functions can be expressed as a perturbation expansion, which is the key to approximate the self-energies. Green's functions provide a very powerful technique for evaluating properties of many-particle systems both in thermodynamic equilibrium and nonequilibrium situations. This formalism has been successfully used to investigate the characteristics of nanoscale transistors [7, 8], CNT-FETs [4, 9], and molecular transistors [10].

Section 4 discusses the numerical implementation of the NEGF formalism to study quantum transport in CNT-FETs. The discretization of the transport equations in both the spatial and energy domain are discussed in detail. We employed a tight-binding Hamiltonian and applied a mode-space transformation to reduce the computational cost.

The calculation of self-energies owing to electron-phonon interactions is also presented. Finally, the iterative method for self-consistent simulation and its convergence rate is studied.

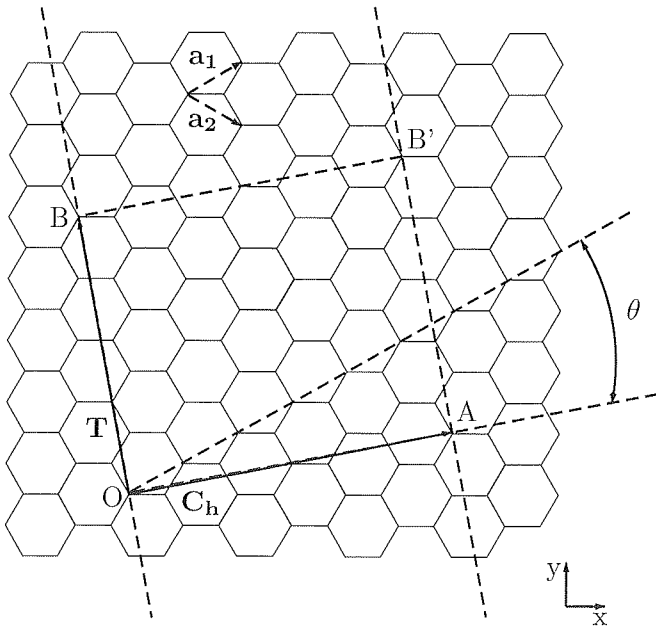
By means of the described methodology, the physics of CNT-FET is explored in Section 5. A comprehensive study of the role of electron-phonon interaction on the performance of CNT-FETs is presented.

## 2. FUNDAMENTALS OF CNTS

CNTs are unique nanostructures that can be considered conceptually a prototype one-dimensional quantum wire. The fundamental building block of CNTs is the very long, all-carbon cylindrical SWNT, one atom in wall thickness and tens of atoms around the circumference (typical diameter  $\sim 1.4$  nm). Initially, CNTs gained great interest in research community because of their exotic electronic properties, and this interest continued as other remarkable properties were discovered and promise of practical applications developed. In this Section, some basic definitions relevant to the structural properties of CNTs are provided, and applications of CNTs in electronics, especially CNT-based transistors, are discussed.

### 2.1. Theoretical Background

The structure of CNTs has been explored early after their discovery by high-resolution transmission electron microscopy techniques yielding direct confirmation that the CNTs are seamless cylinders derived from the honeycomb lattice representing a single atomic layer of crystalline graphite, called a *graphene sheet*. The structure of a SWNT is conveniently explained in terms of its one-dimensional unit cell, defined by the vectors  $\mathbf{C}_h$  and  $\mathbf{T}$  as shown in Figure 1.



**Figure 1.** The chiral vector  $\mathbf{C}_h = n\mathbf{a}_1 + m\mathbf{a}_2$  is defined on the honeycomb lattice of carbon atoms by unit vectors  $\mathbf{a}_1$  and  $\mathbf{a}_2$  and the chiral angle  $\theta$  with respect to the zig-zag axis ( $\theta = 0$ ). The diagram is constructed for  $(n, m) = (4, 2)$ .

The circumference of any CNT is expressed in terms of the chiral vector  $\mathbf{C}_h = n\mathbf{a}_1 + m\mathbf{a}_2$  which connects two crystallographically equivalent sites on a two-dimensional graphene sheet [11]. The construction in Figure 1 depends uniquely on the pair of integers  $(n, m)$  which specify the chiral vector. The chiral angle  $\theta$  is defined as the angle between the chiral vector  $\mathbf{C}_h$  and the zig-zag direction ( $\theta = 0$ ). Three distinct types of CNT structures can be generated by rolling up the graphene sheet into a cylinder. The zig-zag and armchair CNTs correspond to chiral angles of  $\theta = 0$  and  $\theta = 30^\circ$ , respectively, and chiral CNTs correspond to  $0 < \theta < 30^\circ$ . The intersection of the vector  $\overline{OB}$  (which is normal to  $\mathbf{C}_h$ ) with the first lattice point determines the fundamental one-dimensional translation vector  $\mathbf{T}$ . The unit cell of the one-dimensional lattice is the rectangle defined by the vectors  $\mathbf{C}_h$  and  $\mathbf{T}$ .

The cylinder connecting the two hemispherical caps of the CNT is formed by superimposing the two ends of the vector  $\mathbf{C}_h$ , and the cylinder joint is made along the two lines  $\overline{OB}$  and  $\overline{AB'}$  in Figure 1. The lines  $\overline{OB}$  and  $\overline{AB'}$  are both perpendicular to the vector  $\mathbf{C}_h$  at each end of  $\mathbf{C}_h$  [11]. In the  $(n, m)$  notation, for  $\mathbf{C}_h = n\mathbf{a}_1 + m\mathbf{a}_2$ , the vectors  $(n, 0)$  or  $(0, n)$  denote zig-zag CNTs, whereas the vectors  $(n, m)$  correspond to chiral CNTs [12]. The CNT diameter  $d_{\text{CNT}}$  is given by

$$d_{\text{CNT}} = \frac{|\mathbf{C}_h|}{\pi} = \frac{\sqrt{3}a_{\text{C-C}}\sqrt{m^2+n^2+mn}}{\pi}, \quad (1)$$

where  $|\mathbf{C}_h|$  is the length of  $\mathbf{C}_h$  and  $a_{\text{C-C}}$  is the C-C bond length (1.42 Å). The chiral angle  $\theta$  is given by  $\theta = \tan^{-1}[\sqrt{3}n/(2m+n)]$ . For the  $(n, n)$  armchair CNT  $\theta = 30^\circ$  and for the  $(n, 0)$  zig-zag CNT  $\theta = 0^\circ$ . From Figure 1 it follows that if one limits  $\theta$  to the range  $0 \leq \theta \leq 30^\circ$ , then by symmetry,  $\theta = 0$  for a zig-zag CNT. Both armchair and zig-zag CNTs have a mirror plane and thus are considered achiral. Differences in the CNT diameter  $d_{\text{CNT}}$  and chiral angle  $\theta$  give rise to different properties of the various CNTs. The number  $N$  of hexagons per unit cell of a CNT, specified by integers  $(n, m)$ , is given by

$$N = \frac{2(m^2 + n^2 + nm)}{d_R} \quad (2)$$

where  $d_R = d$  if  $n-m$  is not a multiple of  $3d$ , and  $d_R = 3d$  if  $n-m$  is a multiple of  $3d$ , and  $d$  is defined as the greatest common divisor (gcd) of  $(n, m)$ . Each hexagon in the honeycomb lattice contains two carbon atoms. The unit cell area of the CNT is  $N$  times larger than that for a graphene layer and consequently the unit cell area for the CNT in reciprocal space is correspondingly  $1/N$  times smaller. Table 1 provides a summary of relations useful for describing the structure of SWNTs [13, 14].

### 2.2. Electronic Structure

The energy-dispersion relations of SWNTs can be calculated using zone folding [15, 16], the tight-binding method

**Table 1.** Structural properties for CNTs [13].

Symbol	Description	Formula
$a$	Length of unit vectors	$a = \sqrt{3}a_{\text{C-C}} = 2.49 \text{ \AA}$ ,
$a_{\text{C-C}} = 1.42 \text{ \AA}$ $\mathbf{a}_1, \mathbf{a}_2$	Unit vectors	$\left(\frac{\sqrt{3}}{2}, \frac{1}{2}\right)a, \left(\frac{\sqrt{3}}{2}, -\frac{1}{2}\right)a$
$\mathbf{b}_1, \mathbf{b}_2$	Reciprocal lattice vectors	$\left(\frac{1}{\sqrt{3}}, 1\right)\frac{2\pi}{a}, \left(\frac{1}{\sqrt{3}}, -1\right)\frac{2\pi}{a}$
$\mathbf{C}_h$ ( $0 \leq  m  \leq n$ )	Chiral vector	$\mathbf{C}_h = n\mathbf{a}_1 + m\mathbf{a}_2 \equiv (n, m)$ ,
$d_{\text{CNT}}$	Diameter	$d_{\text{CNT}} =  \mathbf{C}_h /\pi = a\sqrt{n^2 + m^2 + nm}$
$\theta$	Chiral angle	$\tan(\theta) = \sqrt{3}/(2n + m)$
$d_R$	$\gcd(2n + m, 2m + n)$	
$\mathbf{T}$	Translational vector	$\mathbf{T} = t_1\mathbf{a}_1 + t_2\mathbf{a}_2, \quad t_1 = \frac{2m + n}{d_R}, \quad t_2 = -\frac{2n + m}{d_R}$
$T$	Length of $\mathbf{T}$	$T =  \mathbf{T}  = \sqrt{3}L/d_R$
$N$	Number of hexagons in the unit-cell	$N = 2(n^2 + m^2 + nm)/d_R$

[17], and density functional theory [18, 19]. In the simplest method the energy-dispersion relations of CNTs are obtained by folding those of graphene.

### 2.2.1. Electronic Band Structure of Graphene

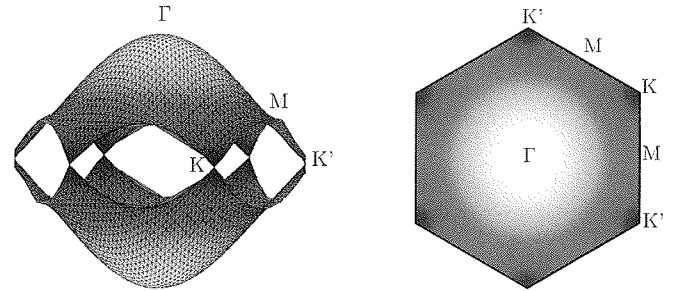
Within the tight-binding method, the two-dimensional energy-dispersion relations of graphene can be calculated by solving the eigenvalue problem for a Hamiltonian  $H_{\text{g-2D}}$  associated with the two carbon atoms in the graphene unit cell [13]. In the Slater-Koster scheme one gets (We consider only the valence and the conduction  $\pi^*$  energy band of graphene and CNTs.)

$$H_{\text{g-2D}} = \begin{bmatrix} 0 & f(k) \\ -f^\dagger(k) & 0 \end{bmatrix}. \quad (3)$$

where  $f(k) = -t(1 + e^{ik \cdot \mathbf{a}_1} + e^{ik \cdot \mathbf{a}_2}) = -t(1 + 2e^{\sqrt{3}k_x a/2} \cos(k_y a/2))$  and  $t$  is the nearest-neighbor C-C tight-binding overlap energy (Experimentally the value  $|t| = 2.7\text{eV}$  has been reported [20].) [16]. Solution of the secular equation  $(H_{\text{g-2D}} - EI) = 0$  leads to

$$E_{\text{g-2D}}^\pm(\mathbf{k}) = \pm t \sqrt{1 + 4\cos\left(\frac{\sqrt{3}k_x a}{2}\right)\cos\left(\frac{k_y a}{2}\right) + \cos^2\left(\frac{k_y a}{2}\right)}, \quad (4)$$

where the  $E_{\text{g-2D}}^+$  and  $E_{\text{g-2D}}^-$  correspond to the  $\pi^*$  and the  $\pi$  energy bands, respectively. Figure 2 shows the electronic energy-dispersion relations for graphene as a function of the



**Figure 2.** The energy-dispersion relations for graphene are shown through the whole region of the Brillouin zone. The lower and the upper surfaces denote the valence  $\pi$  and the conduction  $\pi^*$  energy bands, respectively. The coordinates of high symmetry points are  $\Gamma = (0, 0)$ ,  $K = (0, 2\pi/3a)$ , and  $M = (2\pi/3a, 0)$ . The energy values at the  $K$ ,  $M$ , and  $\Gamma$  points are  $0$ ,  $t$ , and  $3t$ , respectively.

two-dimensional wave vector  $\mathbf{k}$  in the hexagonal Brillouin zone.

### 2.2.2. Electronic Band Structure of SWNTs

The electronic structure of an SWNT can be obtained from that of graphene. Assuming periodic boundary conditions in the circumferential direction characterized by the chiral vector  $\mathbf{C}_h$ , the wave vector associated with the  $\mathbf{C}_h$  direction becomes quantized, while the wave vector associated with the direction of the translational vector  $\mathbf{T}$  (along the CNT axis) remains continuous for a CNT of infinite length. Thus, the energy bands consist of a set of one-dimensional energy-dispersion relations which are cross sections of those of graphene. Expressions for the reciprocal lattice

vectors,  $\mathbf{K}_2$  along the CNT axis, ( $\mathbf{C}_h \cdot \mathbf{K}_2 = 0$ ,  $\mathbf{T} \cdot \mathbf{K}_2 = 2\pi$ ), and  $\mathbf{K}_1$  in the circumferential direction, ( $\mathbf{C}_h \cdot \mathbf{K}_2 = 0$ ,  $\mathbf{T} \cdot \mathbf{K}_2 = 2\pi$ ) are given by (see Table 1)

$$\mathbf{K}_1 = \frac{1}{N}(-t_2 \mathbf{b}_1 + t_1 \mathbf{b}_2), \quad \mathbf{K}_2 = \frac{1}{N}(m \mathbf{b}_1 - n \mathbf{b}_2). \quad (5)$$

The one-dimensional energy-dispersion relations of an SWNT can be written as

$$E_{\text{CNT}}^v(k) = E_{g-2D} \left( k \frac{\mathbf{K}_2}{|\mathbf{K}_2|} + v \mathbf{K}_1 \right), \quad (6)$$

where  $-\pi/T < k < \pi/T$  is a one-dimensional wave vector along the CNT axis and  $v = 1, \dots, N$ . The periodic boundary condition for a CNT gives  $N$  discrete  $k$  values in the circumferential direction. The  $N$  pairs of energy-dispersion curves given by (6) correspond to the cross sections of the two-dimensional energy-dispersion surface of graphene. If the cutting line passes through a  $\mathbf{K}$  point of the two-dimensional Brillouin zone, where the  $\pi$  and  $\pi^*$  energy bands of graphene are degenerated by symmetry, then the one-dimensional energy bands have a zero energy gap. When the  $\mathbf{K}$ -point is located between two cutting lines,  $\mathbf{K}$  is always located in a position one-third of the distance between the two adjacent  $\mathbf{K}_1$  lines [16], and thus a semi-conducting CNT with a finite energy gap is formed. If for a  $(n, m)$  CNT,  $n-m$  is exactly divisible by 3 the CNT is metallic. CNTs with residuals 1 and 2 of the division  $n-m$  by 3 are semiconducting.

Figure 3 shows the energy-dispersion relations for the (5, 5) armchair, the (9, 0) zigzag, and the (10, 0) zigzag CNTs. In general  $(n, n)$  armchair CNTs yield  $4n$  energy subbands with  $2n$  conduction and  $2n$  valence bands. Of these  $2n$  bands, two are nondegenerated and  $n-1$  are doubly degenerated. The degeneracy comes from the two subbands with the same energy dispersion, but different  $v$  values. All armchair CNTs have a band degeneracy between the highest valence and the lowest conduction

band (Fig. 3(a)). In zig-zag CNTs the lowest conduction and the highest valence bands are doubly degenerated (Fig. 3(b) and Fig. 3(c)).

In armchair and zig-zag CNTs, the bands are symmetric with respect to  $k = 0$ . Since the band of an armchair CNT has a minimum at point  $k = 2\pi/3a$ , it has a mirror minimum at point  $k = -2\pi/3a$  and therefore two equivalent valleys are present around the point  $\pm 2\pi/3a$ . The bands of zig-zag and chiral CNTs can have at most one valley (Fig. 3(b) and Fig. 3(c)).

In armchair CNTs, the bands cross the Fermi level at  $k = \pm 2\pi/3a$ . Thus, they are expected to exhibit metallic conduction [13]. There is no energy gap for the (9, 0) CNT at  $k = 0$ , whereas the (10, 0) CNT indeed shows an energy gap.

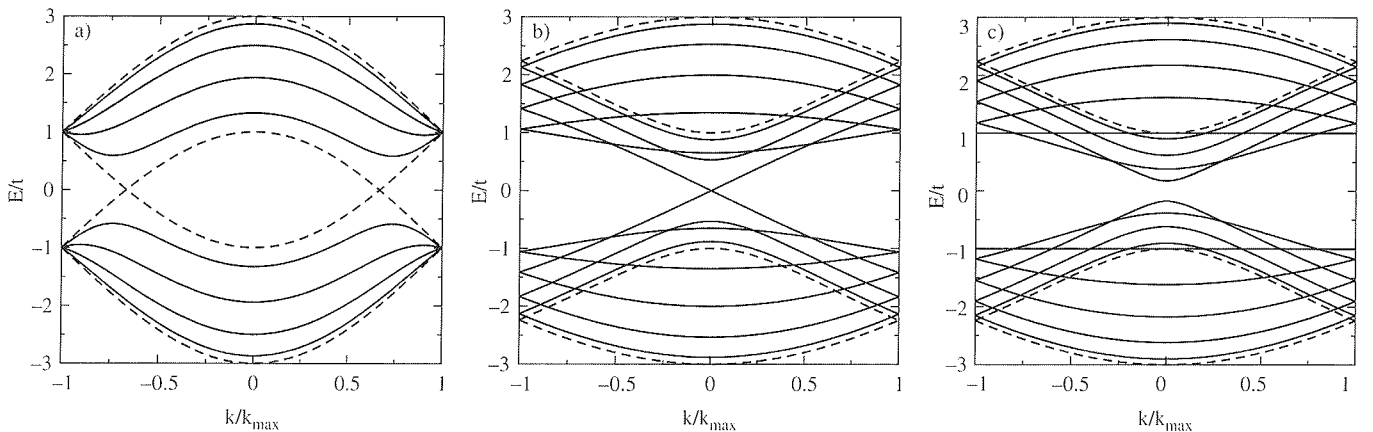
Electrical conduction is determined by states around the Fermi energy. Therefore, it is useful to develop an approximate relation that describes the dispersion relations in the regions around the Fermi energy  $E_F = 0$ . This can be done by replacing the expression for  $f(k) = -t \left( 1 + 2e^{\sqrt{3}k_x a/2} \cos(k_y a/2) \right)$  in (3) with a Taylor expansion around the point  $(0, \pm 4\pi/3a)$ , where the energy gap is zero and  $f(k) = 0$ . It is straightforward to show that  $f(k) \approx (i\sqrt{3}at/2)(k_x \mp i\beta_y)$ , with  $\beta_y \equiv k_y \mp (4\pi/3a)$ . The corresponding energy-dispersion relation can be written as [21]

$$E_{g-2D}(\mathbf{k}) = \pm |f(k)| = \pm \frac{\sqrt{3}at}{2} \sqrt{k_x^2 + \beta_y^2} \quad (7)$$

The energy bands for  $(n, 0)$  zig-zag CNTs can be obtained by imposing the periodic boundary conditions, which define the number of allowed wave vectors  $k_y$  in the circumferential direction as  $nk_y a = 2\pi v$ , ( $v = 1, \dots, 2n$ ). This yields the one-dimensional dispersion relations for the  $4n$  states of the  $(n, 0)$  zig-zag CNT

$$E^v(k_x) = \pm \frac{\sqrt{3}at}{2} \sqrt{k_x^2 + \left[ \frac{4\pi}{3a} \left( \frac{3v}{2n} - 1 \right) \right]^2}, \quad (8)$$

$$-\pi/\sqrt{3}a < k_x < \pi/\sqrt{3}a.$$



**Figure 3.** One-dimensional energy-dispersion relations of (a) the (5, 5) armchair CNT, (b) the (9, 0) zig-zag CNT, and (c) the (10, 0) zig-zag CNT.  $k_{\text{max}}$  for armchair and zig-zag CNTs correspond to  $k_{\text{max}} = \pi/a$  and  $k_{\text{max}} = \pi/\sqrt{3}a$ , respectively. Solid lines denote degenerate bands and dashed lines nondegenerate bands.

Therefore, the energy gap for subband  $\nu$  can be written as the difference between the energies of the + and – branches at  $k_x = 0$

$$E_g^\nu = \sqrt{3}at \frac{2\pi}{na} \left( \nu - \frac{2n}{3} \right). \quad (9)$$

The energy gap has a minimum value of zero corresponding to  $\nu = 2n/3$ . If  $n$  is not a multiple of three, the minimum value of  $\nu - 2n/3$  is equal to  $1/3$ . This means that the minimum energy gap is then given by

$$E_g = \frac{\sqrt{3}at}{3} \frac{2\pi}{na} = \frac{2a_{C-C}t}{d_{CNT}} \approx \frac{0.8 \text{ eV nm}}{d_{CNT}} \quad (10)$$

where  $d_{CNT} = na/\pi$  is the diameter of the CNT in nanometers. Based on (8) and (9), the DOS for semiconducting zig-zag CNTs is given by

$$g(E) = \sum_{\nu} \frac{8}{3\pi a_{C-C}t} \frac{E}{\sqrt{E - E_g^\nu/2}}, \quad (11)$$

which is an approximation valid as long as  $(E - E_g) \ll t$  [22].

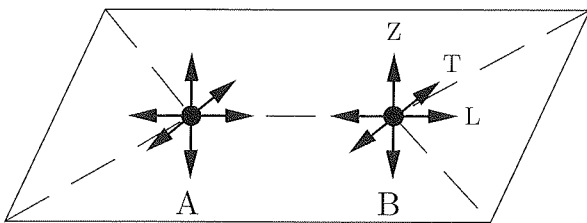
### 2.3. Phonon Properties

The phonon-dispersion relations of SWNTs can be calculated using zone folding [16], tight-binding methods [23–26], density functional theory [27–35], and symmetry-adapted models [36–40]. The phonon-dispersion relations of SWNTs can be understood by zone folding of the phonon-dispersion branches of graphene.

#### 2.3.1. Phonon-Dispersion Relations of Graphene

Since there are two carbon atoms,  $A$  and  $B$ , in the unit cell of graphene, one must consider six coordinates. The secular equation to be solved is thus a dynamical matrix of rank 6, such that six phonon branches are achieved.

The phonon-dispersion relation of the graphene comprises three acoustic (A) branches and three optical (O) branches. The modes are associated with out-of-plane (Z), inplane longitudinal (L), and inplane transverse (T) atomic motions (Fig. 4(a)).



**Figure 4.** Atomic motions of carbon atoms in graphene can be along the out-of-plane (Z), inplane transverse (T), and inplane longitudinal (L) directions.

Figure 4(b) shows the phonon-dispersion branches of graphene. The three phonon-dispersion branches, which originate from the  $\Gamma$ -point of the Brillouin zone correspond to acoustic modes: an out-of-plane mode (ZA), an inplane transverse mode (TA), and inplane longitudinal acoustic (LA), listed in the order of increasing energy. The remaining three branches correspond to optical modes: one out-of-plane mode (ZO), and two inplane modes (TO) and (LO) [13].

While the TA and LA modes display the normal linear dispersion around the  $\Gamma$ -point, the ZA mode shows a  $q^2$  energy dispersion which is explained in [13] as a consequence of the  $D_{6h}$  point-group symmetry of graphene. Another consequence of the symmetry is the linear crossings of the ZA–ZO and the LA–LO modes at the  $K$  point.

#### 2.3.2. Phonon-Dispersion Relations of SWNTs

The phonon-dispersion relations for an SWNT can be determined by folding that of a graphene layer (see Section 2.2.2). Since there are  $2N$  carbon atoms in the unit cell of a CNT,  $6N$  phonon-dispersion branches for the three-dimensional vibrations of atoms are achieved. The corresponding one-dimensional phonon energy–dispersion relation for the CNT is given by

$$\omega_{CNT}^{\mu\lambda}(q) = \omega_{g-2D}^{\lambda} \left( q \frac{\mathbf{K}_2}{|\mathbf{K}_2|} + \mu \mathbf{K}_1 \right), \quad (12)$$

where  $\lambda = 1, \dots, 6$  denotes the polarization,  $\mu = 0, \dots, N-1$  is the azimuthal quantum number, and  $-\pi/T < q \leq \pi/T$  is the wave vector of phonons. However, the zone-folding method does not always give the correct dispersion relation for a CNT, especially in the low-frequency region. For example, the out-of-plane tangential acoustic (ZA) modes of a graphene sheet do not give zero energy at the  $q = 0$  when rolled into a CNT. Here, at  $q = 0$ , all the carbon atoms of the CNT move radially in and out-of-plane radial acoustic vibration, which corresponds to a breathing mode (radial breathing mode, RBM) with a nonzero frequency [23]. To avoid these difficulties, one can directly diagonalize the dynamical matrix.

Fundamental phonon polarizations in CNTs are radial (R), transverse (T), and longitudinal (L). Zone center phonons, also referred to as  $\Gamma$ -point phonons, can belong to the transverse acoustic (TA), the LA, the RBM, the out-of-plane optical branch (RO), the transverse optical (TO), or the longitudinal optical (LO) phonon branch. The LO phonon branch near the  $\Gamma$ -point has an energy of  $\approx 190$  eV, whereas the energy of the RBM phonon branch is inversely proportional to the CNT diameter

$$\hbar\omega_{RBM} \approx 28 \text{ meV} / d_{CNT} \quad (13)$$

where  $d_{CNT}$  is the diameter of the CNT in nanometer [27, 28]. Zone boundary phonons, also referred to as  $K$ -point phonons, are found to be a mixture of fundamental polarizations [41].

### 2.3.3. Electron-Phonon Interaction

The electron-phonon interaction Hamiltonian for CNTs can be written as [42]

$$\hat{H}_{\text{el-ph}} = \sum_{k',v',k,v,\lambda} M_{k',v',k,v,\lambda} c_{k',v'}^\dagger c_{k,v} \times (b_{q,\mu,\lambda} + b_{-q,-\mu,\lambda}^\dagger), \quad (14)$$

where  $c_{k',v'}$  and  $c_{k,v}$  are the electron creation and annihilation operators, respectively,  $b_{-q,-\mu,\lambda}$  and  $b_{q,\mu,\lambda}$  are the phonon creation and annihilation operators, respectively, and  $M_{k',v',k,v,\lambda}$  is the electron-phonon matrix element

$$M_{k',v',k,v,\lambda} = \sqrt{\frac{\hbar}{2\rho_{\text{CNT}}L\omega_{q,\mu,\lambda}}} \tilde{M}_{k',v',k,v,\lambda}, \quad (15)$$

where  $\rho_{\text{CNT}}$  is the mass density of the CNT,  $L$  is the normalization length, and  $\tilde{M}_{k',v',k,v,\lambda}$  is the reduced electron-phonon matrix element of the transition from the initial electronic state  $k,v$  to the final state  $k',v'$ , where  $k$  is the wave vector and  $v$  is the azimuthal quantum number of electrons. Because of energy conservation for a scattering event it holds  $E_{k',v'} - E_{k,v} = \pm \hbar\omega_{\pm\mu,\lambda}$ , where  $q$  is the wave vector,  $\mu$  is the azimuthal quantum number, and  $\lambda$  is the polarization of the phonon. The matrix element obeys selection rules arising from wave vector and azimuthal quantum number conservation,  $q = k' - k$  and  $\mu = v - v'$ .

Because in the CNT two degrees of freedom are confined, an electron can only be scattered forward or backward in the axial direction, preserving or changing the sign of the band velocity, respectively. The scattering processes invoke either *intrasubband* or *intersubband* transitions. The intrasubband processes are important for the electrical and the heat transport in CNTs and for the relaxation of an excited electron or hole in the same subband. The intersubband processes contribute to the radiationless relaxation of electrons (holes) from a given subband to a subband with a lower (higher) energy [42]. The scattering of electrons can take place within a given valley or between two valleys, the two possibilities being termed *intravalley* and *intervalley* scattering processes.

### 2.3.4. Electron-Phonon Matrix Elements

An important case is the intrasubband scattering of electrons,  $v' = v$ , therefore,  $\mu = 0$  and  $\lambda$  can be any of six different phonon polarizations. One can omit the index  $\mu$  and write the phonon frequency as  $\omega_\lambda(q)$  and the reduced electron-phonon matrix element for a given band as  $\tilde{M}_{k',v',k,v,\lambda} = \tilde{M}_\lambda(q)$ , where the weak dependence on  $k$  is neglected.

For intravalley processes, most of the phonons have  $q \approx 0$  and are referred to as  $\Gamma$ -point phonons. Near the  $\Gamma$  point a linear dispersion relation for acoustic phonons is assumed,

$$\omega_{\text{AP}}(q) \approx v_{\text{AP}} |q|, \quad (16)$$

where  $v_{\text{AP}}$  is the acoustic phonon velocity. For OP the energy is assumed to be independent of the phonon wave vector

$$\omega_{\text{OP}}(q) \approx \omega_{\text{OP}} \quad (17)$$

Near the  $\Gamma$ -point the reduced electron-phonon matrix elements can be approximated by

$$\tilde{M}_{\text{AP}}(q) \approx \tilde{M}_{\text{AP}} |q| \quad (18)$$

for acoustic phonons and by

$$\tilde{M}_{\text{OP}}(q) \approx \tilde{M}_{\text{OP}} \quad (19)$$

for OP [42]. Phonons inducing intervalley processes have a wave vector of  $|q| \approx q_K$ , where  $q_K$  is a wave number corresponding to the  $K$ -point of the Brillouin zone of graphite. For such phonons one can neglect the  $q$ -dependence,  $\omega_K(q) \approx \omega_K$  and  $\tilde{M}_K(q) \approx \tilde{M}_K$  [42].

To calculate the electron-phonon matrix elements, one can employ the orthogonal tight-binding [43], the nonorthogonal tight-binding [42], and density functional theory [44] for the band structure and a force constant model for the lattice dynamics [13, 45]. Electron-phonon matrix elements depend on the chirality and the diameter of the CNT [42–44].

## 2.4. CNTs in Electronics

The one-dimensional nature of CNTs severely reduces the phase space for scattering, allowing CNTs to realize maximum possible bulk mobility of this material. The low-scattering probability and high mobility are responsible for high on-current of CNT transistors. Furthermore, the chemical stability and perfection of the CNT structure suggests that the carrier mobility at high gate fields may not be affected by processing and roughness scattering as in the conventional semiconductor channel. Similarly, low scattering together with the strong chemical bonding and high thermal conductivity allows metallic CNTs to withstand extremely high-current densities (up to  $\sim 10^9$  A cm $^{-2}$ ).

Electrostatics is improved in these devices as well. The fact that there are no dangling bond states at the surface of CNTs allows for a much wider choice of gate insulators other than conventional SiO $_2$ . This improved gate control without any additional gate leakage becomes very important in scaled devices with effective SiO $_2$  thickness below 1 nm. Also, the strong one-dimensional electron confinement and full depletion in the nanometer-scale diameter of the SWNTs (typically 1–2 nm) should lead to a suppression of short-channel effects in transistors [4].

The combined impact of transport and electrostatic benefits together with the fact that semiconducting CNTs are, unlike silicon, direct-gap materials, suggest applications in optoelectronics as well [5, 6]. As far as integration is concerned, semiconducting CNTs benefit from their band structure which exhibits essentially the same effective

mass for electrons and holes. This should enable similar mobilities and performance of n- and p-type transistors which is necessary for a complementary metal-oxide semiconductor (CMOS)-like technology. Finally, since CNTs can be both metallic and semiconducting, an all-CNT electronics can be envisioned. In this case, metallic CNTs could act as high current-carrying local interconnects [46], while semiconducting CNTs would form the active devices. The most important appeal of this approach is an ability to fabricate one of the critical device dimensions (the CNT diameter) reproducibly.

#### 2.4.1. Fabrication and Performance of CNT-FETs

The first CNT field-effect transistors (CNT-FETs) were reported only a few years after the initial discovery of CNTs [2, 3]. These early devices were relatively simple in structure: Noble metal (gold or platinum) electrodes were lithographically patterned atop an oxide-coated, heavily doped silicon wafer, and a SW-CNT was deposited atop the electrodes. The metal electrodes served as the source and drain, and the CNT was the active channel. The doped substrate served as the gate electrode, separated from the CNT channel by a thick (~100–200 nm) oxide layer. These devices displayed clear p-type transistor action, with gate voltage modulation of the drain current over several orders of magnitude. The devices displayed high parasitic resistance ( $\geq 1 \text{ M}\Omega$ ), low drive current, low transconductance ( $g_m \sim 1 \text{ nS}$ ), high subthreshold slope ( $S = [d(\log_{10} I_d) / dV_g]^{-1} \sim 1 \text{ V/decade}$ ), and no current saturation. Owing to the thick gate dielectric, these devices required large values of the gate voltage (several volts) to turn on, making them unattractive for practical applications.

Following these initial results, advances in CNT-FET device structures and processing yielded improvements in their electrical characteristics. Rather than laying the CNT down upon the source and drain electrodes, relying on weak van der Waals forces for contact, the CNTs were first deposited on the substrate and the electrodes were patterned on top of the CNTs. In addition to Au, Ti, and Co were used [47–49] with a thermal annealing step to improve the metal–CNT contact. In the case of Ti, the thermal processing leads to the formation of TiC at the metal–CNT interface [48], resulting in a significant reduction in the contact resistance from several  $\text{M}\Omega$  to  $\sim 30 \text{ k}\Omega$ . On-state currents  $\sim 1 \mu\text{A}$  were measured, with a transconductance of  $\sim 0.3 \mu\text{S}$  — an improvement of more than two orders of magnitude relative to the van der Waals contacted devices. This CNT-FET device configuration has been extensively studied in the literature. More recently, it has been found that Pd forms a low resistance contact to CNTs for p-type devices [50]. It is speculated [50] that Pd offers improved sticking or wetting interaction to the CNT surface relative to other metals, as well as good Fermi level alignment relative to the CNT conduction band, see Section 2.4.2.

As mentioned above, early CNT-FETs were p type in air (hole conduction). However, it was found that n-type conduction could be achieved by doping from an alkali (electron

donor) gas [51] or by thermal annealing in vacuum [48]. In addition, it is possible to achieve an intermediate state, in which both electron and hole injection occur, resulting in ambipolar conduction [48]. The ability to controllably fabricate both p- and n-type CNT-FETs is a key to the formation of logic circuits.

Early experiments on CNT-FETs were built upon oxidized silicon wafers, with the substrate itself serving as the gate and a thermally grown oxide film, typically  $\sim 100 \text{ nm}$  or thicker, serving as the gate dielectric. The thick gate oxide required relatively high gate voltages ( $\sim 10 \text{ V}$ ) to turn on the devices, and the use of the substrate as the gate implied that all CNT-FETs must be turned on and off together, precluding the implementation of complex circuits. A more advanced CNT-FET structure [52] is shown in Figure 5. The device comprises a top-gate separated from the CNT channel by a thin-gate dielectric. The top-gate allows independent addressing of individual devices, making it more amenable to integration in complex circuits, while the thin-gate dielectric improves the gate to channel coupling, enabling low-voltage operation. In addition, the reduction of the capacitance owing to gate-source and gate-drain overlap suggests that such a device structure would be appropriate for high-frequency operation. Such a CNT-FET can also be switched using the conductive substrate as a bottom gate, allowing for direct comparison between top and bottom gate operation. Operating the device with the top-gate yields distinctly superior performance relative to bottom gate operation, with a lower threshold voltage ( $-0.5 \text{ V}$  vs.  $-12 \text{ V}$ ), higher transconductance ( $3.25 \mu\text{S}$  vs.  $0.1 \mu\text{S}$ ), and superior subthreshold behavior ( $130 \text{ mV/decade}$  vs.  $2 \text{ V/decade}$ ) [52].

In order to gauge whether or not CNT-FETs have potential for future nanoelectronic applications, it is important to compare their electrical performance to those of advanced silicon devices. Wind et al. [52] demonstrated that although the device structure is far from optimized, the electrical characteristics, such as the on-current and the transconductance of the device shown in Figure 5, exceed those of the state-of-the-art silicon MOSFETs. Further enhancements to CNT-FET structures, such as the use of high dielectric constant gate insulators [53, 54] and additional improvements in the metal–CNT contact resistance at the source and drain [50], have lead directly to improved CNT-FET performance. Such improvements can also be applied to n-type CNT-FETs [55].

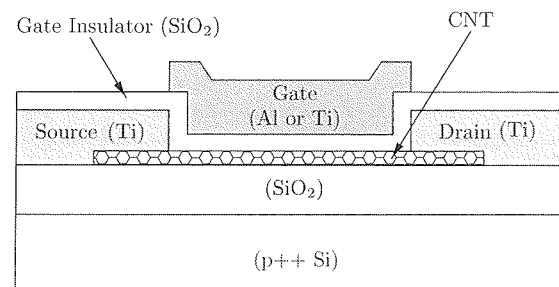


Figure 5. Schematic cross section of top-gate CNT-FET.

### 2.4.2. SB Model of CNT-FET Operations

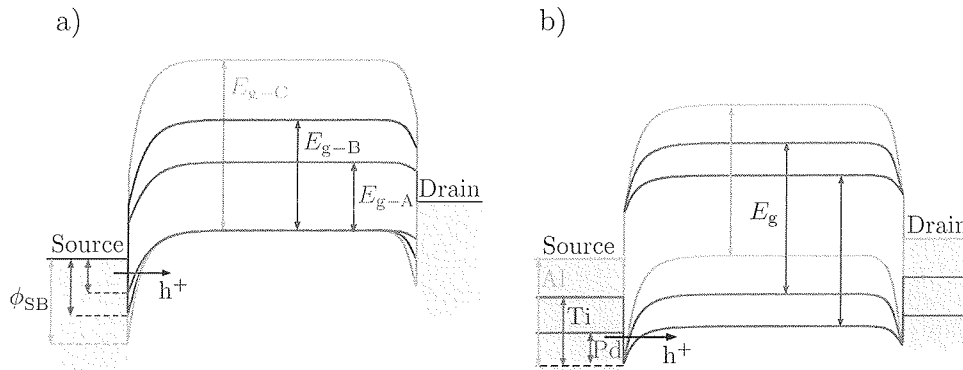
In general a charge transfer will take place at the metal–CNT interface leading to band bending and the creation of an SB. For example, a CNT-FET with titanium-carbide contacts shows equal hole and electron currents depending on the sign of the applied gate bias, so-called ambipolar conduction [48]. This suggests the existence of two barriers, one for electrons and one for holes, of approximately equal height, implying that each must be about half the band gap ( $E_g/2 \approx 300$  meV). Applying conventional semiconductor analysis, which assumes that thermionic emission contributes mostly to the total current through a SB, indeed yields similar thermal activation barriers for electrons and holes; however, on the order of 10 meV [48]. This finding suggested that the thermionic contribution alone cannot account for the observed current levels, which is supported by modeling results showing that SBs in one dimension are much thinner than their planar analogues [56, 57]. Consequently, carrier tunneling through these thin barriers becomes the dominant conduction mechanism and cannot be neglected when quantifying the barrier height [58].

Similar conclusions can be drawn from the subthreshold behavior of CNT-FETs, in particular when plotted as a function of gate oxide thickness. The switching behavior of an MOSFET is described by the inverse subthreshold slope,  $S = (k_B T / q) \ln(10) (1 + C_D / C_g)$  where  $C_D$  and  $C_g$  are the depletion and gate capacitance, respectively. In the case of a fully depleted device,  $C_D$  is zero and, therefore,  $S$  depends only on the temperature, having a value of 65 mV/decade at room temperature. The original CNT-FETs with thick gate oxides in back-gated geometry had unexpectedly high  $S$  values of approximately 1 V/decade. On the other hand, when devices are fabricated using thinner oxides, such as the top-gated CNT-FET in [52], the value of  $S$  dropped significantly into the range of 100–150 mV/decade [52], 80 mV/decade [59], and 67–70 mV/decade [60]. Such a dependence of  $S$  is not consistent with the bulk switching mechanism which should give 65 mV/decade in the long channel limit. Instead, this kind of scaling of the subthreshold slope with oxide thickness is compatible with the existence of sizeable SBs at the metal–CNT interfaces, and theoretical modeling showed that the gate field impact on this

interface is responsible for the observed improvement in  $S$  [61, 62].

Further evidence of the presence of SBs in CNT-FET devices is found in local gating experiments, where the on-current is shown to increase significantly by application of a local potential from a metal-coated scanning probe tip only at the positions above the metal–CNT interface [63]. Similarly, the impact of SBs in the subthreshold characteristics of the CNT-FET is clearly observed in transistors with multiple top-gates [64]. In this case, local gates over the metal–CNT interface are used to electrostatically thin the SBs and reduce the value of  $S$  closer to that of the bulk switching device [64].

Hole (electron) injection into the CNT depends on the line-up of the metal Fermi level and the valence (conduction) band of the CNT, which is defined here as the SB height. In this picture, other details of the contacts such as any changes in the metal–CNT coupling as a function of the curvature of the CNT are incorporated in an *effective* Schottky tunneling barrier height. This barrier height depends on a number of material parameters such as the band gap of the CNT, work-function difference, as well as the interface quality. The CNT band gap is inversely proportional to the diameter of the CNT, according to (10). Figure 6(a) shows qualitative band diagrams for CNT-FETs with different diameters. Assuming a constant work function for all CNTs (The work function is defined as the sum of the CNT electron affinity and half of the band gap in the bulk.), the SB increases linearly with increasing band gap. On a log scale, current injection through the SB is inversely proportional to the barrier height. Therefore, the CNT-FET with a small diameter delivers low on-current. The choice of the metal contacts also affects the device performance. Figure 6(b) depicts the band diagrams for CNT-FETs using different source and drain contact materials. Identical energy band gaps are drawn here to represent CNTs of the same diameter. CNT-FETs with Pd contacts deliver the highest on-current, since Pd has the highest work function (5.1 eV), which forms a low SB height to the valence band of the CNT. The trend shown follows that of the clean metal work functions: 4.3 eV for Ti and 4.1 eV for Al [65].



**Figure 6.** (a) Schematic band diagram showing the different SB heights for (b) CNT-FETs with the same contact, but with different CNT diameters and (c) CNT-FETs with the same diameter, but using Pd, Ti, and Al contacts, respectively.



The effect of ambient air on the performance and functionality of CNT-FETs can also be understood within the framework of the SB model of conduction. In particular, this model helps to clarify and separate the effects owing to the bulk of the CNT channel from those arising from the effects at the contact between the metal electrode and the CNT.

It has been proposed that, for example, oxygen adsorption leads to doping of CNTs [66]. However, the effect of oxygen on the transport properties of a CNT-FET is a reversible transition from p (devices prepared in air) to n type after annealing the transistor in vacuum [67]. In contrast, the deposition of an n-type dopant such as potassium shifts the transfer characteristics with respect to the gate voltage. It is known that the work function of a metal surface is altered significantly upon the adsorption of gases owing to the formation of interface dipoles. Thus, the local work function of the metal electrode can be modified considerably by the adsorption of oxygen at the contacts. If the work function of the metal electrode changes the line-up of the metal Fermi energy with the CNT, bands will shift (Note that this is unique for the contact between a metal and a CNT. In a conventional, planar semiconductor device the position of the Fermi energy is pinned by metal-induced gap states [56].) [57].

### 2.4.3. Prototype CNT-FET Circuits

The promising characteristics of individual CNT-FETs have lead to initial attempts at integration of several CNT-FETs into useful circuits that can perform a logic operation, or function as memories [68] or sensors [69]. In the following, we limit our discussion to advances in logic circuitry. The CNT logic gates have been, in most cases, based on a complementary technology analogous to silicon CMOS, which is important as it may ease integration of CNTs onto this well-established technology.

The first complementary (CMOS-like) logic gates were reported by Derycke and coworkers [51]. In that work, two different techniques were used to produce n-type devices. An inverter gate was created by combining two CNT-FETs: a p-type device in the ambient and a vacuum-annealed n-type device. A more compact and integrated approach uses potassium doping to convert one of the two CNT-FETs built on the same CNT to n type. The masking of the other transistor which remained p type was accomplished by photoresist. The circuit had a voltage gain of about two, suggesting that integration, without signal degradation, of many devices along a single CNT can be accomplished. Shortly thereafter, Bachtold and coworkers [70] used p-type CNT-FETs along with resistors to build prototype logic gates based on an older transistor-resistor scheme. They went a step further in complexity and wired three such inverter gates to form a ring oscillator. The large parasitic capacitances severely degraded the performance of the circuit which oscillated at only about 5 Hz. Later Javey and coworkers [54] used another scheme for converting p- into n-type CNT-FETs and to wire up CMOS inverters with gains in excess of 10 and CMOS ring oscillators with frequencies in the 100 Hz range. Very recently Chen and coworkers [71] reported CMOS ring oscillators operating at frequencies up to 52 MHz. However, these reported frequencies are well below the expected AC response of CNT-FETs, which is difficult to assess because of the relatively small current signals

in these devices. Most recently, the nonlinear current-voltage characteristics of CNT-FETs were used to demonstrate that DC characteristics of CNT-FETs are not affected by AC fields at least up to 500 MHz [72]. However, operation up to 10 GHz [73] and later 50 GHz [74] have been reported, albeit with considerable signal attenuation.

## 3. QUANTUM TRANSPORT MODELS

This section outlines the theory of the NEGF techniques for modeling transport phenomena in semiconductor devices. The NEGF techniques, initiated by Schwinger [75], Kadanoff, and Baym [76], allow one to study the time evolution of a many-particle quantum system. Knowledge of the single-particle Green's function provides both the complete equilibrium or nonequilibrium properties of the system and the excitation energies of the systems containing one more or one less particle. The many-particle information about the system is cast into self-energies, parts of the equations of motion for Green's functions. Green's functions can be expressed as a perturbation expansion, which is the key to approximate the self-energies. The NEGF techniques provide a very powerful technique for evaluating properties of many-particle systems both in thermodynamic equilibrium and also in nonequilibrium situations.

The basic approach developed in the early 1970s has become increasingly popular during the last 10 years. The motivation for the development of the NEGF tunneling formalism was the metal-insulator-metal tunneling experiments that received much attention during the 1960s [77]. The accelerated use of the approach was motivated by experimental investigations of mesoscopic physics made possible by high-quality semiconductor heterostructures grown by molecular-beam epitaxy. In 1988, Kim and Arnold were the first to apply the NEGF formalism to such a system, specifically, a resonant-tunneling diode [78]. As experimental methods progressed, allowing finer manipulation of matter and probing into the nanoscale regime, the importance of quantum effects and tunneling continuously increased. The theory was adapted to address the current systems of interest ranging from mesoscopic to single-electronics, nanoscaled FETs, and molecular electronics.

The general formalism for NEGF calculations of current in devices was first described in a series of papers in the early 1970s [79–82]. The partitioning of an infinite system into left contact, device, and right contact, and the derivation of the open boundary self-energies for a tight-binding model was presented in [79]. This theory was rederived for a continuum representation in [79B], tunneling through localized impurity states was considered in [81], and a treatment of phonon-assisted tunneling was derived in [82]. In 1976, the formalism was first applied to a multiband model (two bands) to investigate tunneling [83] and diagonal disorder [84], and in 1980 it was extended to model time-dependent potentials [85].

The applications of the NEGF techniques have been extensive including quantum optics [86], quantum corrections to the Boltzmann transport equation [87, 88], high field transport in bulk systems [89], and electron transport

through nanoscaled systems. Over the last decade, NEGF techniques have become widely used for modeling high-bias, and quantum electron and hole transport in a wide variety of materials and devices: III–V resonant tunnel diodes [78, 90–103], electron waveguides [104], superlattices used as quantum cascade lasers [105], Si tunnel diodes [106, 107], ultrascaled Si MOSFETs [7, 108–110], Si nanopillars [111, 112], CNTs [113–124], metal wires [125, 126], organic molecules [127–143], superconducting weak links [144], and magnetic leads [115, 145, 146]. Physics that have been included are open-system boundaries [79], full band structure [98, 99, 106, 107, 147], band tails [107], the self-consistent Hartree potential [92, 148], exchange-correlation potentials within a density functional approach [10, 114, 125, 128, 129, 131], acoustic, optical, intravalley, intervalley, and interband phonon scattering, alloy disorder and interface roughness scattering in Born-type approximations [93–98, 105–107], photon absorption and emission [105], energy and heat transport [101], single-electron charging and nonequilibrium Kondo systems [149–155], shot noise [96, 100, 156], A.C. [91, 157–162], and transient response [159, 163]. Time-dependent calculations are described further in [164]. General tutorials on the NEGF techniques [165, 166] and the applications can be found in [98, 167, 168].

This section continues with a brief description of standard expressions, where one shall rely on the *second quantization* formulation. Various formulations of many-particle Green's function theory exist. For instance, in equilibrium theory there is both a zero-temperature as well as a finite-temperature (Matsubara) formalism [164, 169, 170]. The formulation of the more general nonequilibrium finite-temperature theory which also applies to equilibrium situations as a special case is introduced next, and the kinetic equations for this formalism are discussed. Applying Wick theorem, a perturbation expansion of Green's functions can be achieved. Such expansions provide methods to approximate self-energies owing to various scattering mechanisms. Finally, a comparison of Green's function formalism with other transport models is presented.

### 3.1. NEGFs

In this section, a general formalism for systems at finite temperature under nonequilibrium conditions is presented. First, the ensemble average of an operator for nonequilibrium is defined. Then the contour-ordered NEGF formalism is introduced and the equation of motion for Green's function is presented.

#### 3.1.1. Equilibrium Ensemble Average

In many-particle problems it is common to use the interaction representation (see Appendix A)

$$\hat{H} = \hat{H}_0 + \hat{H}^{\text{int}}, \quad (20)$$

where  $H_0$  is the noninteracting Hamiltonian and  $H^{\text{int}}$  is the perturbation, which contains all the interactions, that is  $\hat{H}^{\text{int}} = \hat{H}_{\text{el-el}} + \hat{H}_{\text{el-ph}}$ . At finite temperatures, one assumes that the particle, either electron or phonon, is interacting

with a bath of other particles. The exact state of all these other particles is not known, since they are fluctuating between different configurations. At finite temperature under thermodynamic equilibrium the state of a system is described by the equilibrium density operator  $\hat{\rho}$  [171]. In treating such systems, it will be most convenient to use the grand canonical ensemble, which allows for a variable number of particles. Therefore, the system is considered to be in contact with a heat bath of temperature  $T$  and a particle reservoir characterized by the Fermi energy  $E_F$ . With the definition  $\hat{K} = \hat{H} - E_F \hat{N}$ , where  $\hat{N}$  is the particle number operator, the statistical operator can be written as

$$\hat{\rho} = \frac{e^{-\beta \hat{K}}}{\text{Tr} [e^{-\beta \hat{K}}]}, \quad (21)$$

where the short-hand notation  $\beta = 1/k_B T$  is used. The operator  $\hat{K}$  may be interpreted as a *grand canonical* Hamiltonian. Given the density operator, the ensemble average of any operator  $\hat{O}$  can be calculated as

$$\begin{aligned} \langle \hat{O} \rangle &= \text{Tr} [\hat{\rho} \hat{O}], \\ &= \frac{\text{Tr} [e^{-\beta \hat{K}} \hat{O}]}{\text{Tr} [e^{-\beta \hat{K}}]}. \end{aligned} \quad (22)$$

#### 3.1.2. Nonequilibrium Ensemble Average

We employ the standard device for obtaining a nonequilibrium state. At time  $t_0$ , prior to which the system is assumed to be in thermodynamic equilibrium with a reservoir, the system is exposed to a disturbance represented by the contribution  $\hat{H}^{\text{ext}}$  to the Hamiltonian. The external perturbation can, for instance, be a time-varying electric field, a light excitation pulse, and so forth. The total Hamiltonian is thus given by

$$\hat{\mathcal{H}}(t) = \hat{H}_0 + \hat{H}^{\text{int}} + \hat{H}^{\text{ext}} = \hat{H} + \hat{H}^{\text{ext}}, \quad (23)$$

where  $\hat{H}^{\text{ext}} = 0$  for  $t < t_0$ . One is not restricted to using the statistical equilibrium state at times prior to  $t_0$  as the initial condition. As shown by [172], a nonequilibrium situation can be maintained through contact with a reservoir. A discussion of the coupling of a system to a reservoir has been studied in [173]. Nonequilibrium statistical mechanics is concerned with calculating average values  $\langle \hat{O}_{\neq}(t) \rangle$  of physical observables for times  $t > t_0$ . Given the density operator  $\hat{\rho}$ , the average of any operator  $\hat{O}$  is then defined as

$$\langle \hat{O}_{\neq}(t) \rangle \equiv \text{Tr} [\hat{\rho} \hat{O}_{\neq}(t)], \quad (24)$$

where  $\hat{O}_{\neq}(t)$  is an operator in the Heisenberg picture. The time-ordered NEGF can be defined as

$$G(\mathbf{r}, t, \mathbf{r}', t') = -\frac{i}{\hbar} \left\langle T_1 \left\{ \hat{\psi}_{\neq}(\mathbf{r}, t) \hat{\psi}_{\neq}^{\dagger}(\mathbf{r}', t') \right\} \right\rangle, \quad (25)$$

where  $\hat{\psi}$  is the field operator in the Heisenberg picture evolving with the Hamiltonian  $\mathcal{H}$  defined in (23), the bracket  $\langle \dots \rangle$  is the statistical average with the density operator defined in (24), and  $T$  is the time-ordering operator.

One can evaluate Green's functions by using Wick's theorem, which enables us to decompose many-particle Green's functions into sums and products of single-particle Green's functions [169]. The restriction of the Wick theorem necessitates that the field operators and the density operator have to be represented in the interaction picture, or equivalently, their time evolution is governed by the noninteracting Hamiltonian  $\hat{H}_0$ . The contour-ordered Green's function, which is introduced next, provides a suitable framework for this purpose.

### 3.2. Contour-Ordered Green's Function

To express the field operators in the interaction representation an operator  $\hat{S}$  is defined (see Appendix A.4). The time integral in the  $\hat{S}$  operator can be written as two parts: one goes from  $(-\infty, t)$ , while the second goes from  $(t, -\infty)$  and one then obtains the contour  $C_K$  introduced by Keldysh [174]. The integration path is a contour, which starts and ends at  $-\infty$ . Instead of the time-ordering operator (154), a contour-ordering operator can be employed. The contour-ordering operator  $T_C$  orders the time labels according to their order on the contour  $C$ . Under equilibrium condition the contour-ordered method gives results that are identical to the time-ordered method. The main advantage of the contour-ordered method is in describing nonequilibrium phenomena using Green's functions. Nonequilibrium theory is entirely based upon this formalism, or equivalent methods [164].

Any operator  $\hat{O}_H$  in the Heisenberg picture can be transformed into the interaction picture (see (146))

$$\begin{aligned}\hat{O}_H(t) &= T_C \left\{ \hat{S}_C^{\text{ext}} \hat{O}_H \right\}, \\ &= T_C \left\{ \hat{S}_C^{\text{ext}} \hat{S}_C^{\text{int}} \hat{O}_I \right\}, \\ &= T_C \left\{ \hat{S}_C \hat{O}_I \right\},\end{aligned}\quad (26)$$

$$G(\mathbf{r}, t; \mathbf{r}', t') = -\frac{i}{\hbar} \frac{\left\langle \sum_{n=0}^{\infty} \frac{1}{n!} \left( \frac{-i}{\hbar} \right)^n \int dt_1 \dots \int dt_n T \left\{ \hat{H}_I^{\text{int}}(t_1) \dots \hat{H}_I^{\text{int}}(t_n) \hat{\psi}_I(\mathbf{r}, t) \hat{\psi}_I^\dagger(\mathbf{r}', t') \right\} \right\rangle}{\left\langle \sum_{n=0}^{\infty} \frac{1}{n!} \left( \frac{-i}{\hbar} \right)^n \int dt_1 \dots \int dt_n T \left\{ \hat{H}_I^{\text{int}}(t_1) \dots \hat{H}_I^{\text{int}}(t_n) \right\} \right\rangle}. \quad (30)$$

The expansion of the numerator of Green's function in (30) can be written as

$$\begin{aligned}G_N &= \underbrace{\left\langle T_t \left\{ \hat{\psi}_I(\mathbf{r}, t) \hat{\psi}_I^\dagger(\mathbf{r}', t') \right\} \right\rangle}_{G_N^0} \\ &+ \underbrace{\left\langle T_t \left\{ -\frac{i}{\hbar} \int dt_1 \hat{H}_I^{\text{int}}(t_1) \hat{\psi}_I(\mathbf{r}, t) \hat{\psi}_I^\dagger(\mathbf{r}', t') \right\} \right\rangle}_{G_N^1} + \dots,\end{aligned}\quad (31)$$

where the superscript denotes the order of perturbation. The zero-order perturbation leads to noninteracting

where

$$\begin{aligned}\hat{S}_C &= \exp \left( -\frac{i}{\hbar} \int_C dt \hat{H}_I^{\text{ext}}(t) \right) \exp \left( -\frac{i}{\hbar} \int_C dt \hat{H}_I^{\text{int}}(t) \right), \\ &= \hat{S}_C^{\text{ext}} \hat{S}_C^{\text{int}}\end{aligned}\quad (27)$$

Therefore, Green's function in (25) is given by

$$G(\mathbf{r}, t; \mathbf{r}', t') = -\frac{i}{\hbar} \left\langle T_C \left\{ \hat{S}_C \hat{\psi}_I(\mathbf{r}, t) \hat{\psi}_I^\dagger(\mathbf{r}', t') \right\} \right\rangle_0, \quad (28)$$

where  $\langle \dots \rangle_0$  represents the statistical average with respect to  $\hat{\rho}_0$  (noninteracting density operator). From here we assume that all statistical averages are with respect to  $\hat{\rho}_0$  and drop the 0 from the brackets  $\langle \dots \rangle_0$

### 3.3. Perturbation Expansion of Green's Function

It was shown that Green's functions can be written in terms of the  $\hat{S}$  operator, see (28), where  $\hat{S}$  includes the effects of interactions and external perturbations, see (27). Unfortunately, it is not possible to give an analytical solution for  $G(\mathbf{r}, t; \mathbf{r}', t')$ , unless the interaction perturbation  $\hat{H}^{\text{int}}$  is set equal to zero. This gives the noninteracting Green's function (see Appendix B)

$$G_0(\mathbf{r}, t; \mathbf{r}', t') = -\frac{i}{\hbar} \left\langle T \left\{ \hat{\psi}_I(\mathbf{r}, t) \hat{\psi}_I^\dagger(\mathbf{r}', t') \right\} \right\rangle, \quad (29)$$

which is central for any perturbation expansion.

Green's function by expanding the  $\hat{S}$  operator as series of products of  $\hat{H}_I^{\text{int}}$  in the numerator and the denominator. By expanding the  $\hat{S}$  operator one obtains (see (153))

Green's function  $G_N^0 = i\hbar G_0$ . Wick's theorem allows us to write each of these brackets in terms of noninteracting Green's function and the interaction potential. The same procedure can be applied to the denominator. The terms

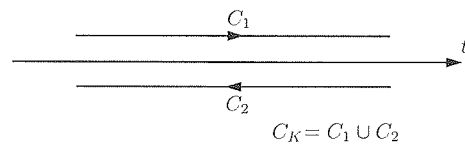


Figure 7. Keldysh contour branches  $C_1 = (-\infty, \infty)$  and  $C_2 = (\infty, -\infty)$ .

in the expansion of the denominator,  $\langle S(\infty, -\infty) \rangle$ , are called *vacuum polarization terms* [175].

Different expansion terms achieved from the Wick theorem can be translated into Feynman diagrams. Feynman introduced the idea of representing different contributions obtained from the Wick decomposition by drawings. These drawings, called diagrams, are very useful for providing an insight into the physical processes which these terms represent. The Feynman diagrams provide an illustrative way to solve many-body problems and the perturbation expansion of Green's functions.

### 3.3.1. Wick Theorem

The Wick decomposition allows a perturbation expansion of Green's functions. It always holds for zero-temperature Green's functions and only under the condition that field operators must be given in the interaction picture (Appendix A.2). Their time evolution is governed by the noninteracting Hamiltonian  $\hat{H}_0$ , and  $\hat{H}^{\text{int}}$  is treated as a perturbation. If these conditions are fulfilled, Wick's theorem states that the expectation values of the products of the field operators are equal to the sum of the expectation values of all the possible pairs of the operators and that each of these pairs will be a noninteracting single-particle Green's function

$$\langle T_I \{ \hat{O}_1 \hat{O}_2 \dots \hat{O}_n \} \rangle = \sum_{P_d} \langle T_I \{ \hat{O}_1 \hat{O}_2 \} \rangle \langle T_I \{ \hat{O}_3 \hat{O}_4 \} \rangle \dots \langle T_I \{ \hat{O}_{n-1} \hat{O}_n \} \rangle. \quad (32)$$

The sum runs over all  $P_d$  distinct permutations of the  $n$  indices. It should be noticed that brackets such as (32) vanish if the number of creation and annihilation operators is not the same. If the number of annihilated particles is not the same as the number of created particles, then the system will not come back to its ground state. As a result the expectation value over the ground state vanishes. With the same reasoning one concludes that if both the operators appearing in a bracket are annihilation or creation operators, the expectation value disappears, otherwise one obtains an expression proportional to the noninteracting Green's function  $G_0$ . The most general proof of this theorem is owing to [176].

### 3.4. Dyson Equation

The Dyson equation can be achieved by classifying the various contributions in arbitrary Feynman diagrams. Dyson's equation summarizes the Feynman–Dyson perturbation theory in a particularly compact form. The exact Green's function can be written as the noninteracting Green's function plus all connected terms with a noninteracting Green's function at each end. By introducing the concept of self-energy  $\Sigma$ , the Dyson equation takes the form shown in Figure 8.

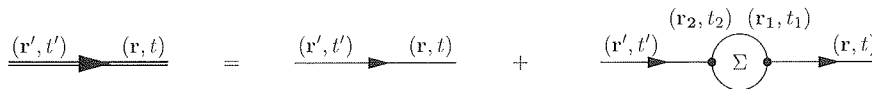


Figure 8. Feynman diagrams showing the general structure of  $G$ .

The corresponding analytic expression is given by

$$G(\mathbf{r}, t; \mathbf{r}', t') = G_0(\mathbf{r}, t; \mathbf{r}', t') + \int d1 \int d2 G_0(\mathbf{r}, t; 1) \Sigma(12) G_0(2; \mathbf{r}', t'), \quad (33)$$

where the abbreviations  $1 \equiv (r_1, t_1)$  and  $\int d1 \equiv \int dr_1 \int dt_1$  are used. The self-energy  $\Sigma$  describes the renormalization of single-particle states owing to the interaction with the surrounding many-particle system and the Dyson equation determines the renormalized Green's function.

Another important concept is the *proper self-energy insertion* which is a self-energy insertion that cannot be separated into two pieces by cutting a single-particle line. By definition, the proper self-energy is the sum of all the proper self-energy insertions, and will be denoted by  $\Sigma^*$ . Using the perturbation expansion, one can define the proper self-energy  $\Sigma^*$  as an irreducible part of Green's function. Based on this definition first-order proper self-energies, which are resulted from the first-order expansion of Green's function, are shown in Figure 9. These diagrams are referred to as the Hartree ( $\Sigma^H$ ) and the Fock ( $\Sigma^F$ ) self-energies.

The self-energy can also, in principle, be introduced variationally [176]. It follows from these definitions that the self-energy consists of a sum of all possible repetitions of the proper self-energy

$$\Sigma(\mathbf{r}, t; \mathbf{r}', t') = \Sigma^*(\mathbf{r}, t; \mathbf{r}', t') + \int d1 \int d2 \Sigma^*(\mathbf{r}, t; 1) G_0(12) \Sigma^*(2; \mathbf{r}', t') + \dots \quad (34)$$

Correspondingly, Green's function in (33) can be rewritten as

$$G(\mathbf{r}, t; \mathbf{r}', t') = G_0(\mathbf{r}, t; \mathbf{r}', t') + \int d1 \int d2 G_0(\mathbf{r}, t; 1) \Sigma^*(12) G_0(2; \mathbf{r}', t') + \dots, \quad (35)$$

which can be summed formally to yield an *integral equation* (Dyson equation) for the exact Green's function

$$G(r, t; r', t') = G_0(r, t; r', t') + \int d1 \int d2 G_0(r, t; 1) \Sigma^*(12) G(2; r', t'). \quad (36)$$

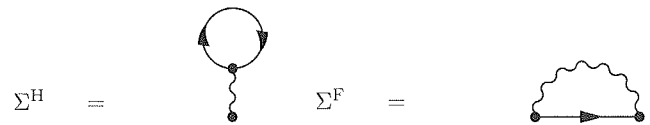


Figure 9. Feynman diagrams of the first-order proper self-energies.

In a similar manner, one can show that the Dyson equation can also be written as

$$G(r, t; r', t') = G_0(r, t; r', t') + \int d1 \int d2 G(r, t; 1) \Sigma^*(1, 2) G_0(2; r', t'). \quad (37)$$

### 3.5. Approximation of the Self-Energy

An exact evaluation of the self-energy is possible only for some rather pathological cases. For real systems one has to rely on approximation schemes. Hence, a natural approach is to retain the single-particle picture and assume that each particle moves in a single-particle potential that comes from its average interaction with all of the other particles. Thus, as a first-order approximation one can keep just the first-order contribution to the proper self-energy  $\Sigma^* \approx \Sigma_1$  (see Fig. 9). This approximation corresponds to summing an infinite class of diagrams containing arbitrary iterations of  $\Sigma_1$ . Therefore, any approximation for  $\Sigma^*$  generates an *infinite-order* series for Green's function.

However, using noninteracting Green's function in self-energies, which is referred to as *Born approximation*, is not fully consistent, since the background particles contributing to the self-energy are treated as noninteracting. In reality, of course, these particles also move in an average potential coming from the presence of all the other particles. Thus, instead of noninteracting Green's functions (single lines), one has to use the exact Green's function (double line) in the proper self-energy, as shown in Figure 10. Since the exact Green's function  $G$  both determines and is determined by the proper self-energy  $\Sigma^*$ , this approximation is known as the *self-consistent Born approximation*. The self-consistent approach preserves conservation laws, for example, the continuity equation holds valid (Section 3.8.4). Throughout this work the self-consistent Born approximation is applied.

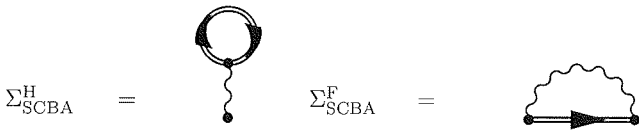


Figure 10. Feynman diagrams of the self-consistent first-order proper self-energies.

#### 3.5.1. Electron–Electron Interaction

The self-consistent Hartree self-energy owing to electron–electron interaction is given by [177]

$$\begin{aligned} \Sigma_{el-el}(r_1, t_1) &= -i\hbar \int dt_3 \int dr_3 \delta_{t_1, t_3} V(r_1 - r_3) G(r_3, t_3; r_3, t_3), \\ &= -i\hbar \int dr_3 V(r_1 - r_3) G(r_3, t_1; r_3, t_1), \\ &= \int dr_3 V(r_1 - r_3) n(r_3, t_1), \\ &= \int dr_3 \frac{q^2}{4\pi \epsilon |r_1 - r_3|} \frac{\varrho(r_3, t_1)}{-q}, \\ &= -q\phi(r_1). \end{aligned} \quad (38)$$

where  $\varrho(\mathbf{r}, t) / (-q) = n(\mathbf{r}, t) = -i\hbar G(\mathbf{r}, t, \mathbf{r}, t)$  (see Section 3.8.1).

The potential  $\phi$  resulting from the Hartree self-energy is in fact the solution of Poisson's equation with the charge density  $\varrho$ . The Hartree self-energy is instantaneous.

#### 3.5.2. Electron–Phonon Interaction

The electron–phonon interaction Hamiltonian can be written as [170]

$$\hat{H}_I^{\text{el-ph}}(t_1) = \int d\mathbf{r}_1 \hat{\psi}_I^\dagger(\mathbf{r}_1, t_1) \left( \sum_{\mathbf{q}, \lambda} e^{i\mathbf{q}_1 \cdot \mathbf{r}_1} M_{\mathbf{q}_1, \lambda} \hat{A}_{\mathbf{q}_1}(t_1) \right) \hat{\psi}_I(\mathbf{r}_1, t_1), \quad (39)$$

where  $\hat{A}_{\mathbf{q}, \lambda}(t) = (b_{\mathbf{q}, \lambda} e^{-i\omega_{\mathbf{q}, \lambda} t} + b_{-\mathbf{q}, \lambda}^\dagger e^{+i\omega_{\mathbf{q}, \lambda} t})$ ,  $b_{\mathbf{q}, \lambda}$  and  $b_{\mathbf{q}, \lambda}^\dagger$  are the annihilation and creation operators for phonons with wave vector  $\mathbf{q}$ , polarization  $\lambda$ , and energy  $\hbar\omega_{\mathbf{q}, \lambda}$ , and  $M_{\mathbf{q}, \lambda}$  is the electron–phonon interaction matrix element. The zero-order perturbation gives the noninteracting Green's function. The first-order term of the perturbation expansion must vanish because it contains the factor  $\langle \hat{A}_{\mathbf{q}, \lambda} \rangle$  which is zero since the factors  $\langle b_{\mathbf{q}, \lambda} \rangle$  and  $\langle b_{\mathbf{q}, \lambda}^\dagger \rangle$  are zero [170]. Similarly, all the odd terms vanish because their time-ordered bracket for phonons contains an odd number of  $\hat{A}_{\mathbf{q}, \lambda}$  factors. Applying Wick's theorem (Section 3.3.1), only the even terms contribute to the perturbation expansion for the electron–phonon interaction where  $F_N^1$  is the time-ordered products of electron operators. Notice that, owing to the properties of the annihilation and

$$\begin{aligned} G_N^1 &= \left\langle T_1 \left\{ \frac{1}{2} \left( \frac{-i}{\hbar} \right)^2 \int dt_1 \hat{H}_I^{\text{el-ph}}(t_1) \hat{\psi}_I(\mathbf{r}, t) \hat{\psi}_I^\dagger(\mathbf{r}', t') \right\} \right\rangle \\ &= \frac{1}{2} \left( \frac{-i}{\hbar} \right)^2 \int dt_1 \int dt_2 \int d\mathbf{r}_1 \int d\mathbf{r}_2 \\ &\quad \times \left\langle T_1 \left\{ \hat{\psi}_I^\dagger(\mathbf{r}_1, t_1) \hat{\psi}_I^\dagger(\mathbf{r}_2, t_2) \hat{\psi}_I(\mathbf{r}_2, t_2) \hat{\psi}_I(\mathbf{r}_2, t_2) \hat{\psi}_I(\mathbf{r}, t) \hat{\psi}_I^\dagger(\mathbf{r}', t') \right\} \right\rangle \\ &\quad \times \underbrace{\sum_{\mathbf{q}_1, \mathbf{q}_2, \lambda} e^{i\mathbf{q}_1 \cdot \mathbf{r}_1} e^{i\mathbf{q}_2 \cdot \mathbf{r}_2} M_{\mathbf{q}_1, \lambda} M_{\mathbf{q}_2, \lambda} \langle A_{\mathbf{q}_1, \lambda}(t_1) A_{\mathbf{q}_2, \lambda}(t_2) \rangle}_{K_N^1}, \end{aligned} \quad (40)$$

creation operators for Bosons [170],  $\langle A_{q_1,\lambda}(t_1)A_{q_2,\lambda}(t_2) \rangle = 0$  unless  $q_2 = -q_1$ , therefore, one obtains

$$K_N^1 = \sum_{q_1,\lambda} e^{iq_1(r_1-r_2)} M_{q_1,\lambda}^2 i\hbar D_{\lambda_0}(q_1, t_1, t_2), \quad (41)$$

where  $D_{\lambda_0}(q_1, t_1, t_2)$  is the noninteracting phonon Green's function (see Appendix B). The lowest-order self-energies owing to electron-phonon interaction are also referred to as Hartree and Fock self-energy by analogy to the treatment of the electron-electron interaction. However, the Hartree self-energy owing to electron-phonon interaction is zero since it corresponds to phonons with  $q = 0$ , but such phonon is either a translation of the crystal or a permanent strain, and neither of these meant to be in the Hamiltonian. The analytical expression regarding the contribution of the self-consistent Fock self-energy (Fig. 10) is given by [95]

$$\sum_{\text{el-ph}}(r_1, t_1; r_2, t_2) = i\hbar \sum_{q_1,\lambda} e^{iq_1(r_1-r_2)} M_{q_1,\lambda}^2 G(r_1, t_1; r_2, t_2) D_{\lambda}(q_1; t_1, t_2). \quad (42)$$

### 3.6. Analytical Continuation

The contour representation is rather impractical in calculations, and one prefers to work with real-time integrals. The procedure of converting the contour into real-time integrals is called analytic continuation [164]. We followed here the formulation by Langreth [178]. In this section we are only concerned with temporal variables, therefore, spatial variables have been suppressed.

#### 3.6.1. Real-Time Formalism

The contour  $C_K$  depicted in Figure 7 consists of two branches,  $C_1$  and  $C_2$ . Each of the time arguments of Green's function can reside either on the first or second part of the contour. Therefore, contour-ordered Green's function thus contains four different Green's functions

$$G(t, t') = \begin{cases} G^>(t, t') & t \in C_2, t' \in C_1 \\ G^<(t, t') & t \in C_1, t' \in C_2 \\ G_t(t, t') & t, t' \in C_1 \\ G_{\bar{t}}(t, t') & t, t' \in C_2 \end{cases}. \quad (43)$$

The *greater* ( $G^>$ ), *lesser* ( $G^<$ ), *time-ordered* ( $G_t$ ), and *antitime-ordered* ( $G_{\bar{t}}$ ) Green's functions can be defined as

$$\begin{aligned} G^>(t, t') &= -i\hbar^{-1} \langle \hat{\psi}_H(t) \hat{\psi}_H^\dagger(t') \rangle, \\ G^<(t, t') &= +i\hbar^{-1} \langle \hat{\psi}_H^\dagger(t') \hat{\psi}_H(t) \rangle, \\ G_t(t, t') &= -i\hbar^{-1} \langle T_t \{ \hat{\psi}_H(t) \hat{\psi}_H^\dagger(t') \} \rangle, \\ &= -\theta(t-t') i\hbar^{-1} \langle \hat{\psi}_H(t) \hat{\psi}_H^\dagger(t') \rangle \\ &\quad + \theta(t'-t) i\hbar^{-1} \langle \hat{\psi}_H^\dagger(t') \hat{\psi}_H(t) \rangle, \\ &= +\theta(t-t') G^>(t, t') + \theta(t'-t) G^<(t, t'), \end{aligned}$$

$$\begin{aligned} G_{\bar{t}}(t, t') &= -i\hbar^{-1} \langle T_{\bar{t}} \{ \hat{\psi}_H(t) \hat{\psi}_H^\dagger(t') \} \rangle, \\ &= -\theta(t'-t) i\hbar^{-1} \langle \hat{\psi}_H(t) \hat{\psi}_H^\dagger(t') \rangle \\ &\quad + \theta(t-t') i\hbar^{-1} \langle \hat{\psi}_H^\dagger(t') \hat{\psi}_H(t) \rangle, \\ &= +\theta(t'-t) G^>(t, t') + \theta(t-t') G^<(t, t'), \end{aligned} \quad (44)$$

where the time-ordering operator  $T_t$  is defined in (154). The antitime-ordering operator  $T_{\bar{t}}$  can be defined in a similar manner. Since  $G_t + G_{\bar{t}} = G^> + G^<$ , there are only three linearly independent functions. The freedom of choice reflects itself in the literature, where a number of different conventions can be found. For our purpose the most suitable functions are the  $G^>$ , and the retarded ( $G^r$ ) and advanced ( $G^a$ ) Green's functions defined as

$$\begin{aligned} G^r(t, t') &= +\theta(t-t') [G^>(t, t') - G^<(t, t')], \\ G^a(t, t') &= +\theta(t-t') [G^<(t, t') - G^>(t, t')]. \end{aligned} \quad (45)$$

It is straightforward to show that  $G^r - G^a = G^> - G^<$ .

The next step is replacing contour by real-time integrals in the Dyson equation. In that equation one encounters the following contour integrals:

$$D(t, t') = \int_C d\tau A(t, \tau) B(\tau, t'), \quad (46)$$

and their generalizations involving products of three or more terms. For that purpose one can employ the Langreth theorem [178] (see Table 2).

### 3.7. Quantum Kinetic Equations

In this section the equations of motion for the NEGFs are introduced. There are two different, but equivalent, formulations: the Kadanoff-Baym and the Keldysh

**Table 2.** Rules for analytic continuation derived from the Langreth theorem.

Contour	Real axis
$D = \int_C AB$	$D^> = \int_t [A^> B^> + A^< B^a]$ $D^r = \int_t A^r B^r$
$D = \int_C ABC$	$D^> = \int_t [A^r B^r C^> + A^r B^< C^a + A^< B^a C^a]$ $D^r = \int_t A^r B^r C^r$
$D(\tau, \tau') = A(\tau, \tau') B(\tau, \tau')$	$D^>(t, t') = A^>(t, t') B^>(t, t')$ $D^r(t, t') = A^<(t, t') B^r(t, t') + A^r(t, t') \times B^<(t, t') + A^r(t, t') B^r(t, t')$

formulation. These are treated in the following subsections. Finally, kinetic equations for steady-state condition are presented.

### 3.7.1. The Kadanoff–Baym Formulation

The starting point of the derivation is the differential form of the Dyson equation. By assuming that  $[i\hbar\partial_{t_1} - H_0(1)]G_0(12) = \delta_{1,2}$ , eqns (36) and (37) can be rewritten as [176]

$$[+i\hbar\partial_{t_1} - \hat{H}_0(1)]G(12) = \delta_{1,2} - i\hbar \int_C d3 \Sigma(13)G(32), \quad (47)$$

$$[-i\hbar\partial_{t_2} - \hat{H}_0(2)]G(12) = \delta_{1,2} - i\hbar \int_C d3 G(13)\Sigma(32). \quad (48)$$

Note that the singular part of the self-energy on the contour, which corresponds to the Hartree self-energy (Section 3.5), does not appear explicitly in the kinetic equations, but is included in the potential energy of the single-particle Hamiltonian  $\hat{H}_0$ .

Using the Langreth rules (Table 2) and fixing the time arguments of Green's functions in (47) and (48) at opposite sides of the contour, one obtains the Kadanoff–Baym equations [76, 176]

$$[+i\hbar\partial_{t_1} - \hat{H}_0(1)]G^<(12) = \int d3 \Sigma^<(13)G^<(32) + \int d3 \Sigma^<(13)G^a(32), \quad (49)$$

$$[-i\hbar\partial_{t_2} - \hat{H}_0(2)]G^<(12) = \int d3 G^r(13)\Sigma^<(32) + \int d3 G^<(13)\Sigma^a(32). \quad (50)$$

The Kadanoff–Baym equations determine the time evolution of Green's functions, but they do not determine the consistent initial values. This information is contained in the original Dyson equations (36) and (37), and lost in the derivation. To have a closed set of equations, the Kadanoff–Baym equations must be supplemented with Dyson equations for  $G^r$  and  $G^a$ . By subtracting (50) from (49), one finds the equation satisfied by  $G^r$  [176]

$$[+i\hbar\partial_{t_1} - \hat{H}_0(1)]G^r(12) - \int d3 \Sigma^r(13)G^r(32) = \delta_{1,2}, \quad (51)$$

$$[-i\hbar\partial_{t_2} - \hat{H}_0(2)]G^r(12) - \int d3 \Sigma^r(13)G^r(32) = \delta_{1,2}, \quad (52)$$

Similar relations hold for the advanced Green's functions.

### 3.7.2. The Keldysh Formulation

For certain applications in classical transport theory it is advantageous to write the Boltzmann equation as an integral equation, rather than an integrodifferential equation. An analogous situation holds in quantum kinetics. Instead of working with the Kadanoff–Baym equations (49) and (50), it may be useful to consider their integral forms. Historically, Keldysh [174] derived his alternative form almost

simultaneously and independently of Kadanoff and Baym. However, the Keldysh and Kadanoff–Baym formalisms are equivalent.

By applying Langreth's rules to the Dyson equation (36) one obtains

$$G^< = G_0^< + G_0^r \Sigma^r G^< + G_0^r \Sigma^< G^a + G_0^< \Sigma^a G^a. \quad (53)$$

For convenience, a notation where a product of two terms is interpreted as a matrix product in the internal variables (space, time, etc.) has been used. One can proceed by iteration with respect to  $G^<$ . Iterating once and regrouping the terms, one obtains

$$G^< = (1 + G_0^r \Sigma^r) G_0^< (1 + \Sigma^a G^a) + (G_0^r + G_0^r \Sigma^r G_0^r) \Sigma^< G^a + G_0^r \Sigma^r G_0^r \Sigma^r G^<. \quad (54)$$

The form of (54) suggests that infinite-order iterations result in [164]

$$G^< = (1 + G^r \Sigma^r) G_0^< (1 + \Sigma^a G^a) + G^r \Sigma^< G^a. \quad (55)$$

Eqn (55) is equivalent to Keldysh's results. In the original work; however, it was written for another function,  $G_K \equiv G^< + G^>$ . This difference is only of minor significance [164].

The first term on the right-hand side of (55) accounts for the initial conditions. One can show that this term vanishes for steady-state systems, if the system was in a noninteracting state in the infinite past [164]. Thus, in many applications it is sufficient to only keep the second term.

Similar steps can be followed to obtain the kinetic equation for  $G^>$ . In integral form these equations can be written as

$$G^>(12) = \int d3 \int d4 G^r(13) \Sigma^>(34) G^a(42). \quad (56)$$

The relation between the Keldysh equation and the Kadanoff–Baym equation is analogous to the relation between an ordinary differential equation plus a boundary condition and the corresponding integral equation.

### 3.7.3. Steady-State Kinetic Equations

In the steady state, Green's functions depend on time differences. One usually Fourier transforms the time difference coordinate,  $\tau = t - t'$ , to energy,

$$G(\mathbf{r}_1, \mathbf{r}_2; E) = \int \frac{d\tau}{\hbar} e^{iE\tau/\hbar} G(\mathbf{r}_1, \mathbf{r}_2; \tau). \quad (57)$$

Under steady-state condition the quantum kinetic equations, (51), (52), and (56), can be written as [165]:

$$[E - \hat{H}_0(r_1)]G^r(r_1, r_3; E) - \int d\mathbf{r}_3 \Sigma^r(r_1, r_3; E)G^r(r_3, r_2; E) = \delta_{r_1, r_2}, \quad (58)$$

$$G^>(\mathbf{r}_1, \mathbf{r}_2; E) = \int d\mathbf{r}_3 \int d\mathbf{r}_4 G^r(\mathbf{r}_1, \mathbf{r}_3; E) \Sigma^<(\mathbf{r}_3, \mathbf{r}_4; E) G^a(\mathbf{r}_4, \mathbf{r}_2; E), \quad (59)$$

where  $\Sigma$  is the total self-energy. A similar transformation can be applied to self-energies. However, to obtain self-energies one has to first apply Langreth's rules and then Fourier transform the time difference coordinate to energy. We consider the self-energies discussed in Section 3.5. The evaluation of the Hartree self-energy owing to electron-electron interaction is straightforward, since it only includes the electron Green's function. However, the lowest-order self-energy owing to electron-phonon interaction contains the products of the electron and phonon Green's functions. Using Langreth's rules (Table 2) and then Fourier transforming the self-energies owing to electron-phonon interaction, (42) takes the form

$$\Sigma_{\text{el-ph}}^>(\mathbf{r}_1, \mathbf{r}_2; E) = i \sum_{\mathbf{q}, \lambda} \int \frac{d(\hbar\omega_{\mathbf{q}, \lambda})}{2\pi} e^{i\mathbf{q}(\mathbf{r}_1 - \mathbf{r}_2)} M_{\mathbf{q}, \lambda}^2 G^>(\mathbf{r}_1, \mathbf{r}_2; E - \hbar\omega_{\mathbf{q}, \lambda}) D_{\lambda}^>(\mathbf{q}, \hbar\omega_{\mathbf{q}, \lambda}). \quad (60)$$

To calculate the retarded self-energy, however, it is more straightforward to Fourier transform the relation  $\Sigma^r(\tau) = \theta(\tau)[\Sigma^>(\tau) - \Sigma^<(\tau)]$ , see (45). By defining the broadening function  $\Gamma$ ,

$$\Gamma(\mathbf{r}_1, \mathbf{r}_2; E) = i[\Sigma^>(\mathbf{r}_1, \mathbf{r}_2; E) - \Sigma^<(\mathbf{r}_1, \mathbf{r}_2; E)] = 2\Im \Sigma^<(\mathbf{r}_1, \mathbf{r}_2; E), \quad (61)$$

the retarded self-energy is given by the convolution of  $-i\Gamma(E)$  and the Fourier transform of the step function [21]

$$\Sigma^r(E) = -i\Gamma(E) \otimes \left( \frac{\delta(E)}{2} + \frac{i}{2\pi E} \right), \quad (62)$$

where  $\otimes$  denotes the convolution. Therefore, the retarded self-energy is given by [98]

$$E^r(\mathbf{r}_1, \mathbf{r}_2; E) = -\frac{i}{2}\Gamma(\mathbf{r}_1, \mathbf{r}_2; E) + P \int \frac{dE'}{2\pi} \frac{\Gamma(\mathbf{r}_1, \mathbf{r}_2; E')}{E - E'}, \quad (63)$$

where  $P$  stands for principal part.

### 3.8. Relation to Observables

Observables such as particle and current densities are directly linked to the greater and lesser Green's functions.

In this section some of the most important observables and their relations to Green's functions are discussed.

#### 3.8.1. Electron and Hole Density

The electron and hole concentration are, respectively, given by

$$n(\mathbf{r}, t) = \langle \hat{\psi}^\dagger(\mathbf{r}, t) \hat{\psi}(\mathbf{r}, t) \rangle = -i\hbar G^<(\mathbf{r}, t; \mathbf{r}, t), \quad (64)$$

$$p(\mathbf{r}, t) = \langle \hat{\psi}(\mathbf{r}, t) \hat{\psi}^\dagger(\mathbf{r}, t) \rangle = +i\hbar G^>(\mathbf{r}, t; \mathbf{r}, t). \quad (65)$$

In the steady state (see Section 3.7.3) these relations can be written as [165]

$$n(\mathbf{r}) = -i \int \frac{dE}{2\pi} G^<(\mathbf{r}, E). \quad (66)$$

$$p(\mathbf{r}) = +i \int \frac{dE}{2\pi} G^>(\mathbf{r}, E) \quad (67)$$

The total space charge density is given by

$$\varrho(\mathbf{r}) = q(p(\mathbf{r}) - n(\mathbf{r})). \quad (68)$$

#### 3.8.2. Spectral Function and Local Density of States

The spectral function is defined as

$$A(\mathbf{r}, \mathbf{r}'; E) = i[G^r(\mathbf{r}, \mathbf{r}'; E) - G^a(\mathbf{r}, \mathbf{r}'; E)]. \quad (69)$$

The spectral function provides information about the nature of the allowed electronic states, regardless of whether they are occupied or not, and can be considered as a generalized density of states. The diagonal elements of the spectral function give the local density of states

$$\rho(\mathbf{r}; E) = \frac{1}{2\pi} A(\mathbf{r}, \mathbf{r}; E) = -\frac{1}{\pi} \Im [G^r(\mathbf{r}, \mathbf{r}; E)]. \quad (70)$$

#### 3.8.3. Current Density

To derive an equation for the current density, one uses the conservation law of quantum mechanical variables [76]. The starting point is the subtraction of eqn (49) from (50)

$$\begin{aligned} i\hbar(\partial_{t_1} + \partial_{t_2})G^<(12) + \frac{\hbar^2}{2m}[(\nabla_{\mathbf{r}_1} + \nabla_{\mathbf{r}_2})(\nabla_{\mathbf{r}_1} + \nabla_{\mathbf{r}_2})]G^<(12) - [U(1) - U(2)]G^<(12) = \\ + \int d3 [\Sigma^r(13)G^<(32) + \Sigma^<(13)G^a(32) + G^r(13)\Sigma^<(32) + G^<(13)\Sigma^a(32)], \end{aligned} \quad (71)$$



where  $H_0(1) = -\hbar^2 / 2m \nabla_1^2 + U(1)$  has been assumed. By taking the limit  $1 \rightarrow 2$  ( $r_2 \rightarrow r_1$  and  $t_2 \rightarrow t_1$ ) and assuming that the right-hand side of (71) approaches zero in this limit, one obtains

$$i\hbar \lim_{t_2 \rightarrow t_1} [\partial_{t_1} G^<(12) + \partial_{t_2} G^<(12)] + \nabla \cdot \left( \frac{\hbar^2}{2m} \lim_{r_2 \rightarrow r_1} (\nabla_{r_1} - \nabla_{r_2}) G^<(12) \right) = 0. \quad (72)$$

By multiplying both sides by  $-q$  and recalling the definition of the charge density, one recovers the continuity equation,

$$\partial_{t_1} \rho(\mathbf{r}_1, t_1) + \nabla \cdot \mathbf{J}(\mathbf{r}_1, t_1) = 0, \quad (73)$$

where the current density is defined as

$$\mathbf{J}(\mathbf{r}_1, t_1) = -\frac{i\hbar^2 q}{2m} \lim_{r_2 \rightarrow r_1} (\nabla_{r_1} - \nabla_{r_2}) G^<(\mathbf{r}_1, t_1; \mathbf{r}_2, t_1). \quad (74)$$

In the steady state the current density takes the form [165]

$$J(r_1) = -\frac{i\hbar q}{2m} \int \frac{dE}{2\pi} \lim_{r_2 \rightarrow r_1} (\nabla_{r_1} - \nabla_{r_2}) G^<(r_1, r_2, E). \quad (75)$$

### 3.8.4. Current Conservation

The current is conserved as long as the right-hand side of (71) approaches zero as  $2 \rightarrow 1$

$$\lim_{2 \rightarrow 1} \int d3 \left[ \Sigma^r(13) G^<(32) + \Sigma^<(13) G^a(32) \right] = 0, \quad (76)$$

The current is obviously conserved if there is no interaction, whereas the situation is different in the interacting case. As described in Section 3.5, the interactions are described in terms of appropriate self-energies. However, self-energies can often be obtained approximately only. Therefore, one could choose an approximation which violates the continuity equation, which, of course, is not physical. It is straightforward to show that the approximated self-energy owing to electron-phonon interaction within the self-consistent Born approximation (42) preserves the current continuity.

## 4. IMPLEMENTATION

Novel structures and materials such as decananometer Si bulk MOSFETs, multiple-gate MOSFETs, CNT-FETs, and molecular-based transistors are expected to be introduced to meet the requirements for scaling. A deep understanding of quantum effects in nanoelectronic devices helps to improve their functionality and to develop new device types. For that purpose extensive computer simulations are required.

A multipurpose quantum-mechanical solver, the Vienna Schrödinger-Poisson (VSP) solver, with the aim to aid theoretical as well as experimental research on nanoscale electronic devices, has been developed [179]. VSP is a quantum-mechanical solver for closed as well as open boundary problems. The software is written in C++ using state-of-the-art software design techniques. The chosen software architecture allows one to add new models easily. Critical numerical calculations are performed with stable numerical libraries such as BLAS, LAPACK, and ARPACK. VSP holds a graphical user interface written in JAVA, as well as an XML-based interface. Furthermore, VSP has an open software application interface and can be used within third party simulation environments.

This section describes the implementation of the outlined NEGF formalism into VSP. For an accurate analysis it is essential to solve the coupled system of transport and Poisson equations self-consistently [142]. The discretization of Poisson's equation and the quantum transport equation is studied.

A tight-binding Hamiltonian is used to describe transport phenomena in CNT-FETs. The mode-space transformation used in this work reduces the computational cost considerably. The mode-space approach takes only a relatively small number of transverse modes into consideration. To reduce the computational cost even further, we used the local scattering approximation [95]. In this approximation the scattering self-energy terms are diagonal in coordinate representation. We show that the local approximation is well justified for electron-phonon scattering caused by deformation potential interaction.

We investigate methods of generating adaptive energy grids for the transport equations and their effect on the convergence behavior of the self-consistent iteration. Our results indicate that for accurate and fast convergent simulations, the energy grid must be carefully adapted.

### 4.1. Electrostatic Potential and the Poisson Equation

Planar CNT-FETs constitute the majority of devices fabricated to date, mostly owing to their relative simplicity and moderate compatibility with existing manufacturing technologies. However, coaxial devices (see Fig. 11) are of special interest because their geometry allows for better electrostatic control than their planar counterparts. These devices would exhibit wrap-around gates that maximize capacitive coupling between the gate electrode and the

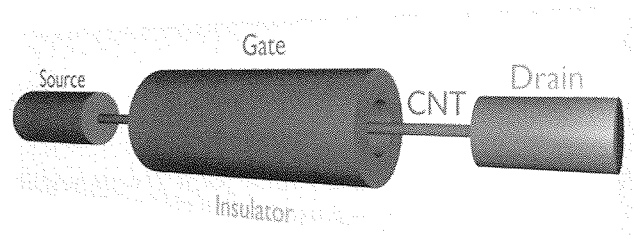


Figure 11. Schematic of the coaxial CNT-FET.

CNT channel. Presently, the closest approximations to this geometry are the electrolyte-gated devices [180, 181]. Alternative structures that place CNTs vertically with respect to the substrate have already been used for field-emission applications [182, 183]. Coaxial CNT-FETs could be fashioned by placing CNTs inside the cavities of a porous material such as alumina, surrounding them by gate electrodes.

Appropriate treatment of the electrostatic potential in the device is essential for accurately predicting the device characteristics [184]. The electron–electron interaction self-energy of lowest order yields the Hartree potential, which is the solution of the Poisson equation,

$$\nabla \cdot \epsilon \nabla \phi = -\rho \quad (77)$$

where  $\rho$  is the total charge density, given (in cylindrical coordinates) by [185]

$$\rho = \frac{q[p(z) - n(z)]\delta(r - r_{\text{CNT}})}{2\pi r}. \quad (78)$$

Here,  $n$  and  $p$  denote the electron and hole concentration per unit length, respectively, and  $r_{\text{CNT}}$  is the radius of the CNT. In (78)  $\delta(r)/r$  is the Dirac delta function in cylindrical coordinates, implying that the carrier density is approximated by a sheet charge distributed along the insulator–CNT interface [185]. Owing to azimuthal symmetry in wrap-around gate devices, the carriers are uniformly distributed as a function of the azimuthal angle.

#### 4.1.1. Discretization of the Poisson Equation

To solve partial differential equations numerically, they are usually discretized. For that reason, the domain  $v$  where the equations are posed has to be partitioned into a finite number of subdomains  $v_i$ , which are usually obtained by a Voronoi tessellation [186]. In order to obtain the solution with a desired accuracy, the equation system is approximated in each of these subdomains by algebraic equations. The unknowns of this system are approximations of the continuous solutions at the discrete grid points in the domain [187]. Several approaches for the discretization of the partial differential equations have been proposed. It has been found to be advantageous to apply the finite boxes discretization scheme for semiconductor device simulation [187]. This method considers the integral form of the equation for each subdomain, which is the so-called control volume  $v_i$  associated with the grid point  $p_i$ . By applying the Gauss integral theorem, the Poisson equation (77) is integrated as

$$\oint_{\partial v} \epsilon \nabla \phi \cdot d\mathbf{A} + \int_v \rho dV = 0. \quad (79)$$

Finally, the discretized equation for point  $i$  with neighbor points  $j$  can be written implicitly as

$$F_i = \sum_j \bar{\epsilon}_{ij} \frac{\phi_j - \phi_i}{d_{ij}} A_{ij} + \rho_i V_i = 0, \quad (80)$$

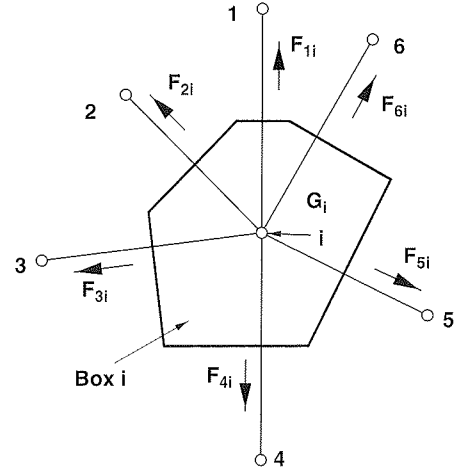


Figure 12. Box for node  $i$  with six neighbor nodes.

with  $d_{ij}$  the distance between grid point  $P_i$  and  $P_j$ ,  $A_{ij}$  the interface area between the domains  $v_i$  and  $v_j$ , and  $V_i$  the volume of the domain  $V_i$ . For position-dependent  $\epsilon$ , one can use here some average, for example  $\bar{\epsilon}_{ij} = (\epsilon_i + \epsilon_j)/2$ . Eqn (80) can be generally written as

$$F_i = \sum_j F_{ji} + G_i = 0, \quad (81)$$

where  $j$  runs over all neighboring grid points in the same segment,  $F_{ji}$  is the flux between points  $j$  and  $i$ , and  $G_i$  is the source term (see Fig. 12).

Grid points on the boundary  $\partial v$  are defined as having neighbor grid points in other segments. Thus, for boundary elements, (81) does not represent the complete control function  $F_i$ , since all the fluxes into the contact or the other segment are missing. For that reason, the information for these boxes has to be completed by taking the boundary conditions into account.

#### 4.1.2. Boundary Conditions

Dirichlet boundary conditions are introduced at the source, drain, and gate contacts. Potentials are conveniently measured relative to the source potential. The amount of bending of the vacuum energy level along the length of the CNT is given by  $E_{\text{vac}}(z) = -q\phi(z)$ , since we assume that the local electrostatic potential rigidly shifts the CNT band structure. The conduction and valence band edges of the CNT are given by

$$\begin{aligned} E_c(z) &= -q\phi(z) + q\Phi_{\text{Be}}, \\ E_v(z) &= -q\phi(z) - q\Phi_{\text{Bh}}. \end{aligned} \quad (82)$$

The SB heights for electrons ( $q\Phi_{\text{Be}}$ ) and holes ( $q\Phi_{\text{Bh}}$ ) at the metal–CNT interface are given by (see Fig. 13)

$$\begin{aligned} q\Phi_{\text{Be}} &= +q\Phi_{\text{M}} - q\chi_{\text{CNT}}, \\ q\Phi_{\text{Bh}} &= -q\Phi_{\text{Be}} + E_g, \end{aligned} \quad (83)$$

where  $q\Phi_{\text{M}}$  is the work function of the metal contact,  $q\chi_{\text{CNT}}$  is the electron affinity of the CNT, and  $E_g$  is the

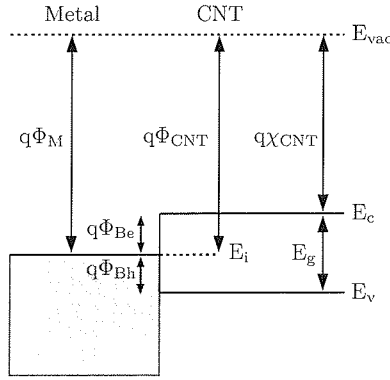


Figure 13. Band diagram at the metal–CNT interface with  $\Phi_M = \Phi_{CNT}$ .

band gap of the CNT. The work function of CNT  $q\Phi_{CNT}$  is defined as the sum of the CNT electron affinity and half of the band gap in the bulk. Figure 13 shows the band diagram at the metal–CNT interface with  $\Phi_M = \Phi_{CNT}$ . The work function of the CNT is assumed to be  $\Phi_{CNT} = 4.5\text{eV}$  [3]. In an intrinsic CNT (undoped) the Fermi level of the CNT is located in the middle of the band gap. Under these conditions, equal SB heights for both electrons and holes are achieved. If the work functions of metal and CNT do not align, band bending near the contact occurs, and the SB heights for electrons and holes will be different. For example, if the work function of the metal contact is larger than that of the CNT ( $\Phi_M > \Phi_{CNT}$ ), the SB height for holes is smaller than that for electrons. As a result, a p-type CNT-FET is achieved, where holes are the majority carriers.

#### 4.1.3. Interface Conditions

To account for interface conditions, grid points located at the boundary of the segments (see Fig. 14(a)) are assigned three values, one for each segment (see Fig. 14(b)) and a third value assigned directly to the interface, which can be used to formulate more complicated interface conditions such as, for example, interface charges.

As discussed in Section 4.1, charges on the CNT are approximated as sheet charges at the CNT–insulator interface. The boundary flux owing to interface charges is simply added to the segment fluxes given by (81)

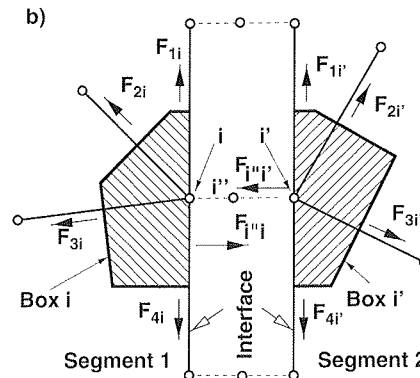
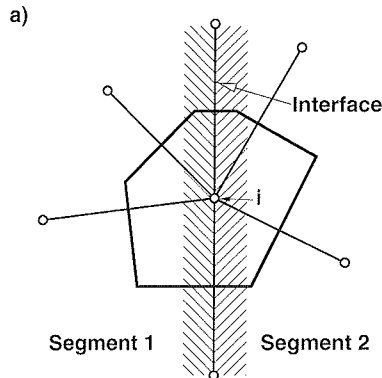


Figure 14. Interface points as given in (a) are split into three different points having the same geometrical coordinates (b).

$$\begin{aligned} F_i + F_{i'',i} &= 0, \\ F_{i'} + F_{i'',i'} &= 0, \end{aligned} \quad (84)$$

with

$$\begin{aligned} F_{i'',i} &= \frac{\rho_{i''} A_i}{2}, \\ F_{i'',i'} &= \frac{\rho_{i''} A_{i'}}{2}, \end{aligned} \quad (85)$$

where  $A_i = A_{i'}$  and  $\rho_{i''}$  is the interface charge density at some point  $i''$  along the insulator–CNT interface given by (78). This method satisfies the condition of the discontinuity of the electric displacement in the presence of interface charges. For the CNT, the free-space relative permittivity,  $\epsilon_{CNT} = 1$ , is assumed [188].

#### 4.2. Basis Functions and Matrix Representation

For the purpose of discretization, one can expand the Hamiltonian, self-energies, and Green's functions in terms of some basis functions to obtain the corresponding matrices. In the *tight-binding method* one can take the basis functions to be any set of localized functions, such as atomic s and p orbitals [98, 103], Wannier functions [189], and so forth. A common approximation used to describe the Hamiltonian of layered structures consists of nonvanishing interactions only between the *nearest-neighbor* layers. That is, each layer  $i$  interacts only with itself and its nearest-neighbor layers  $i-1$  and  $i+1$ . Then the single-particle Hamiltonian of the layered structure is a *block tridiagonal matrix*, where diagonal blocks  $\underline{H}_i$  represent the Hamiltonian of layer  $i$  and off-diagonal blocks  $t_{i,i+1}$  represent the interaction between layers  $i$  and  $i+1$

$$\underline{H} = \begin{bmatrix} \ddots & \ddots & \ddots & \ddots & \ddots \\ \bullet & \underline{H}_1 & t_{-12} & \ddots & \ddots \\ \bullet & t_{-12}^\dagger & \underline{H}_2 & t_{-2,3} & \ddots \\ & \ddots & \ddots & \ddots & \ddots \\ & & t_{-N-2,N-1}^\dagger & \underline{H}_{N-1} & t_{-N-1,N} \\ & & & t_{-N-1,N}^\dagger & \underline{H}_N \\ & & & & \ddots & \ddots \end{bmatrix} \quad (86)$$

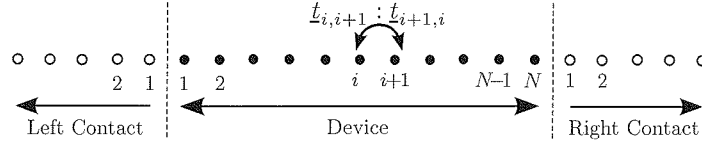


Figure 15. Partitioning of the simulation domain into device region and left and right contacts. Each point corresponds to a layer.

where  $t_{i+1,i} = t_{i,i+1}^\dagger$ . The matrix representation of the kinetic equations (58) and (59) are

$$\left[ EI - \underline{H} - \underline{\Sigma}_{\text{Scat}}^r \right] \underline{G}^r = \underline{I}, \quad (87)$$

$$\underline{G}^< = \underline{G}^r \underline{\Sigma}_{\text{Scat}}^< \underline{G}^a, \quad (88)$$

where  $\underline{\Sigma}_{\text{Scat}}$  is the self-energy owing to scattering processes and  $\underline{G}^a = [\underline{G}^r]$  [165]. One can partition the layered structure into left contact with index  $L$ , device region with index  $D$ , and right contact with index  $R$  (Fig. 15). The device corresponds to the region where one solves the transport equations and the contacts are the highly conducting regions connected to the device. While the device region consists of only  $N$  layers, the matrix equation corresponding to (87) is infinitely dimensional owing to the semiinfinite contacts. As shown in Appendix C the influence of the semiinfinite contacts can be folded into the device region by adding a self-energy to the device region. This can be viewed as an additional self-energy owing to the transitions between the device and the contacts. In the next sections the matrix representations of the Hamiltonian and self-energies are discussed in detail.

### 4.3. Tight-Binding Hamiltonian

The general form of the tight-binding Hamiltonian for electrons in a CNT can be written as

$$\hat{H} = \sum_{i,p} U_i^p c_{i,p}^\dagger c_{i,p} + \sum_{i,j,p,q} t_{i,j}^{p,q} c_{i,p}^\dagger c_{j,q}. \quad (89)$$

The sum is taken over all the rings  $i, j$  along the transport direction, which is assumed to be the  $z$ -direction of the cylindrical coordinate system, and over all the atomic locations  $p, q$  in a ring. We use a nearest-neighbor tight-binding  $\pi$ -bond model [9, 190]. Each atom in an  $sp^2$ -coordinated CNT has three nearest neighbors, located  $a_{\text{C-C}}$  away. The band structure consists of  $\pi$  orbitals only, with the hopping parameter  $t = V_{\text{pp}\pi} = -2.77$  eV and zero on-site potential. Furthermore, it is assumed that the electrostatic potential  $U$  rigidly shifts the on-site potentials. Such a tight-binding model is adequate to model transport properties in undeformed CNTs.

In this work we consider zig-zag CNTs. However, this method can be readily extended to armchair or chiral

CNTs. Within the nearest-neighbor approximation, only the following parameters are nonzero [9]:

$$\begin{aligned} t_{i,i-1}^{p,q} &= t_{i-1,i}^{p,q} = t \delta_{p \pm a/2, q}, \quad \forall i = 2k, \\ t_{i,i+1}^{p,q} &= t_{i+1,i}^{p,q} = t \delta_{p, q}, \quad \forall i = 2k. \end{aligned} \quad (90)$$

Figure 16 shows that a zig-zag CNT is composed of rings (layers) of A- and B-type carbon atoms, where  $A$  and  $B$  represent the two carbon atoms in a unit cell of graphene. Each A-type ring is adjacent to a B-type ring. Within nearest-neighbor tight-binding approximation, the total Hamiltonian matrix is block tridiagonal [190]

$$\underline{H} = \begin{bmatrix} \underline{H}_1 & t_2 & & & \\ t_2^\dagger & \underline{H}_2 & t_1 & & \\ & t_1 & \underline{H}_3 & t_2^\dagger & \\ & & t_2 & \underline{H}_4 & t_1 \\ & & & t_1 & \underline{H}_5 & \ddots \\ & & & & \ddots & \ddots \end{bmatrix}, \quad (91)$$

where the diagonal blocks,  $\underline{H}_i$ , describe the coupling within an A- or B-type carbon ring and off-diagonal blocks,  $t_1$  and  $t_2$ , describe the coupling between the adjacent rings. It should

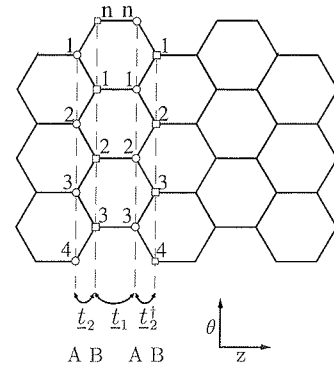
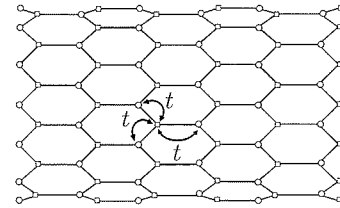


Figure 16. Layer layout of a  $(n, 0)$  zig-zag CNT. Circles are rings of A-type carbon atoms and squares rings of B-type carbon atoms. The coupling coefficient between nearest-neighbor carbon atoms is  $t$ . The coupling matrices between rings are denoted by  $t_1$  and  $t_2$ , where  $t_1$  is a diagonal matrix and  $t_2$  is nondiagonal.

be noted that the odd-numbered Hamiltonians  $\underline{H}_i$  refer to A-type rings and the even numbered ones to B-type rings. Each A-type ring couples to the next B-type ring according to  $t_2$  and to the previous B-type ring according to  $t_1$ . Each B-type ring couples to the next A-type ring according to  $t_1$  and to the previous A-type ring according to  $t_2$ . In a  $(n,0)$  zig-zag CNT, there are  $n$  carbon atoms in each ring, thus, all the submatrices in (91) have a size of  $n \times n$ .

In the nearest-neighbor tight-binding approximation, carbon atoms within a ring are not coupled to each other so that  $\underline{H}_i$  is a diagonal matrix. The value of a diagonal entry is the potential at that carbon atom site. In the case of a coaxially gated CNT, the potential is constant along the CNT circumference. As a result, the submatrices  $\underline{H}_i$  are given by the potential at the respective carbon ring times the identity matrix,

$$\underline{H}_i = \underline{U}_i = U_i \underline{I}, \quad (92)$$

There are two types of coupling matrices between the nearest carbon rings,  $t_1$  and  $t_2$ . As shown in Figure 16, the first type,  $t_1$ , only couples an A(B) carbon atom to its B(A) counterpart in the neighboring ring. The coupling matrix is just the tight-binding coupling parameter times an identity matrix,

$$t_1 = t \underline{I}. \quad (93)$$

The second type of coupling matrix,  $t_2$ , couples an A(B) atom to two B(A) neighbors in the adjacent ring. The coupling matrix is

$$t_2 = \begin{bmatrix} t & & t \\ t & t & \\ & t & t \\ & & \bullet \bullet \end{bmatrix}. \quad (94)$$

The period of the zig-zag CNT in the longitudinal direction contains four rings, ABAB, and has a length of  $3a_{C-C}$ . Therefore, the average distance between the rings is

$$\Delta z = \frac{3a_{C-C}}{4}. \quad (95)$$

#### 4.4. Mode-Space Transformation

A mode-space approach significantly reduces the size of the Hamiltonian matrix [8]. Owing to the quantum confinement along the CNT circumference, circumferential modes appear, and transport can be described in terms of these modes. If  $M$  modes contribute to transports, and if  $M < n$ , then the size of the problem is reduced from  $n \times N$  to  $M \times N$ , where  $N$  is the number of carbon rings along the CNT. If the potential profile does not vary sharply along the CNT, subbands are decoupled [8], and one can solve  $M$  one-dimensional problems of size  $N$ .

Mathematically, one performs a basis transformation on the Hamiltonian of the  $(n,0)$  zig-zag CNT to decouple the problem into  $n$  one-dimensional mode-space lattices [190]

$$\underline{H}' = \begin{bmatrix} \underline{S}^\dagger & & & & \\ & \underline{S}^\dagger & & & \\ & & \underline{S}^\dagger & & \\ & & & \underline{S}^\dagger & \\ & & & & \underline{S}^\dagger \end{bmatrix} \begin{bmatrix} \underline{U}_1 & t_2 & & & \\ t_2^\dagger & \underline{U}_2 & t_1 & & \\ & t_1 & \underline{U}_3 & t_2^\dagger & \\ & & t_2 & \underline{U}_4 & t_1 \\ & & & t_1 & \underline{U}_5 & \bullet \\ & & & & \bullet & \bullet \end{bmatrix} \quad (96)$$

$$\begin{bmatrix} \underline{S} & & & & \\ & \underline{S} & & & \\ & & \underline{S} & & \\ & & & \underline{S} & \\ & & & & \underline{S} \end{bmatrix} = \begin{bmatrix} \underline{U}_1' & t_2' & & & \\ t_2'^\dagger & \underline{U}_2' & t_1' & & \\ & t_1' & \underline{U}_3' & t_2'^\dagger & \\ & & t_2' & \underline{U}_4' & t_1' \\ & & & t_1' & \underline{U}_5' & \bullet \\ & & & & \bullet & \bullet \end{bmatrix},$$

with

$$\begin{aligned} \underline{U}_i' &= \underline{S}^\dagger \underline{U}_i \underline{S}, \\ t_1' &= \underline{S}^\dagger t_1 \underline{S}, \\ t_2' &= \underline{S}^\dagger t_2 \underline{S}, \end{aligned} \quad (97)$$

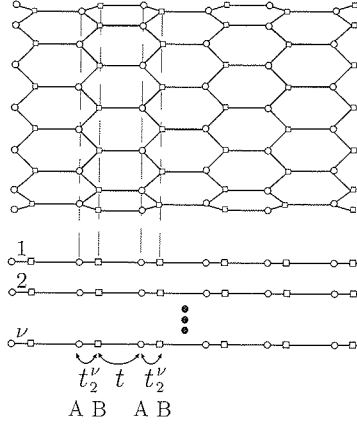
where  $\underline{S}$  is the transformation matrix from the real-space basis to the mode-space basis. The purpose is to decouple the modes after the basis transformation, that is, to make the Hamiltonian matrix blocks between the different modes equal to zero. This requires that after the transformation, the matrices  $\underline{U}_i$ ,  $t_1$ , and  $t_2$  become diagonal. Since  $\underline{U}_i$  and  $t_1$  are identity matrices multiplied by a constant, they remain unchanged and diagonal after any basis transformation,  $\underline{U}_i = \underline{U}_i'$  and  $t_1 = t_1'$ . To diagonalize  $t_2$ , elements of the transformation matrix  $\underline{S}$  have to be the eigenvectors of  $t_2$ . These eigenvectors are plane waves with wave vectors satisfying the periodic boundary condition around the CNT. The eigenvalues are

$$t_2^v = 2te^{-i\pi v/n} \cos(\pi v/n), \quad (98)$$

where  $v = 1, 2, \dots, n$  [190]. The phase factor in (98) has no effect on the results such as charge and current density, thus it can be omitted and  $t_2^v = 2t \cos(\pi v/n)$  can be used instead.

After the basis transformation all submatrices,  $\underline{U}_i$ ,  $t_1$ , and  $t_2$  are diagonal. By reordering the basis according to the modes, the Hamiltonian matrix takes the form

$$\underline{H}' = \begin{bmatrix} \underline{H}^1 & & & & \\ & \underline{H}^2 & & & \\ & & \bullet & & \\ & & & \underline{H}^v & \\ & & & & \bullet \end{bmatrix}, \quad (99)$$



**Figure 17.** Zig-zag CNT and the corresponding one-dimensional chain with two sites per unit cell with hopping parameters  $t$  and  $t_2^\nu = 2t \cos(\pi\nu/n)$ .

where  $\underline{H}^\nu$  is the Hamiltonian matrix for the  $\nu$ th mode [190]

$$\underline{H}^\nu = \begin{bmatrix} U_1 & t_2^\nu & & & & \\ t_2^\nu & U_2 & t & & & \\ & t & U_3 & t_2^\nu & & \\ & & t_2^\nu & U_4 & t & \\ & & & t & U_5 & \ddots \\ & & & & \ddots & \ddots \end{bmatrix}. \quad (100)$$

The one-dimensional tight-binding Hamiltonian  $H^\nu$  describes a chain of atoms with two sites per unit cell and on-site potential  $U$  and hopping parameters  $t$  and  $t_2^\nu$  (Fig. 17). The spatial grid used for device simulation corresponds to the circumferential rings of carbon atoms. Therefore, the ranks of the matrices for each subband are equal to the total number of these rings  $N$ . Self-energies can also be transformed into mode-space  $\Sigma^\nu$ , see Sections 4.5 and 4.6. Green's functions can therefore be defined for each subband (mode) and one can solve the system of transport equations for each subband independently

$$\left[ E\underline{I} - \underline{H}^\nu - \underline{\Sigma}^{r\nu} \right] \underline{G}^{r\nu} = \underline{I}, \quad (101)$$

$$\underline{G}^{<\nu} = \underline{G}^{r\nu} \underline{\Sigma}^{<\nu} \underline{G}^{a\nu}. \quad (102)$$

## 4.5. Contact Self-Energies

Boundary conditions have to be specified to model the contacts, which act as a source or drain for electrons. While the matrix representing the Hamiltonian of the device has a finite dimension, the total Hamiltonian matrix is infinite dimensional owing to the semi-infinite contacts.

The influence of the contacts can be folded into the device region. Owing to the transitions between the device and the contacts, the influence of the contacts can be demonstrated by adding contact self-energies to the total self-energy

[165]. The self-energy matrices for the contacts and the Hamiltonian matrix for the device have the same rank, but the self-energy matrices are highly sparse. For example, only one carbon ring at the source end of the channel couples to the source, thus only one submatrix is nonzero for the source self-energy. Similarly, only one submatrix is nonzero for the drain self-energy. As shown in Appendix C, nonzero blocks of the contact self-energies are given by:

$$\underline{\Sigma}_L^r = t_{LD}^\dagger \underline{g}_{L,1}^r t_{LD}, \quad (103)$$

$$\underline{\Sigma}_R^r = t_{RD}^\dagger \underline{g}_{R,1}^r t_{RD},$$

$$\underline{\Sigma}_L^< = +i\underline{\Gamma}_L f_L, \quad (104)$$

$$\underline{\Sigma}_R^< = +i\underline{\Gamma}_R f_R,$$

$$\underline{\Sigma}_L^> = -i\underline{\Gamma}_L (1 - f_L), \quad (105)$$

$$\underline{\Sigma}_R^> = -i\underline{\Gamma}_R (1 - f_R),$$

where subscripts  $L$  and  $R$  denote the left (source) and right (drain) contacts, respectively,  $f_{L,R}$  the Fermi factor of the contacts,  $\underline{g}_{L,R}^r$  the surface Green's function of the contacts, and finally the broadening functions are defined as

$$\begin{aligned} \underline{\Gamma}_L &= i(\underline{\Sigma}_L^r - \underline{\Sigma}_L^a) = -2\Im[\underline{\Sigma}_L^r], \\ \underline{\Gamma}_R &= i(\underline{\Sigma}_R^r - \underline{\Sigma}_R^a) = -2\Im[\underline{\Sigma}_R^r]. \end{aligned} \quad (106)$$

Surface Green's functions can be calculated using a recursive relation described in Appendix C.3. In this section two types of contacts are discussed: semi-infinite CNTs acting as the source and drain contacts and Schottky-type metal-CNT contacts. The respective surface Green's functions and self-energies for the both contact types are derived next.

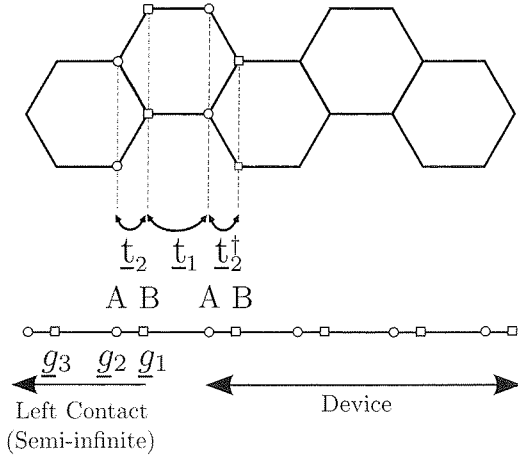
In mode-space representation (see Section 4.4), the matrices in (103)–(106) become one-dimensional. Thus, the respective quantities for each mode can be treated as numbers, and the computational cost decreases considerably.

### 4.5.1. Semi-Infinite CNT Contacts

Figure 18 shows carbon rings of A or B type coupled to a semi-infinite CNT acting as a contact. Each circle (rectangle) represents a carbon ring consisting of A- or B-type carbon atoms. The carbon ring couples to the nearest ring, with a coupling matrix of  $t_1$  or  $t_2$ , and  $\underline{g}_{L,i}^r$  is the surface Green's function for the  $i$ th ring in the left extension, ordered from the channel-contact interface. The recursive relation (196) can be applied to the CNT in Figure 18 and gives

$$\begin{aligned} \left[ \underline{A}_{L_1} - t_2 \underline{g}_{L,2}^r t_2^\dagger \right] \underline{g}_{L,1}^r &= \underline{I}, \\ \left[ \underline{A}_{L_2} - t_1 \underline{g}_{L,3}^r t_1^\dagger \right] \underline{g}_{L,2}^r &= \underline{I}, \end{aligned} \quad (107)$$

where  $\underline{A}_{L_i} = E\underline{I}_i - \underline{U}_{L_i} - \underline{\Sigma}_{\text{Scat},i}^r$  (see Appendix C.1), and  $t_1$  and  $t_2$  are given by (93) and (94), respectively. Since the potential is invariant inside the contact,  $\underline{A}_{L_1} = \underline{A}_{L_2}$ . Further-



**Figure 18.** Computing the surface Green's function for the left contact. The surface Green's function for the  $i$ th ring inside the contact is  $g_i$ .

more,  $\underline{g}_{L_{3,3}}^r = \underline{g}_{L_{1,1}}^r$  owing to the periodicity of the CNT lattice. Using these relations, (107) represents two coupled matrix equations with two unknowns,  $\underline{g}_{L_{1,1}}^r$  and  $\underline{g}_{L_{2,2}}^r$ , which can be solved by iteration. However, in mode-space representation matrices,  $t_1$  and  $t_2$  are replaced by the numbers  $t_1 = t$  and  $t_2^v$ , respectively. As a result, the surface Green's function for each mode can be calculated analytically by solving a quadratic equation,

$$g_{L_{1,1}}^{rv} = \frac{A_{L_1}^2 + t_1^2 - t_2^{v^2} - \sqrt{[A_{L_1}^2 + t_1^2 - t_2^{v^2}]^2 - 4A_{L_1}^2 t_1^2}}{2A_{L_1} t_1^2}. \quad (108)$$

The self-energy of the left contact for the  $v$ th mode is therefore given by

$$\Sigma_L^{rv} = t_1^2 g_{L_{1,1}}^{rv} \quad (109)$$

A similar relation holds for the right contact self-energy.

#### 4.5.2. Schottky-Type Metal–CNT Contacts

At the metal–CNT interface, a Schottky barrier (SB) forms, which governs the operation of CNT-FETs (Section 2.4.2). The metal region acts as a source and a sink of electrons in the device region. In this work Pd contacts are assumed. For transport calculation based on a simplified tight-binding Hamiltonian, describing only the interaction between  $d_z$  orbitals of Pd and  $p_z$  orbitals of the carbon atoms, the self-energy for this SB contact can be written as

$$\Sigma_{SB}^{rv} = t_{M-C}^2 g_{M_{1,1}}^r \quad (110)$$

where  $t_{M-C}$  is the hopping parameter between the metal and the carbon atoms and  $g_{M_{1,1}}^r$  is the surface Green's function of the metal contact. The contact model in (110) assumes injection from the contact into all the CNT subbands.

Based on *ab initio* calculations, it has been shown that the electronic band structure of the Pd–graphene system near the Fermi level can be reproduced by considering the hybridization between the graphene and the Pd bands, using  $t_{Pd-C} = 0.15\text{eV}$  [191].

The surface Green's function contains information about the band structure of the metal contact. To calculate the surface Green's function, one has to specify an appropriate Hamiltonian for the contacts, for example, one can employ the tight-binding method [135], density functional theory [191], or extended Hückel theory [192]. Contacts can be approximated as semi-infinite leads along the transport axis and infinite in the transverse direction. Therefore, the surface Green's function can be calculated iteratively along the transport direction (Section C.3).

#### 4.6. Scattering Self-Energies

The lowest-order electron–phonon self-energies are introduced in Section 3.5, and the steady-state form of these self-energies has been derived in Section 3.7.3. By transforming the self-energies (60) into mode-space, one obtains [9]

$$\begin{aligned} \Sigma_{\text{el-ph},i,j}^{>v}(E) = & i \sum_{q,\lambda} \int \frac{d(\hbar\omega_{q,\lambda})}{2\pi} e^{iq(z_i - z_j)} M_{q,\lambda,v}^2 \\ & \times G_{i,j}^{>v}(E - \hbar\omega_{q,\lambda}) D_{\lambda}^{>}(q, \hbar\omega_{q,\lambda}), \end{aligned} \quad (111)$$

where  $z_i$  is the position of some lattice point  $i$  along the CNT axis. Note that owing to the one-dimensional nature of CNTs, the coordinate and wave vector variables are all one-dimensional. In (111) only intrasubband scattering processes are considered (see Section 2.3.3). To include intersubband scattering processes, the summation in (111) would have to run over all the subbands  $v'$  with the electron–phonon matrix elements  $M_{q,\lambda,v,v'}$ .

The electron–phonon self-energies in the self-consistent Born approximation are expressed in terms of the full electron and phonon Green's functions. One should therefore study the influence of the bare electron states on the phonons first and then calculate the effect on the electrons of the renormalized phonon states [170]. In this work we assume that the phonon renormalization can be neglected. By doing so, we miss to capture a possible reduction of the phonon lifetime. The above considerations also appeal to the Migdal theorem [193], which states that the phonon-induced renormalization of the electron–phonon vertex scales with the ratio of the electron mass to the ion mass [169]. Therefore, one can assume that the phonon bath is in thermal equilibrium so that the full phonon Green's function  $D_{\lambda}$  can be replaced by the noninteracting Green's functions  $D_{\lambda_0}$  from (173). As a result (111) can be written as

$$\begin{aligned} \Sigma_{\text{el-ph},i,j}^{>v}(E) = & \sum_{q,\lambda} e^{iq(z_i - z_j)} M_{q,\lambda}^2 \\ & \times \left[ (n_B(\hbar\omega_{q,\lambda}) + 1) G_{i,j}^{<v}(E + \hbar\omega_{q,\lambda}) \right. \\ & \left. + n_B(\hbar\omega_{q,\lambda}) G_{i,j}^{<v}(E - \hbar\omega_{q,\lambda}) \right], \end{aligned} \quad (112)$$

$$\begin{aligned} \Sigma_{\text{el-ph}_{i,j}}^{\geq v}(E) = & \sum_{q,\lambda} e^{iq(z_i - z_j)} M_{q,\lambda}^2 \\ & \times \left[ (n_B(\hbar\omega_{q,\lambda}) + 1) G_{i,j}^v(E - \hbar\omega_{q,\lambda}) \right. \\ & \left. + n_B(\hbar\omega_{q,\lambda}) G_{i,j}^v(E + \hbar\omega_{q,\lambda}) \right], \end{aligned} \quad (113)$$

where the first term on the right-hand side is owing to the phonon emission and the second term owing to the phonon absorption. The summation over wave vector  $q$  in (112) and (113) can be generally transformed into an integral over the first Brillouin zone,

$$\sum_q = \frac{L}{2\pi} \int dq, \quad (114)$$

where  $L$  is the normalization length and the limits of the integral are  $\pm 3a_{C-C}$  which is the period of the carbon rings, see (4.3). To calculate electron-phonon self-energies, the integral in (114) must be evaluated.

#### 4.6.1. Scattering with OP

In this section the self-energies owing to the interaction of electrons with optical phonons (OPs) are evaluated. As discussed in Section 2.3.3, the phonon energy and the reduced electron-phonon matrix elements for OP phonons are approximately constant and independent of the phonon wave vector. Under this assumption all the terms except the exponential term in (112) and (113) can be taken out of the integral (114), and one obtains [41]

$$\begin{aligned} & \int_{-\pi/(3a_{CC})}^{\pi/(3a_{CC})} \frac{dq}{2\pi} \exp(iq(z_i - z_j)) \\ & = \begin{cases} 1, & z_i - z_j = 0 \\ 0, & z_i - z_j = k \times 3a_{C-C} \end{cases}, \end{aligned} \quad (115)$$

where  $k$  is an integer number. One has to multiply the above result by a factor of 4, for the number of rings in the lattice period [41]. Eqn (115) justifies the approximation which only considers diagonal elements of the electron-phonon self-energy. As discussed in Section 4.3, by employing the nearest-neighbor tight-binding method (block), tridiagonal matrices are achieved. Keeping only diagonal elements of the electron-phonon self-energy, the matrices remain (block) tridiagonal. Therefore, an efficient recursive method [7] can be used to calculate the inverse matrices. This implies considerable reduction of the computational cost and memory requirement.

Using the result of (115) and the relations (15) and (19), the self-energy owing to scattering with OP can be written as

$$\begin{aligned} \Sigma_{\text{OP}_{i,j}}^{\leq v}(E) = & \delta_{i,j} D_{\text{OP}} \left[ (n_B(\hbar\omega_{\text{OP}}) + 1) G_{i,j}^{\leq v}(E + \hbar\omega_{\text{OP}}) \right. \\ & \left. + n_B(\hbar\omega_{\text{OP}}) G_{i,j}^{\leq v}(E - \hbar\omega_{\text{OP}}) \right], \end{aligned} \quad (116)$$

$$\begin{aligned} \Sigma_{\text{OP}_{i,j}}^{\geq v}(E) = & \delta_{i,j} D_{\text{OP}} \left[ (n_B(\hbar\omega_{\text{OP}}) + 1) G_{i,j}^{\geq v}(E - \hbar\omega_{\text{OP}}) \right. \\ & \left. + n_B(\hbar\omega_{\text{OP}}) G_{i,j}^{\geq v}(E + \hbar\omega_{\text{OP}}) \right], \end{aligned} \quad (117)$$

where  $D_{\text{OP}}$  is given by

$$\begin{aligned} D_{\text{OP}} = & \frac{\hbar}{2\rho_{\text{CNT}} L \omega_{\text{OP}}} \tilde{M}_{\text{OP}}^2 \frac{L}{\Delta z} \\ = & \frac{\hbar}{2nm_C \omega_{\text{OP}}} \tilde{M}_{\text{OP}}^2, \end{aligned} \quad (118)$$

where  $\Delta z = 3a_{C-C}/4$  (see (4.3)). In the second line in (118), the mass density of a  $(n, 0)$  zig-zag CNT has been replaced  $\rho_{\text{CNT}} = nm_C/\Delta z$ , where  $m_C$  is the mass of a carbon atom.

The retarded self-energy can be calculated as (63)

$$\Sigma_{\text{OP}_{i,j}}^r(E) = -\frac{i}{2} \Gamma_{\text{OP}_{i,j}}(E) + P \int \frac{dE'}{2\pi} \frac{\Gamma_{\text{OP}_{i,j}}(E')}{E - E'}, \quad (119)$$

where

$$\begin{aligned} \Gamma_{\text{OP}_{i,j}}(E) = & i \left[ \Sigma_{\text{OP}_{i,j}}^{\geq}(E) - \Sigma_{\text{OP}_{i,j}}^{\leq}(E) \right] \\ = & 2\Im \left[ \Sigma_{\text{OP}_{i,j}}^{\leq}(E) \right]. \end{aligned} \quad (120)$$

Since the lesser and greater self-energies are assumed to be diagonal, the retarded self-energy is also diagonal.

#### 4.6.2. Scattering with Acoustic Phonons

Interaction with acoustic phonons (APs) can be approximated as elastic scattering,  $E \pm \hbar\omega_{\text{AP}} \approx E$ . As discussed in Section 2.3.3, near the  $\Gamma$  point a linear dispersion relation for acoustic phonons is assumed,  $\omega_{\text{AP}}(q) \approx v_{\text{AP}}|q|$ , where  $v_{\text{AP}}$  is the acoustic phonon velocity. Furthermore, at room temperature low-energy phonons have an appreciable occupation, such that

$$n_B \approx n_B + 1 \approx \frac{k_B T}{\hbar v_{\text{AP}} |q|} \gg 1. \quad (121)$$

With eqn (121) and the elastic approximation, the contributions owing to phonon emission and absorption become equal and can be lumped into one term. As a result, by using the relations (15) and (19), the self-energies owing to acoustic-phonon interaction are written as

$$\Sigma_{\text{AP}_{i,j}}^{\geq v}(E) = \sum_q e^{iq(z_i - z_j)} \frac{\hbar}{2\rho_{\text{CNT}} L v_{\text{AP}} |q|} \tilde{M}_{\text{AP}}^2 |q|^2 \times 2 \frac{k_B T}{\hbar v_{\text{AP}} |q|} G_{i,j}^{\geq v}(E) = \sum_q e^{iq(z_i - z_j)} \frac{k_B T}{\rho_{\text{CNT}} L v_{\text{AP}}^2} \tilde{M}_{\text{AP}}^2 G_{i,j}^{\geq v}(E). \quad (122)$$



With the exception of the exponential term, all terms in (122) can be taken out of the sum and one can convert the sum into an integral over  $q$ , see (114) and (115). The self-energies simplify to

$$\Sigma_{AP,i,j}^{>v}(E) = \delta_{i,j} D_{AP} G_{i,j}^{>v}(E), \quad (123)$$

where similar to (118),  $D_{AP}$  is given by

$$D_{AP} = \frac{k_B T}{\rho_{CNT} L v_{AP}^2} \sim \frac{2}{M_{AP}} \frac{L}{\Delta z} \\ = \frac{k_B T}{nm_C v_{AP}^2} \sim \frac{2}{M_{AP}}. \quad (124)$$

A discussion similar to that in Section 4.6.1 gives a justification to keep only diagonal elements of the self-energy owing to the interaction of electrons with acoustic phonons.

By substituting (123) in (63), the retarded self-energy is obtained as

$$\Sigma_{AP,i,j}^r(E) = \delta_{i,j} D_{AP} G_{i,j}^r(E), \quad (125)$$

Owing to the approximations made, the retarded self-energy for scattering with acoustic phonons is simplified and directly related to the retarded Green's function. Therefore, one does not need to evaluate the integrals such as (119), which implies a considerable saving of computational cost.

## 4.7. Evaluation of Observables

To solve the Poisson equation in a self-consistent scheme, one has to know the carrier density profile in the device. To study the device characteristics, the current through the device needs to be calculated. In this section the numerical evaluation of these two observables is discussed.

### 4.7.1. Carrier Density

Green's function matrices  $G_{i,j}^{>v}$  are defined in the basis set of ring numbers  $i, j$  and subbands  $v$ . Thus, the diagonal elements correspond to the spectrum of carrier occupation (66) of those basis sites with a given energy  $E$ . So the total electron and hole density (per unit length) at a site  $i$  is given by [41]

$$n_i = -4i \sum_v \int \frac{dE}{2\pi} \frac{G_{i,i}^{<v}}{\Delta z}, \quad (126)$$

$$p_i = +4i \sum_v \int \frac{dE}{2\pi} \frac{G_{i,i}^{>v}}{\Delta z}, \quad (127)$$

where the summation runs over all the subbands contributing to transport and  $\Delta z$  is the average distance between the rings (4.3). The factor 4 in (126) and (127) is owing to double spin and double subband degeneracy (Section 2.2.2). To evaluate these integrals numerically, the energy grid should be selected such that the numerical error of the

calculation can be controlled. This issue is discussed in Section 4.8.

### 4.7.2. Current

By expanding Green's function in terms of the basis functions, the continuity equation (73) can be derived as

$$-\frac{i\hbar q}{\Delta z} \lim_{t_2 \rightarrow t_1} \left( \partial_{t_1} G_{i,j}^{<}(t_1, t_2) + \partial_{t_2} G_{i,j}^{<}(t_1, t_2) \right) \\ + \frac{J_{i+1/2}(t_1) - J_{i-1/2}(t_1)}{\Delta z} = 0, \quad (128)$$

where  $J_{i+1/2}$  represents the current passing through a point between  $i+1$  and  $i$ . Note that  $J$  has a unit of A rather than  $A/m^2$  owing to the one-dimensional nature of CNTs. The time derivative of Green's functions can be replaced by the relation (49),

$$\partial_t G_i = -\frac{q}{\Delta z} \sum_j \left\{ \left[ H_{i,j} G_{j,j}^{<}(t, t) - G_{i,j}^{<}(t, t) H_{j,i} \right] \right. \\ \left. + \int dt' \left[ \sum_{i,j}^r(t, t') G_{j,j}^{<}(t', t) + \sum_{i,j}^{<}(t, t') G_{j,j}^a(t', t) \right. \right. \\ \left. \left. + G_{i,j}^r(t, t') \sum_{j,i}^{<}(t', t) + G_{i,j}^{<}(t, t') \sum_{j,i}^a(t', t) \right] \right\}, \\ = -\frac{J_{i+1/2}(t) - J_{i-1/2}(t)}{\Delta z}, \quad (129)$$

where the term inside the integral is zero owing to the condition stated in (76).

The next step is separating  $J_{i+1/2}$  from  $J_{i-1/2}$  by decomposing eqn (128). Caroli proposed the following ansatz in [79]. The current  $J$  is the difference between the flow of particles from left to right and from right to left. This leads to the following expression for  $J_i$  [79]:

$$J_{i+1/2}(t) = -q \sum_{j \geq i+1} \sum_{k \leq i} \left( H_{j,k} G_{k,j}^{<}(t, t) - G_{j,k}^{<}(t, t) H_{k,j} \right) \quad (130)$$

It is straightforward to show that (130) along with an expression for  $J_{i-1}$  satisfies (129).

Under steady-state condition, one can transform the time difference coordinate to energy to obtain

$$J_{i+1/2} = -\frac{q}{\hbar} \sum_{j \geq n+1} \sum_{k \leq n} \int \frac{dE}{2\pi} \left( H_{j,k} G_{k,j}^{<}(E) - G_{j,k}^{<}(E) H_{k,j} \right), \\ = -\frac{q}{\hbar} \sum_{j \geq n+1} \sum_{k \leq n} \int \frac{dE}{2\pi} 2\text{Re} \left[ H_{j,k} G_{k,j}^{<}(E) \right], \quad (131)$$

Based on the nearest-neighbor tight-binding method in mode-space (see Section 4.4), eqn (131) can be simplified to

$$J_{i+1/2}^v = \frac{4q}{\hbar} \sum_v \int \frac{dE}{2\pi} 2\text{Re} \left[ t_{i+1,i}^v G_{i,i+1}^{<v} \right], \quad (132)$$

where the summation runs over all the subbands contributing to transport. The factor 4 in (132) is owing to double spin and double subband degeneracy.

### 4.7.3. Discussion

The carrier concentration is related to the diagonal elements of Green's function. The calculation of the current requires only the nearest off-diagonal elements of Green's function. Furthermore, the Hamiltonian matrix is tridiagonal. Considering these factors one can employ an efficient method, such as the recursive Green's method, to calculate only the required elements of Green's functions.

The recursive method has been proposed in [7, 98]. The operations required to solve for all elements of  $G^r$  with a size of  $N \times N$  scales as  $N^3$ . However, the required operations for the recursive method scale linearly with  $N$  [7].

## 4.8. Selection of the Energy Grid

For a numerical solution of the transport equations, one has to discretize Green's functions in the energy domain. However, owing to the presence of the narrow resonances at some energies, one has to be careful about the selection of the energy grid.

### 4.8.1. Nonadaptive Energy Grid

One can straightforwardly divide the integration domain into  $N_E$  equidistant intervals  $\Delta E = (E_{\max} - E_{\min}) / N_E$ . A disadvantage of this method is that the numerical error cannot be predefined. This problem is more pronounced when the integrand is not smooth. For accurate results, a grid spacing smaller than  $\Gamma$  has to be employed. For example, to resolve a resonance of  $\Gamma \approx 1 \mu\text{eV}$  width in an energy range of 1 eV, more than  $10^6$  energy grid points are required, which would severely increase the computational cost. For even narrower resonances, (e.g.,  $\Gamma \approx 1 \text{ neV}$ ), an equidistant grid is no longer feasible. To avoid these problems an adaptive method needs to be employed.

### 4.8.2. Adaptive Energy Grid

There is a variety of methods available for numerical adaptive integration [194]. Adaptive strategies divide the integration interval into subintervals and, typically, employ a progressive formula in each subinterval with some fixed upper limit on the number of points. If the required accuracy is not achieved by the progressive formula, the subinterval is bisected and a similar procedure carried out on each half. This subdivision process is carried out recursively until the desired accuracy is achieved. An obvious way to obtain an error estimate is based on the comparison between the two quadrature approximations [195]. However, owing to the dependence of such procedures on the underlying integration formulae, this method may not be reliable [196]. Error estimation with sequences of *null rules* has been proposed as a simple solution [197]. In adaptive quadrature algorithms the error estimate governs the decision on whether to accept the current approximation and terminate or to continue. Therefore, both the efficiency and the reliability depend on the error estimation algorithm. The decision to further subdivide a region may be based on either *local* or *global* information. Local information refers only to the region being currently processed, while global information refers obviously to data concerning all regions.

Integration programs based on global subdivision strategies are more efficient and reliable [198]. In this work a global error estimator based on the null rules method has been employed [196].

## 4.9. Self-Consistent Simulations

For an accurate analysis it is essential to solve the coupled system of transport and Poisson equations self-consistently [142]. The iterative method for solving this coupled system is presented.

### 4.9.1. Self-Consistent Iteration Scheme

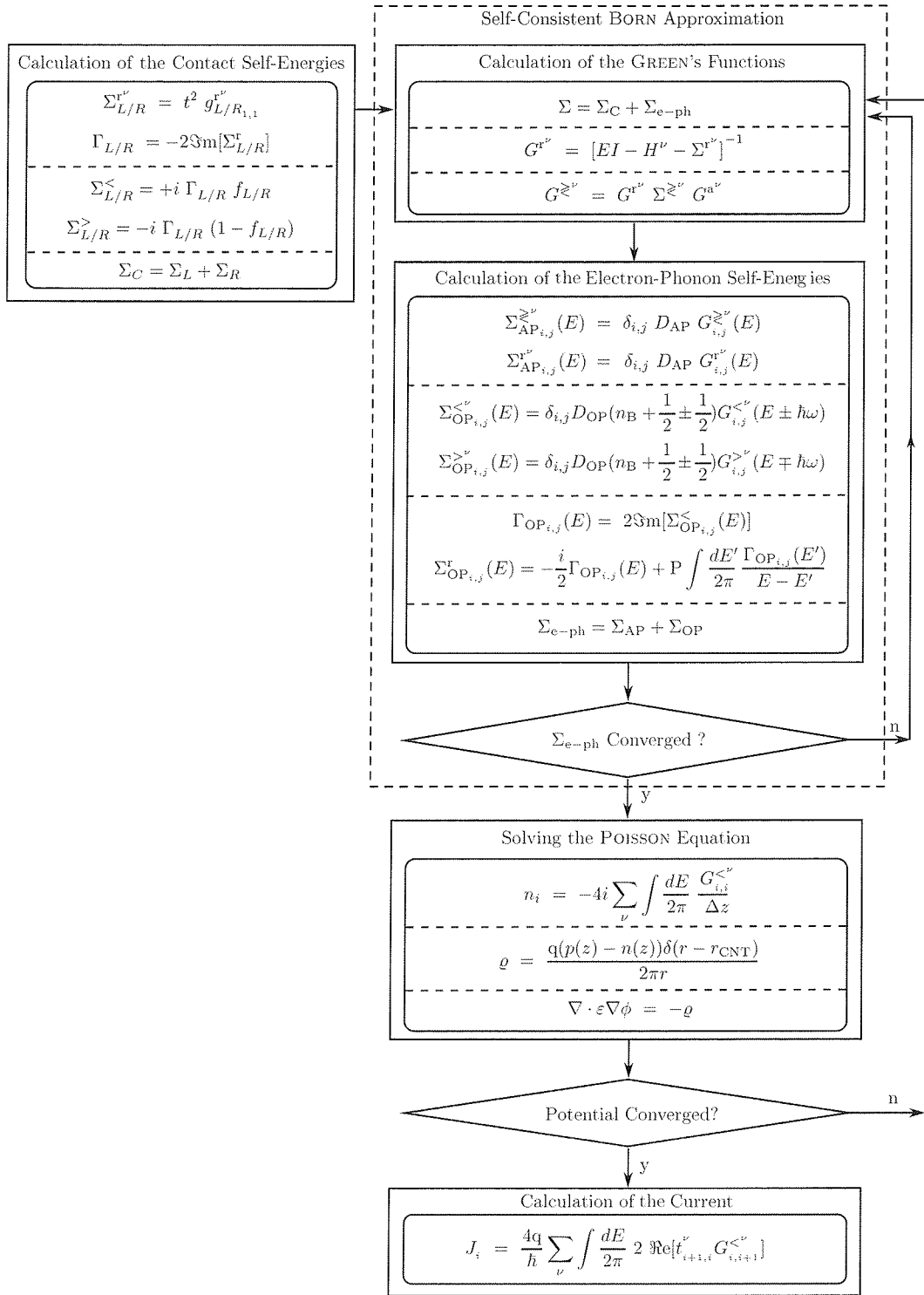
Figure 19 depicts the block diagram of the iterative procedure employed to solve the coupled system of transport and Poisson equations. All the procedures are discussed in the following. We solve the kinetic equations in mode space, (101) and (102), to obtain Green's functions. The required elements for calculating Green's functions are the Hamiltonian, and the electron-phonon and contact self-energies. As discussed in Section 4.4, diagonal elements of the Hamiltonian are potential energies, which can be obtained from the solution of the Poisson equation, and off-diagonal elements represent the coupling between the adjacent rings of the carbon atoms in the CNT. Given the contact properties and the contact-device coupling parameters, the contact self-energy  $\Sigma_C^{r^v}$  can be calculated once at the start of the simulation (see Section 4.5).

The calculation of the electron-phonon self-energy is presented in Section 4.6. Within the self-consistent Born approximation of the electron-phonon self-energy (Section 3.5), the noninteracting Green's function  $G_0$  is replaced by the full Green's function  $G$ . However, the full Green's is not known and has to be calculated. As a result, a coupled system of equations is achieved which can be solved by the iteration,

$$\underline{G}^{r^v(m)} = \left[ E \underline{I} - \underline{H}^v - \underline{\Sigma}_{\text{e-ph}}^{r^v(m-1)} - \underline{\Sigma}_C^{r^v} \right]^{-1}, \quad (133)$$

where  $m$  denotes the iteration number. For the first step the electron-phonon self-energy is assumed to be zero and Green's function is calculated from the kinetic equation. The next iteration starts with the calculation of the electron-phonon self-energy based on Green's function from the previous iteration. The updated electron-phonon self-energy is then used for the calculation of Green's function. This iteration continues till a convergence criterion is satisfied. The mentioned procedure should be followed for each subband (mode), and finally the total charge density is calculated.

In semiclassical simulations, the coupled system of the transport and Poisson equations is solved by Gummel's or Newton's method [187]. Both Gummel's method [199] and a variation of Newton's method [200] can be employed in self-consistent quantum mechanical simulations. While Gummel's method has a fast initial error reduction, for Newton's method it is very important that the initial guess



**Figure 19.** Block diagram of the iterative procedure employed to solve the coupled system of transport and Poisson equations. For the first step an initial guess for the electron-phonon self-energy is required; here we assumed  $\Sigma_{e-ph} = 0$ .

is close to the solution. The computational cost per iteration of Newton's method can be much higher than that for Gummel's method.

We employed Gummel's method, where after convergence of the electron-phonon self-energy, the Poisson

equation is solved once. Based on the updated electrostatic potential, Green's functions and the electron-phonon self-energy are iterated again. These two iterations continue until a convergence criterion is satisfied. Finally, the total current through the device is calculated.

#### 4.9.2. Convergence of the Self-Consistent Simulations

The coupled system of transport and Poisson equation can be solved by iteration with appropriate numerical damping, which terminates if a convergence criterion is satisfied [201]. In this work, the maximum element of the potential update, corresponding to  $L_\infty = |\phi^k - \phi^{k-1}|_\infty$ , is considered as a measure of convergence.

One of the reasons causing convergence problems [202, 203] is the exponential dependence of the carrier concentration on the electrostatic potential,  $n \propto \exp(q\phi / k_B T)$ . A small potential variation causes large variation in the carrier concentration. As a result, a strong damping is required in many cases, which increases the simulation time. To avoid this problem a nonlinear Poisson equation is generally employed [204]. Solving a nonlinear Poisson equation takes that exponential dependence into account. Compared to the linear Poisson equation, it leads to faster convergence in both semiclassical [204] and quantum mechanical [203, 205] transport simulations. In this work the Gummel method along with a nonlinear Poisson equation is employed.

However, we show that an inappropriate energy grid for the discretization of the transport equations can be another reason of convergence problems in quantum transport simulations [206]. It is demonstrated that with adaptive energy grids the iterative solution can converge very fast, and the simulation time can decrease considerably.

### 5. APPLICATION

Based on the quantum transport model outlined in Section 4, both the static and dynamic response of CNT-FETs are investigated. The effect of electron-phonon interactions on the device characteristics is discussed in detail. In agreement with the experimental data, our results indicate

that electron-phonon interactions can affect the switching response of CNT-FETs significantly, while the effect on the DC characteristics is small.

#### 5.1. The Effect of the Electron-Phonon Interaction

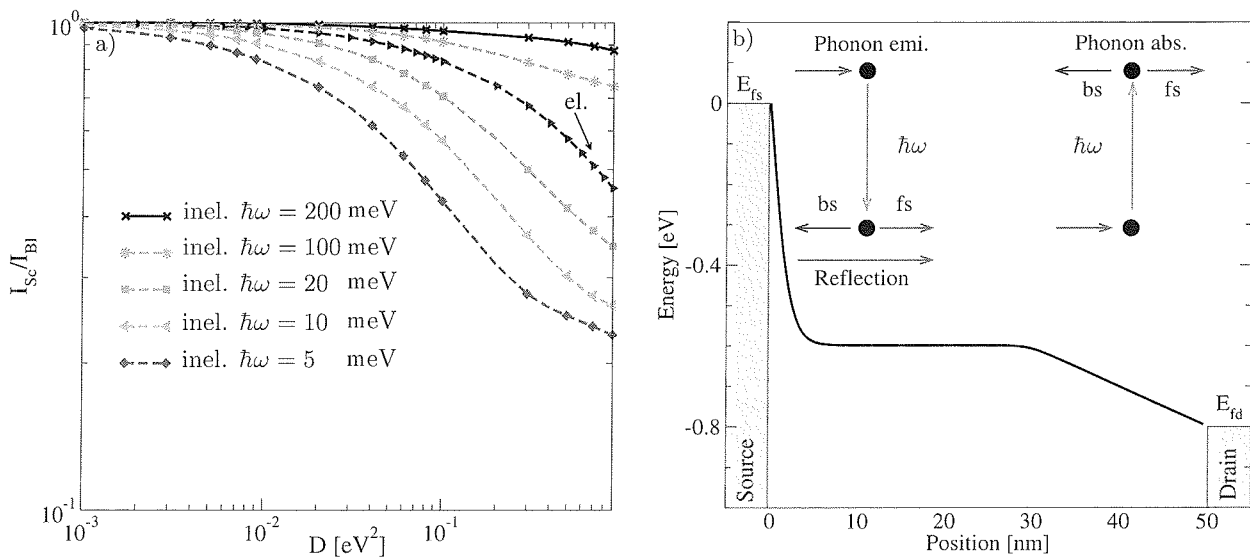
The electron-phonon coupling strength and the phonon energy depend on the chirality and the diameter of the CNT (see Section 2.3.4). In this section the device response of a CNT-FET with a channel length of 50 nm is studied for a wide range of electron-phonon interaction parameters.

##### 5.1.1. Electron-Phonon Coupling Strength

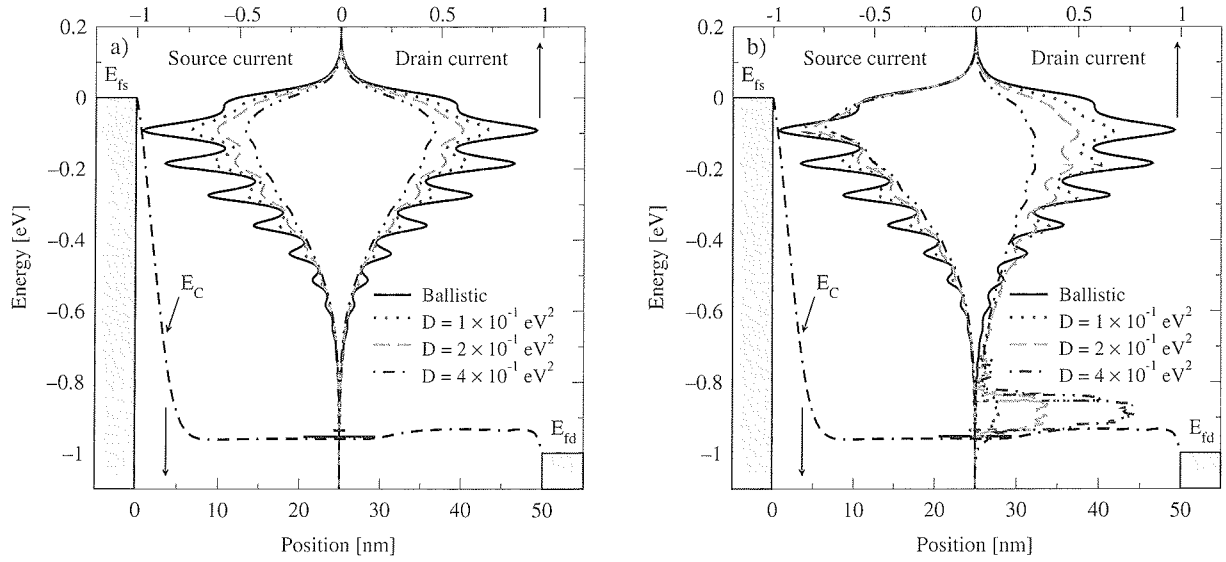
Figure 20(a) shows the ballisticity as a function of the electron-phonon coupling strength. The ballisticity is defined as  $I_{Sc} / I_{Bl}$ , the ratio of the on-current in the presence of electron-phonon interaction to the current in the ballistic case [207].

The left part of Figure 20(b) illustrates an electron losing its kinetic energy by emitting a phonon. The electron will be scattered either forward or backward. In the case of backward scattering, the electron faces a thick barrier near the source contact and will be reflected with high probability, such that its momentum will again be directed toward the drain contact.

Elastic scattering conserves the energy of carriers, but the current decreases owing to elastic backscattering of carriers. Figure 21(a) shows the current spectra at the source and drain contact. For elastic scattering the spectra are symmetric. As the electron-phonon coupling strength increases, resonances in the current spectrum are washed out, and the total current decreases owing to elastic backscattering. In the case of inelastic scattering, carriers acquiring enough kinetic energy can emit a phonon and scatter into lower-energy states. Therefore, as shown in Figure 21(b), the source and drain current spectra are not symmetric.



**Figure 20.** (a) Ballisticity versus electron-phonon coupling strength for a CNT of 50 nm length. Results for both elastic and inelastic scattering with different phonon energies are shown. The operating point is  $V_G = V_D = 1$  V. (b) Sketch of phonon emission and absorption processes in the channel. Reprinted with permission from [208], M. Pourfath et al., *Nanotechnology* 18, 424036 (2007). © 2007, IOP Publishing.



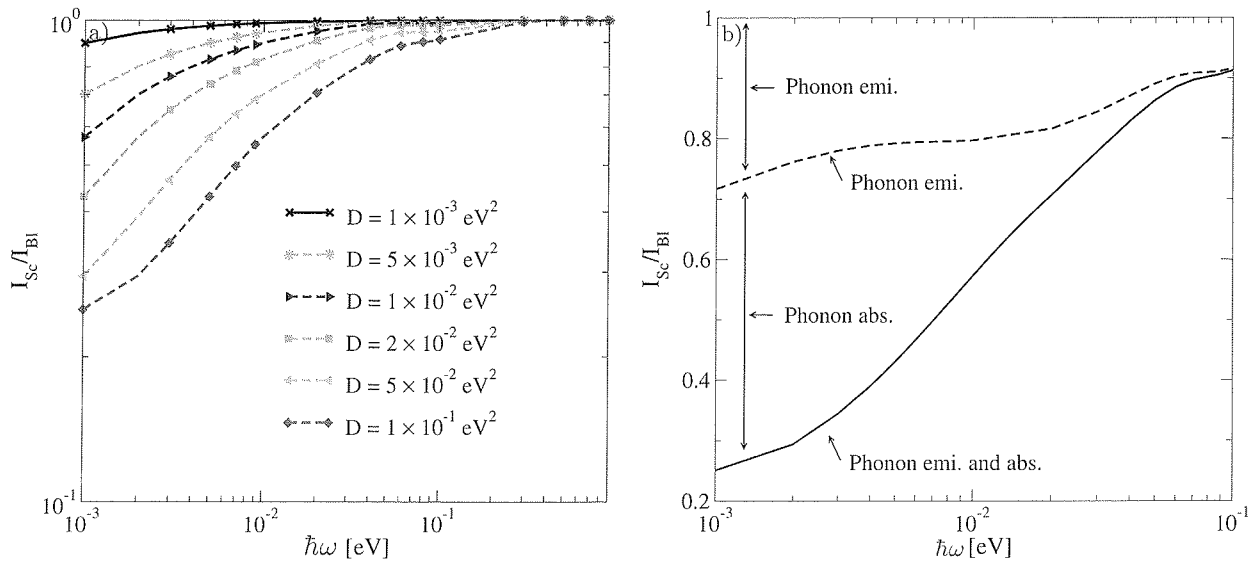
**Figure 21.** The spectra of the source and drain currents: (a) The effect of elastic phonon scattering with different coupling strengths is shown. As the coupling strength increases resonances in the current spectrum wash out and the total current decreases owing to elastic backscattering. (b) The effect of inelastic phonon scattering with different coupling strengths is shown. The phonon energy is  $\hbar\omega = 100$  meV. Carriers acquiring enough kinetic energy can emit phonons and scatter into lower energy states. Since the energy of electrons is not conserved in this process, the source and drain current spectrum are not symmetric. As the coupling strength increases more electrons are scattered into lower energy states. Reprinted with permission from [208], M. Pourfath et al., *Nanotechnology* 18, 424036 (2007). © 2007, IOP Publishing.

### 5.1.2. Phonon Energy

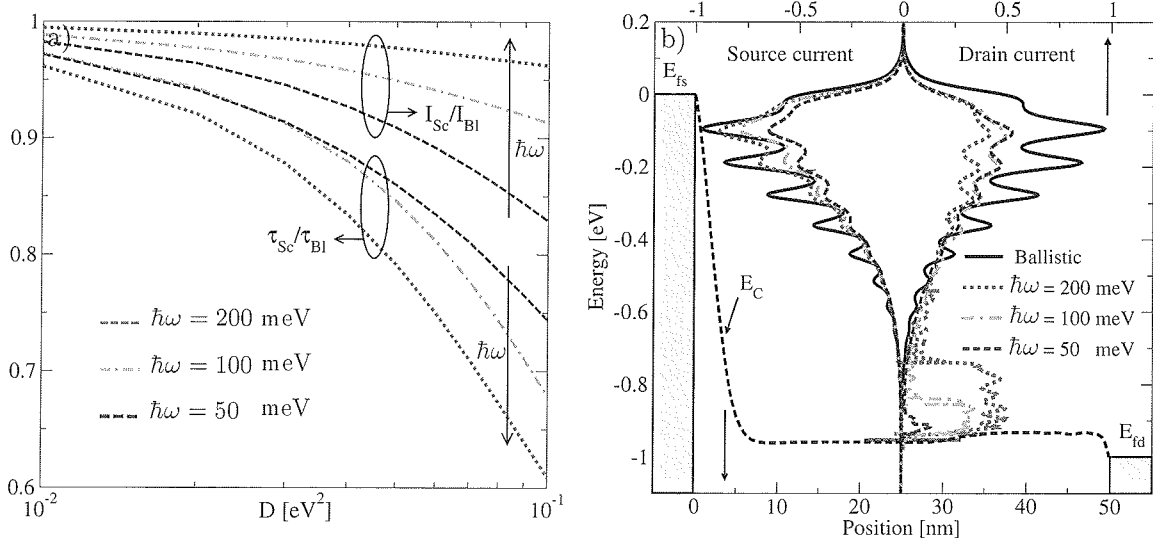
Figure 22(a) shows the dependence of the ballisticity with respect to the phonon energy. With increasing phonon energy the effect of phonon scattering on the current is reduced, because scattered electrons lose more kinetic energy, and the probability for traveling back to the source contact decreases. The considerable decrease of ballisticity for low-energy phonons is owing to the phonon absorption process.

The right part of Figure 20(b) shows an electron-absorbing energy from a phonon and scattering into a higher-energy state. In this case, the probability for arriving at the source contact increases. This process can severely reduce the total current.

Figure 22(b) separately shows the effects of the phonon emission and absorption processes on the ballisticity. As the phonon energy decreases, the phonon occupation number increases exponentially, and the self-energy contributions of these two components increase. However, owing to the higher



**Figure 22.** (a) Ballisticity versus phonon energy for a CNT of 50 nm length. Results for inelastic scattering with different electron–phonon couplings are shown.  $V_G = V_D = 1$  V. (b) Ballisticity versus phonon energy with  $D = 10^{-1}$  eV<sup>2</sup> at the bias point  $V_G = V_D = 1$  V. The contributions owing to phonon absorption and emission are shown. Reprinted with permission from [208], M. Pourfath et al., *Nanotechnology* 18, 424036 (2007). © 2007, IOP Publishing.



**Figure 23.** (a) The ratio of the gate-delay time in the presence of electron-phonon interaction to the gate-delay time in the ballistic case,  $\tau_{sc}/\tau_{bl}$ , as a function of the electron-phonon coupling strength. For comparison, the ratio  $I_{sc}/I_{bl}$  is also shown. As the phonon energy increases the gate-delay time increases. This behavior is owing to the reduction of the electron velocity in the channel and the resulting charge pile up. (b) The spectra of the source and drain currents. The effect of inelastic scattering with different phonon energies is shown. The electron-phonon coupling strength is  $D = 2 \times 10^{-1}$  eV<sup>2</sup>. The figure shows a considerable increase of the electron population close to the conduction band edge as the phonon energy increases. Reprinted with permission from [208], M. Pourfath et al., *Nanotechnology* 18, 424036 (2007). © 2007, IOP Publishing.

probability for backscattering of electrons in the case of phonon absorption, this component reduces the total current more effectively than the phonon emission process does.

Figure 23(a) shows the ratio of the gate-delay time in the presence of electron-phonon interaction to that in the ballistic case,  $\tau_{sc}/\tau_{bl}$ , as a function of the electron-phonon coupling strength. As the phonon energy increases, the gate-delay time increases. This behavior can be attributed to the average electron velocity in the channel, which is high for ballistic electrons and low for electrons scattered to lower energy states.

Figure 23(b) shows the spectra of the source and drain currents for different inelastic phonon energies. Electrons can emit a single phonon or a couple of phonons to reach lower-energy states. The probability of multiple-phonon emissions decreases as the number of interactions increases. Therefore, as the phonon energy increases, the occupation of electrons at lower energy states increases.

As shown in Figure 23(b), the electron population close to the conduction band edge considerably increases as the phonon energy increases. Therefore, as the phonon energy increases the mean velocity of electrons decreases and the carrier concentration in the channel increases (Fig. 24). The increased charge in the channel results in an increased gate-delay time.

### 5.1.3. Diffusive Limit

All the above-discussed results were calculated for a device with a CNT length of 50 nm. In the case of ballistic transport the current is independent of the device length, but in the presence of scattering it decreases as the device length increases. Figure 25(a) shows the ballisticity as a function of the CNT length in the presence of elastic and inelastic electron-phonon interaction. An artificially large value for the

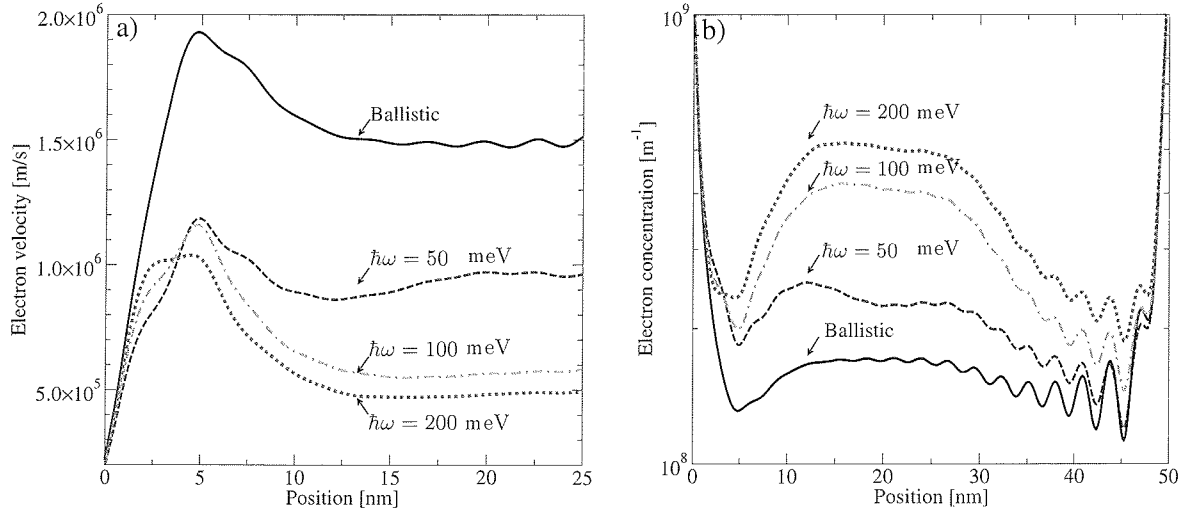
electron-phonon coupling strength and a small value for the phonon energy are chosen to simulate the diffusive limit (see Fig. 25(b)). In this case, the current is expected to be inversely proportional to the device length according to Ohm's law.

### 5.1.4. Discussion

In general the electron-phonon interaction parameters depend on the diameter and the chirality of the CNT (see Section 2.3.4). CNTs with a diameter  $d_{CNT} > 2$  nm have a band gap  $E_G < 0.4$  eV (10), which render them unsuitable as channel for transistors. Since the fabrication of devices with a diameter  $d_{CNT} < 1$  nm is very difficult, we limit our study to zig-zag CNTs with diameters in the range of  $d_{CNT} = 1 - 2$  nm.

Scattering with acoustic phonons is treated as an elastic process. The electron-phonon coupling is also weak for acoustic phonons ( $D_{AP} < 10^{-3}$  eV<sup>2</sup>), which implies that elastic backscattering of carriers is weak. Inelastic scattering is induced by OP, RBM, and K-point phonons (Section 2.3.2). Considering the class of CNTs discussed above, the energies of these phonons are  $\hbar\omega_{OP} \approx 200$  meV,  $\hbar\omega_{RBM} \approx 25$  meV, and  $\hbar\omega_{K_1} \approx 160$  meV and  $\hbar\omega_{K_2} \approx 180$  meV [207, 210]. The corresponding coupling coefficients are  $D_{OP} \approx 40 \times 10^{-3}$  eV<sup>2</sup>,  $D_{RBM} \approx 10^{-3}$  eV<sup>2</sup>, and  $D_{K_1} \approx 10^{-4}$  eV<sup>2</sup>, and  $D_{K_2} \approx 10^{-3}$  eV<sup>2</sup> [207].

As discussed in Section 5.1.2, high-energy phonons such as OP and K-point phonons reduce the on-current only weakly, but can increase the gate-delay time considerably owing to charge pileup in the channel. Low-energy phonons such as the RBM phonon can reduce the on-current more effectively, but have a weaker effect on the gate-delay time. However, owing to strong coupling, scattering processes are mostly owing to electron-phonon



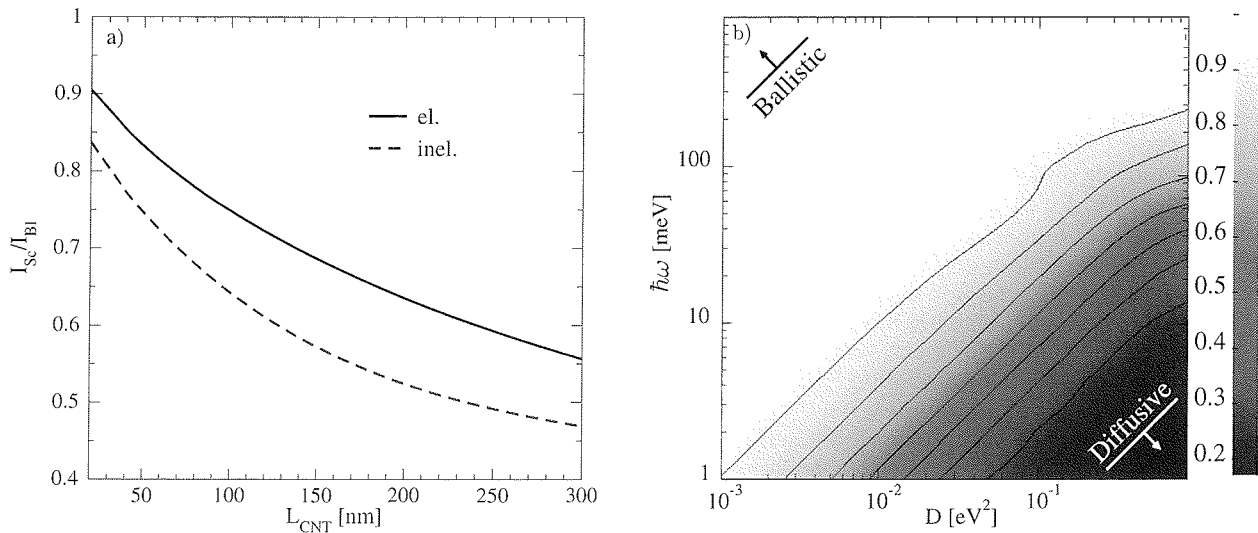
**Figure 24.** (a) The profile of the electron velocity near the source contact. (b) The profile of the electron concentration along the device. The results for the ballistic case and for electron–phonon interaction are shown. As the phonon energy increases the electrons scatter to lower energy states. Therefore, the electron velocity decreases and the carrier concentration increases. The electron–phonon coupling strength is  $D = 10^{-1} \text{ eV}^2$  and the bias point is  $V_G = V_D = 1 \text{ V}$ . Reprinted with permission from [208], M. Pourfath et al., *Nanotechnology* 18, 424036 (2007). © 2007, IOP Publishing.

interaction with high-energy phonons. Therefore, the on-current of short CNT-FETs can be close to the ballistic limit [211] (see Fig. 26), whereas the gate-delay time can be significantly below that limit [72, 73, 212]. The intrinsic (without parasitic capacitances) gate-delay time for the ballistic case can be approximated as  $\tau \approx 1.7 \text{ ps}/\mu\text{m}$ , or equivalently  $f_T \approx 100 \text{ GHz}/\mu\text{m}$  [213]. The highest reported intrinsic cut-off frequency for a device with a length of 300 nm is  $f_T \approx 30 \text{ GHz}$  [214], which is far below the ballistic limit. Inelastic electron–phonon interaction with high-energy phonon has to be considered to explain the results.

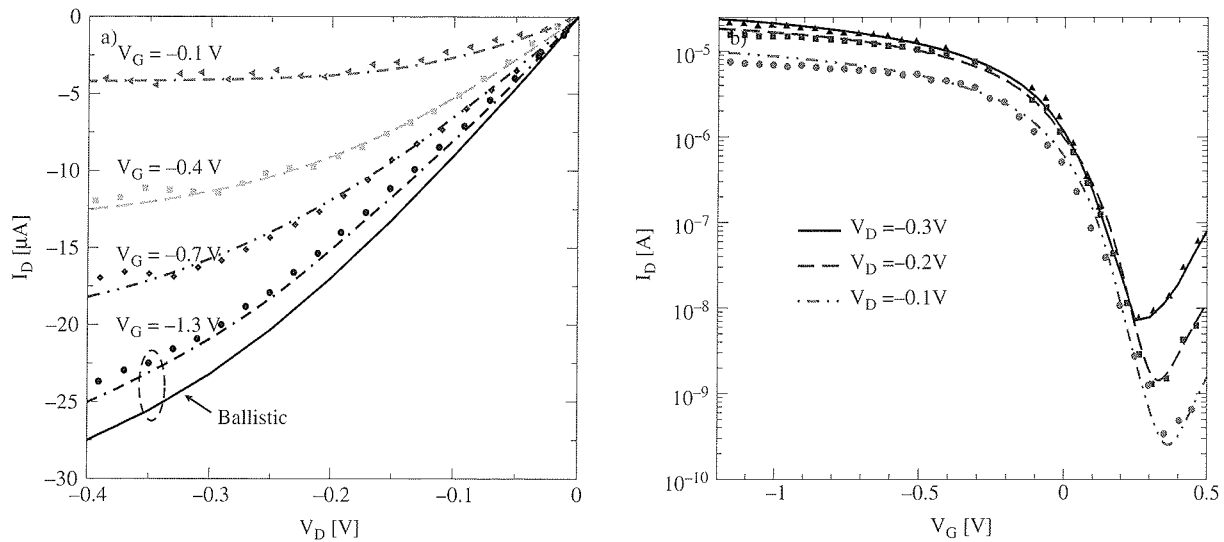
## 6. SUMMARY AND CONCLUSIONS

Ever since the first demonstration of CNT-FETs, their performance is improving very rapidly and the understanding of such devices is evolving. The one-dimensional nature of CNTs severely reduces the phase space for scattering, allowing CNTs to operate close to the ballistic limit even at room temperature. The low-scattering probability and high mobility are responsible for high on-current of CNT-FETs.

Furthermore, the chemical stability and perfection of the CNT structure suggests that the carrier mobility at high gate fields may not be affected by processing and roughness



**Figure 25.** (a) Ballistic transport versus CNT length. The electron–phonon coupling strength for both elastic and inelastic scattering is  $D = 10^{-1} \text{ eV}^2$ , and  $\hbar\omega = 25 \text{ meV}$  for inelastic scattering. These scattering parameters simulate the diffusive regime. In this case the ballisticity is inversely proportional to the device length [209]. (b) Ballistic transport as a function of the electron–phonon coupling strength and phonon energy for inelastic scattering. The scale of the ballisticity is shown in the color bar. The regions of ballistic and diffusive transport are shown. As the strength of the electron–phonon interaction increases, the transport of carriers deviates from the ballistic limit and becomes more diffusive.



**Figure 26.** Comparison of the simulation results and experimental data for the (a) output and (b) transfer characteristics. Lines show the simulation results and symbols show experimental data. The result for  $V_G = -1.3$  V is compared with the ballistic limit. Experimental data have been adopted from [211]. Reprinted with permission from [215], M. Pourfath et al., *J. Phys.: Conf. Ser.* 38, 29 (2006). © 2006, IOP Publishing.

scattering as it is the case in the conventional semiconductor channel. Electrostatic control is improved as well. The fact that there are no dangling bond states at the surface of CNTs allows for a much wider choice of gate insulators beyond the conventional  $\text{SiO}_2$ . Also, the strong one-dimensional electron confinement of the SW-CNTs (typically 1–2 nm diameter) should lead to a suppression of short-channel effects in transistor devices [4].

As far as integration is concerned, semiconducting CNTs benefit from their band structure which gives essentially the same effective mass for electrons and holes. This should enable similar mobilities and performances of n- and p-type transistors, which are necessary for a CMOS-like technology. The most important appeal of this approach is the ability to fabricate one of the critical device dimensions (the CNT diameter) reproducibly using synthetic chemistry.

The purposes of this work are to develop a simulation approach and tools for CNT-FETs and apply them to understand device physics and explore device issues, which are crucial for improving device performance. We employed the NEGF technique for modeling transport phenomena in CNT-FETs. The NEGF technique allows one to study the time evolution of a many-particle quantum system. Knowledge of the single-particle Green's function provides properties of the system and the excitation energies of the system. The many-particle information about the system is cast into self-energies, parts of the equations of motion for Green's functions. Green's functions can be expressed as a perturbation expansion, which is the key to approximate the self-energies. Green's functions provide a very powerful technique for evaluating properties of many-particle systems both in thermodynamic equilibrium and also in nonequilibrium situations.

We solve the coupled system of transport and Poisson equations self-consistently. A tight-binding Hamiltonian is used to describe transport phenomena in CNT-FETs. The mode-space transformation used in this work reduces the computational cost considerably. The mode-space approach

takes only a relatively small number of transverse modes into consideration. To reduce the computational cost even further, we used the local scattering approximation [95]. In this approximation the scattering self-energy terms are diagonal in coordinate representation. We show that the local approximation is well justified for electron-phonon scattering caused by deformation potential interaction.

The carrier concentration is related to the diagonal elements of Green's function. The calculation of the current requires only the nearest off-diagonal elements of Green's function. Furthermore, by using a nearest tight-binding Hamiltonian and assuming the local scattering approximation, the achieved matrix is tridiagonal. Considering these factors we employed the efficient recursive Green's function method to calculate only the required elements of Green's functions.

We also investigated methods of generating energy grids for numerical integration and their effects on the convergence behavior of the self-consistent iteration. Our results indicate that for accurate and fast convergent simulations the energy grid must be carefully adapted. All methods were implemented into the multipurpose quantum-mechanical solver VSP.

Employing the described model, we investigated both the static and dynamic response of CNT-FETs. The effect of electron-phonon interaction on the device characteristics is discussed in detail. In agreement with the experimental data, our results indicate that electron-phonon interaction affects the DC current of CNT-FETs only weakly, whereas the switching response of such devices can be significantly degraded.

## APPENDICES

### A. TIME-EVOLUTION PICTURES

The time evolution of operators and state vectors in quantum mechanics can be expressed in different representations. The Schrödinger, the interaction, and the



Heisenberg representations are useful in analyzing the second-quantized form of the Schrödinger equation. The Hamiltonian is assumed to be of the form (see (20))

$$\hat{H} = \hat{H}_0 + \hat{H}^{\text{int}}, \quad (134)$$

where  $\hat{H}_0$  is the noninteracting part, which is assumed to be exactly solvable.  $\hat{H}^{\text{int}}$  contains all the interactions, such as carrier-carrier, carrier-phonon, impurity scattering, and so forth.

### A.1. Schrödinger Picture

In the Schrödinger picture the operators  $\hat{O}_S$  are time independent

$$\hat{O}_S(t) = \hat{O}_S(t_0) = \hat{O}_S, \quad (135)$$

where  $t_0$  is assumed to be the time reference point. The time dependence of the state vector  $|\Psi_S(t)\rangle$  is obtained from the Schrödinger equation

$$i\hbar\partial_t|\Psi_S(t)\rangle = \hat{H}|\Psi_S(t)\rangle, \quad (136)$$

which has the formal solution

$$|\Psi_S(t)\rangle = e^{-i\hat{H}(t-t_0)/\hbar}|\Psi_S(t_0)\rangle \quad (137)$$

### A.2. Interaction Picture

In the interaction representation both the state vectors and the operators are time dependent. The state vector in the interaction representation is given by

$$|\Psi_I(t)\rangle = e^{i\hat{H}_0 t/\hbar}|\Psi_S(t)\rangle, \quad (138)$$

which is merely a unitary transformation at the time  $t$ . The equation of motion of this state vector is found by taking the time derivative,

$$\begin{aligned} i\hbar\partial_t|\Psi_I(t)\rangle &= -\hat{H}_0 e^{i\hat{H}_0 t/\hbar}|\Psi_S(t)\rangle + e^{i\hat{H}_0 t/\hbar} i\hbar\partial_t|\Psi_S(t)\rangle \\ &= e^{i\hat{H}_0 t/\hbar} \left[ -\hat{H}_0 + \hat{H}_0 + \hat{H}^{\text{int}} \right] e^{-i\hat{H}_0 t/\hbar} |\Psi_I(t)\rangle. \end{aligned} \quad (139)$$

Therefore, one obtains the following set of equations in the interaction picture:

$$\begin{aligned} i\hbar\partial_t|\Psi_I(t)\rangle &= \hat{H}^{\text{int}}(t)|\Psi_I(t)\rangle, \\ \hat{H}^{\text{int}}(t) &\equiv e^{i\hat{H}_0 t/\hbar} \hat{H}^{\text{int}} e^{-i\hat{H}_0 t/\hbar}. \end{aligned} \quad (140)$$

An arbitrary matrix element in the Schrödinger picture can be written as

$$\langle\Psi'_S(t)|\hat{O}_S|\Psi_S(t)\rangle = \langle\Psi'_I(t)|e^{i\hat{H}_0 t/\hbar}\hat{O}_S e^{-i\hat{H}_0 t/\hbar}|\Psi_I(t)\rangle, \quad (141)$$

which suggests the following definition of an operator in the interaction picture:

$$\hat{O}_I(t) = e^{i\hat{H}_0 t/\hbar}\hat{O}_S e^{-i\hat{H}_0 t/\hbar}. \quad (142)$$

### A.3. Heisenberg Picture

In the Heisenberg representation, state vectors are defined as

$$|\Psi_H(t)\rangle = e^{i\hat{H}t/\hbar}|\Psi_S(t)\rangle. \quad (143)$$

Its time derivative may be combined with (136) to yield  $i\hbar\partial_t|\Psi_H(t)\rangle = 0$ , which shows that  $|\Psi_H(t)\rangle$  is time independent. Since an arbitrary matrix element in the Schrödinger picture can be written as

$$\begin{aligned} \langle\Psi'_S(t)|\hat{O}_S|\Psi_S(t)\rangle &= \\ \langle\Psi'_H(t)|e^{i\hat{H}t/\hbar}\hat{O}_S e^{-i\hat{H}t/\hbar}|\Psi_H(t)\rangle, \end{aligned} \quad (144)$$

a general operator in the Heisenberg picture is given by

$$\hat{O}_H(t) = e^{i\hat{H}t/\hbar}\hat{O}_S e^{-i\hat{H}t/\hbar}. \quad (145)$$

Eqn (145) can be rewritten in terms of the interaction picture operators,

$$\hat{O}_H(t) = e^{i\hat{H}t/\hbar} e^{-i\hat{H}_0 t/\hbar} \hat{O}_I(t) e^{i\hat{H}_0 t/\hbar} e^{-i\hat{H}t/\hbar}, \quad (146)$$

or in terms of the operator  $\hat{S}$  derived in the next section

$$\hat{O}_H(t) = \hat{S}(0,t)\hat{O}_I(t)\hat{S}(t,0). \quad (147)$$

### A.4. The Evolution Operator $\hat{S}$

To solve the equations of motion in the interaction picture (140), a unitary operator  $\hat{S}(t,t_0)$  that determines the state vector at time  $t$  in terms of the state vector at time  $t_0$  is introduced,

$$|\Psi_I(t)\rangle = \hat{S}(t,t_0)|\Psi_I(t_0)\rangle, \quad (148)$$

$\hat{S}$  satisfies the initial condition  $\hat{S}(t_0,t_0) = 1$ . For finite times  $\hat{S}(t,t_0)$  can be constructed explicitly by employing the Schrödinger picture,

$$\begin{aligned} |\Psi_I(t)\rangle &= e^{i\hat{H}_0 t/\hbar}|\Psi_S(t)\rangle \\ &= e^{i\hat{H}_0 t/\hbar} e^{-i\hat{H}(t-t_0)/\hbar}|\Psi_S(t_0)\rangle, \\ &= e^{i\hat{H}_0 t/\hbar} e^{-i\hat{H}(t-t_0)/\hbar} e^{-i\hat{H}_0 t_0/\hbar}|\Psi_I(t_0)\rangle, \end{aligned} \quad (149)$$

which therefore identifies

$$\hat{S}(t, t_0) = e^{iH_0 t/\hbar} e^{-i\hat{H}(t-t_0)/\hbar} e^{-iH_0 t_0/\hbar}. \quad (150)$$

Since  $\hat{H}$  and  $\hat{H}_0$  do not commute with each other, the order of the operators must be carefully maintained. Eqn (150) immediately yields several general properties of  $\hat{S}$  [169]:

- $\hat{S}^\dagger(t, t_0)\hat{S}(t, t_0) = \hat{S}(t, t_0)\hat{S}^\dagger(t, t_0) = 1$ , implying that  $\hat{S}$  is unitary  $\hat{S}^\dagger(t, t_0) = \hat{S}^{-1}(t, t_0)$ ,
- $\hat{S}(t_1, t_2)\hat{S}(t_2, t_3) = \hat{S}(t_1, t_3)$ , which shows that  $\hat{S}$  has the group property, and
- $\hat{S}(t, t_0)\hat{S}(t_0, t) = 1$ , implying that  $\hat{S}(t_0, t) = \hat{S}^\dagger(t, t_0)$ .

Although (150) is the formal solution to the problem posed by (148), it is not very useful for computational purposes. Instead, one can construct an integral equation for  $\hat{S}$ , which can then be solved by iteration. It follows from (140) and (148) that  $\hat{S}$  satisfies the differential equation,

$$i\hbar\partial_t \hat{S}(t, t_0) = \hat{H}_I^{\text{int}}(t) \hat{S}(t, t_0). \quad (151)$$

Integrating both sides of (151) with respect to time with the initial condition  $\hat{S}(t_0, t_0) = 1$  yields

$$\begin{aligned} \hat{S}(t, t_0) &= \hat{S}(t_0, t_0) - \frac{i}{\hbar} \int_{t_0}^t dt_1 \hat{H}_I^{\text{int}}(t_1) \hat{S}(t_1, t_0) \\ &= 1 - \frac{i}{\hbar} \int_{t_0}^t dt_1 \hat{H}_I^{\text{int}}(t_1) \hat{S}(t_1, t_0). \end{aligned} \quad (152)$$

By iterating this equation repeatedly, one gets

$$\begin{aligned} \hat{S}(t, t_0) &= 1 - \frac{i}{\hbar} \int_{t_0}^t dt_1 \hat{H}_I^{\text{int}}(t_1) + \left(\frac{-i}{\hbar}\right)^2 \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \\ &\quad \times \hat{H}_I^{\text{int}}(t_1) \hat{H}_I^{\text{int}}(t_2) + \left(\frac{-i}{\hbar}\right)^n \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \int_{t_0}^{t_2} dt_3 \dots \int_{t_0}^{t_{n-1}} dt_n \hat{H}_I^{\text{int}}(t_1) \\ &\quad \times \hat{H}_I^{\text{int}}(t_2) \dots \hat{H}_I^{\text{int}}(t_n) \\ &= \sum_{n=0}^{\infty} \left(\frac{-i}{\hbar}\right)^n \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \int_{t_0}^{t_2} dt_3 \dots \int_{t_0}^{t_{n-1}} dt_n \hat{H}_I^{\text{int}}(t_1) \\ &\quad \times \hat{H}_I^{\text{int}}(t_2) \dots \hat{H}_I^{\text{int}}(t_n). \end{aligned} \quad (152)$$

Eqn (153) has the characteristic feature that the operator containing the latest time stands farthest to the left. At this point it is convenient to introduce the *time-ordering operator* denoted by the symbol  $T_t$

$$\begin{aligned} T_t \{ \hat{A}(t_1) \hat{B}(t_2) \} &= \theta(t_1 - t_2) \hat{A}(t_1) \hat{B}(t_2) \\ &\quad + \theta(t_2 - t_1) \hat{B}(t_2) \hat{A}(t_1). \end{aligned} \quad (154)$$

where  $\theta(t)$  is the step function. The step function is defined as

$$\theta(t) = \begin{cases} 1 & t > 0 \\ \frac{1}{2} & t = 0 \\ 0 & t < 0. \end{cases}$$

Each time two Fermi ons are interchanged, the resulting expression changes its sign. By rearranging the integral using  $T_t$

$$\begin{aligned} &\frac{1}{2!} \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 T_t \{ H_I^{\text{int}}(t_1) H_I^{\text{int}}(t_2) \} \\ &= \frac{1}{2!} \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \hat{H}_I^{\text{int}}(t_1) \hat{H}_I^{\text{int}}(t_2) \\ &\quad + \frac{1}{2!} \int_{t_0}^t dt_2 \int_{t_0}^{t_2} dt_1 \hat{H}_I^{\text{int}}(t_2) \hat{H}_I^{\text{int}}(t_1). \end{aligned} \quad (155)$$

The second term on the right-hand side is equal to the first, which is easy to see by just redefining the integration variables  $t_1 \rightarrow t_2$ ,  $t_2 \rightarrow t_1$ . Thus, one gets

$$\begin{aligned} &\frac{1}{2!} \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 T_t \{ H_I^{\text{int}}(t_1) H_I^{\text{int}}(t_2) \} \\ &= \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 H_I^{\text{int}}(t_1) H_I^{\text{int}}(t_2) S(t, t_0). \end{aligned} \quad (156)$$

Thus, for the expansion of the  $\hat{S}(t, t_0)$ , one obtains

$$\begin{aligned} \hat{S}(t, t_0) &= \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{-i}{\hbar}\right)^n \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \dots \int_{t_0}^{t_{n-1}} dt_n T_t \\ &\quad \times \{ \hat{H}_I^{\text{int}}(t_1) \hat{H}_I^{\text{int}}(t_2) \dots \hat{H}_I^{\text{int}}(t_n) \} \\ &= T_t \left\{ \exp \left( -\frac{i}{\hbar} \int_{t_0}^t dt' \hat{H}_I^{\text{int}}(t') \right) \right\} \end{aligned} \quad (157)$$

## B. NONINTERACTING GREEN'S FUNCTIONS

The noninteracting or free Green's function is used in the perturbation expansions described in Section 3.3. In this appendix real-time Green's functions for both electrons and phonons are derived.

### B.1. Noninteracting Fermions

The Hamiltonian for noninteracting electrons (Fermions) in momentum representation is

$$H_0 = \sum_{\mathbf{k}} \xi_{\mathbf{k}} c_{\mathbf{k}}^\dagger c_{\mathbf{k}}, \quad (158)$$

where  $\xi_{\mathbf{k}} = E_{\mathbf{k}} - E_F$  is the single-particle energy measured with respect to the Fermi energy  $c_{\mathbf{k}}$  and  $c_{\mathbf{k}}^\dagger$  are the Fermion annihilation and creation operators, respectively. The time

evolution of the annihilation operator in the Heisenberg picture is (Appendix A)

$$c_{\mathbf{k}}(t) = e^{iH_0 t/\hbar} c_{\mathbf{k}} e^{-iH_0 t/\hbar}, \quad (159)$$

so the operator obeys the equation,

$$i\hbar \partial_t c_{\mathbf{k}}(t) = [c_{\mathbf{k}}(t), H_0] = \xi_{\mathbf{k}} c_{\mathbf{k}}(t), \quad (160)$$

which has the solution,

$$c_{\mathbf{k}}(t) = e^{-i\xi_{\mathbf{k}} t/\hbar} c_{\mathbf{k}}. \quad (161)$$

The creation operator for Fermions is just the Hermitian conjugate of  $c_{\mathbf{k}}$ , that is

$$c_{\mathbf{k}}^\dagger(t) = e^{i\xi_{\mathbf{k}} t/\hbar} c_{\mathbf{k}}^\dagger. \quad (162)$$

The noninteracting real-time Green's functions (Section 3.6.1) for Fermions in momentum representation are now given by

$$\begin{aligned} G_0^<(\mathbf{k}, t; \mathbf{k}', t') &\equiv +i\hbar^{-1} \langle c_{\mathbf{k}}^\dagger(t') c_{\mathbf{k}}(t) \rangle_0, \\ &= +i\hbar^{-1} e^{-i\xi_{\mathbf{k}}(t-t')/\hbar} n_{\mathbf{k}} \delta_{\mathbf{k}, \mathbf{k}'}, \\ G_0^>(\mathbf{k}, t; \mathbf{k}', t') &\equiv -i\hbar^{-1} \langle c_{\mathbf{k}}(t) c_{\mathbf{k}}^\dagger(t') \rangle_0, \\ &= -i\hbar^{-1} e^{-i\xi_{\mathbf{k}}(t-t')/\hbar} [1 - n_{\mathbf{k}}] \delta_{\mathbf{k}, \mathbf{k}'}, \\ G_0^r(\mathbf{k}, t; \mathbf{k}', t') &\equiv -i\hbar^{-1} \theta(t-t') \langle c_{\mathbf{k}}(t) c_{\mathbf{k}}^\dagger(t') + c_{\mathbf{k}}^\dagger(t') c_{\mathbf{k}}(t) \rangle_0, \\ &= -i\hbar^{-1} \theta(t-t') e^{-i\xi_{\mathbf{k}}(t-t')/\hbar} \delta_{\mathbf{k}, \mathbf{k}'}, \\ G_0^a(\mathbf{k}, t; \mathbf{k}', t') &\equiv +i\hbar^{-1} \theta(t'-t) \langle c_{\mathbf{k}}(t) c_{\mathbf{k}}^\dagger(t') + c_{\mathbf{k}}^\dagger(t') c_{\mathbf{k}}(t) \rangle_0, \\ &= +i\hbar^{-1} \theta(t'-t) e^{-i\xi_{\mathbf{k}}(t'-t)/\hbar} \delta_{\mathbf{k}, \mathbf{k}'}, \end{aligned} \quad (163)$$

where  $n_{\mathbf{k}} = \langle c_{\mathbf{k}}^\dagger c_{\mathbf{k}} \rangle$  is the average occupation number of the state  $\mathbf{k}$ . Green's functions depend only on time differences. One usually Fourier transforms the time difference coordinate,  $t - t'$ , to energy

$$\begin{aligned} G_0^<(\mathbf{k}, E) &= +2\pi i n_{\mathbf{k}} \delta(E - \xi_{\mathbf{k}}), \\ G_0^>(\mathbf{k}, E) &= +2\pi i [1 - n_{\mathbf{k}}] \delta(E - \xi_{\mathbf{k}}), \\ G_0^r(\mathbf{k}, E) &= \frac{1}{E - \xi_{\mathbf{k}} + i\eta}, \\ G_0^a(\mathbf{k}, E) &= \frac{1}{E - \xi_{\mathbf{k}} - i\eta}, \end{aligned} \quad (164)$$

where  $\eta = 0^+$  is a small positive number. Assuming that the particles are in thermal equilibrium, one obtains  $n_{\mathbf{k}} = n_F(\xi_{\mathbf{k}})$ , where  $n_F$  is the Fermi-Dirac distribution function. The result (164) shows that  $G^<$  and  $G^>$  provide information about the statistics, such as occupation  $n_{\mathbf{k}}$  or unoccupation  $1 - n_{\mathbf{k}}$  of the states, and  $G^r$  and  $G^a$  provide information about the states regardless of their occupation. The spectral function  $A_0(\mathbf{k}, E)$  for Fermions is therefore defined as

$$\begin{aligned} A_0(\mathbf{k}, E) &= +i [G_0^r(\mathbf{k}, E) - G_0^a(\mathbf{k}, E)] \\ &= -2\Im [G_0^r(\mathbf{k}, E)], \\ &= +2\pi \delta(E - \xi_{\mathbf{k}}), \end{aligned} \quad (165)$$

where the following relation is used:

$$\frac{1}{x \pm i\eta} = \mathcal{P} \left( \frac{1}{x} \right) \mp i\pi \delta(x) \quad (166)$$

where  $\mathcal{P}$  indicates the principal value. Under equilibrium the lesser and greater Green's functions can be rewritten as

$$\begin{aligned} G_0^<(\mathbf{k}, E) &= i n_F A_0(\mathbf{k}, E), \\ G_0^>(\mathbf{k}, E) &= i [1 - n_F] A_0(\mathbf{k}, E). \end{aligned} \quad (167)$$

## B.2. Noninteracting Bosons

The Hamiltonian for noninteracting phonons (Bosons) in momentum representation is

$$H_0 = \sum_{\mathbf{q}, \lambda} \hbar \omega_{\mathbf{q}, \lambda} \left( b_{\mathbf{q}, \lambda}^\dagger b_{\mathbf{q}, \lambda} + \frac{1}{2} \right), \quad (168)$$

where  $\hbar \omega_{\mathbf{q}, \lambda}$  is the energy of mode  $\mathbf{q}$  with the polarization  $\lambda$ ,  $b_{\mathbf{q}, \lambda}$ , and  $b_{\mathbf{q}, \lambda}^\dagger$  are the Bosons annihilation and creation operators. The time evolution of the annihilation operator in the Heisenberg picture is

$$b_{\mathbf{q}, \lambda}(t) = e^{iH_0 t/\hbar} b_{\mathbf{q}, \lambda} e^{-iH_0 t/\hbar}, \quad (169)$$

so the operator obeys the equation,

$$i\hbar \partial_t b_{\mathbf{q}, \lambda}(t) = [b_{\mathbf{q}, \lambda}(t), H_0] = \hbar \omega_{\mathbf{q}, \lambda} b_{\mathbf{q}, \lambda}(t), \quad (170)$$

which has the solution

$$b_{\mathbf{q}, \lambda}(t) = e^{-i\omega_{\mathbf{q}, \lambda} t} b_{\mathbf{q}, \lambda}. \quad (171)$$

The creation operator for Bosons is the just the Hermitian conjugate of  $b_{\mathbf{q}}$ , that is,

$$b_{\mathbf{q}, \lambda}^\dagger(t) = e^{+i\omega_{\mathbf{q}, \lambda} t} b_{\mathbf{q}, \lambda}^\dagger. \quad (172)$$

The noninteracting real-time Green's functions for Bosons in momentum representation are now given by

$$\begin{aligned} D_{\lambda_0}^<(\mathbf{q}, t; \mathbf{q}', t') &\equiv -i\hbar^{-1} \langle \hat{A}_{\mathbf{q}', \lambda'}^\dagger(t') \hat{A}_{\mathbf{q}, \lambda}(t) \rangle_0, \\ &= -i\hbar^{-1} \langle b_{\mathbf{q}', \lambda'}^\dagger(t') b_{\mathbf{q}, \lambda}(t) \\ &\quad + b_{-\mathbf{q}', \lambda'}(t') b_{-\mathbf{q}, \lambda}^\dagger(t) \rangle_0, \\ &= -i\hbar^{-1} \left[ e^{-i(\omega_{\mathbf{q}, \lambda} t - \omega_{\mathbf{q}', \lambda'} t')} \langle b_{\mathbf{q}', \lambda'}^\dagger b_{\mathbf{q}, \lambda} \rangle_0 \right. \\ &\quad \left. + e^{-i(\omega_{\mathbf{q}', \lambda'} t' - \omega_{\mathbf{q}, \lambda} t)} \langle b_{-\mathbf{q}', \lambda'} b_{-\mathbf{q}, \lambda}^\dagger \rangle_0 \right] \delta_{\mathbf{q}, \mathbf{q}'}, \\ &= -i\hbar^{-1} \left[ e^{-i\omega_{\mathbf{q}, \lambda}(t-t')} n_{\mathbf{q}, \lambda} \right. \\ &\quad \left. + e^{+i\omega_{\mathbf{q}, \lambda}(t-t')} (n_{\mathbf{q}, \lambda} + 1) \right] \delta_{\mathbf{q}, \mathbf{q}'}, \\ &\equiv D_{\lambda_0}^<(\mathbf{q}; t, t'), \end{aligned} \quad (173)$$

$$\begin{aligned}
D_{\lambda_0}^>(\mathbf{q}; t, t') &= D_{\lambda_0}^<(\mathbf{q}; t', t), \\
&= -i\hbar^{-1} \left[ e^{+i\omega_{\mathbf{q},\lambda}(t-t')} n_{\mathbf{q},\lambda} + e^{-i\omega_{\mathbf{q},\lambda}(t-t')} (n_{\mathbf{q},\lambda} + 1) \right], \\
D_{\lambda_0}^r(\mathbf{q}; t, t') &\equiv -i\hbar^{-1} \theta(t-t') \left\langle \hat{A}_{\mathbf{q},\lambda}^\dagger(t') \hat{A}_{\mathbf{q},\lambda}(t) + \hat{A}_{\mathbf{q},\lambda}(t) \hat{A}_{\mathbf{q},\lambda}^\dagger(t') \right\rangle_0, \\
&= -i\hbar^{-1} \theta(t-t') \left[ e^{-i\omega_{\mathbf{q},\lambda}(t-t')} - e^{+i\omega_{\mathbf{q},\lambda}(t-t')} \right], \\
D_{\lambda_0}^a(\mathbf{q}; t, t') &= -i\hbar^{-1} \theta(t'-t) \left[ e^{+i\omega_{\mathbf{q},\lambda}(t-t')} - e^{-i\omega_{\mathbf{q},\lambda}(t-t')} \right],
\end{aligned} \tag{173}$$

where  $\hat{A}_{\mathbf{q},\lambda}(t) = b_{\mathbf{q},\lambda}(t) + b_{-\mathbf{q},\lambda}^\dagger(t)$ ,  $\hat{A}_{\mathbf{q},\lambda}^\dagger(t) = \hat{A}_{-\mathbf{q},\lambda}(t)$ ,  $\omega_{-\mathbf{q},\lambda} = \omega_{\mathbf{q},\lambda}$ , and  $n_{\mathbf{q},\lambda} = \langle b_{\mathbf{q},\lambda}^\dagger b_{\mathbf{q},\lambda} \rangle$  is the occupation number of the state  $(\mathbf{q}, \lambda)$ , where under thermal equilibrium, one obtains  $n_{\mathbf{q},\lambda} = n_B(\hbar\omega_{\mathbf{q},\lambda})$ , with  $n_B$  denoting the Bose–Einstein distribution function. Green's functions depend only on time differences. One usually Fourier transforms the time difference coordinate,  $t, t'$ , to energy

$$\begin{aligned}
D_{\lambda_0}^<(\mathbf{q}, E) &= -2\pi i \left[ n_{\mathbf{q},\lambda} \delta(E - \hbar\omega_{\mathbf{q},\lambda}) + (n_{\mathbf{q},\lambda} + 1) \right. \\
&\quad \left. \times \delta(E + \hbar\omega_{\mathbf{q},\lambda}) \right], \\
D_{\lambda_0}^>(\mathbf{q}, E) &= -2\pi i \left[ n_{\mathbf{q},\lambda} \delta(E + \hbar\omega_{\mathbf{q},\lambda}) + (n_{\mathbf{q},\lambda} + 1) \right. \\
&\quad \left. \times \delta(E - \hbar\omega_{\mathbf{q},\lambda}) \right], \\
D_{\lambda_0}^r(\mathbf{q}, E) &= \frac{1}{E - \hbar\omega_{\mathbf{q},\lambda} + i\eta} - \frac{1}{E + \hbar\omega_{\mathbf{q},\lambda} + i\eta}, \\
D_{\lambda_0}^a(\mathbf{q}, E) &= \frac{1}{E - \hbar\omega_{\mathbf{q},\lambda} - i\eta} - \frac{1}{E + \hbar\omega_{\mathbf{q},\lambda} - i\eta},
\end{aligned} \tag{174}$$

### C. TREATMENT OF CONTACTS

In order to solve the transport equations, boundary conditions have to be specified. An important point is the treatment of contacts, which act as a source or drain for the carriers [7, 80, 98, 160, 216, 217]. Here, the method described in [217] is followed.

One can partition the layered structure into left contact with index  $L$ , device with index  $D$ , and right contact with index  $R$  (Fig. 27). The device corresponds to the region where one solves the transport equations and the contacts are the highly conducting regions connected to the device.

While the device region consists of only  $N$  layers, the matrices corresponding to Green's functions (87) and (88) are infinite dimensional owing to the semiinfinite contacts.

It is shown next that the influence of the semiinfinite contacts can be folded into the device region, where the semiinfinite contacts only affect layers 1 and  $N$  of the device region.

As shown next, the influence of the semiinfinite contacts can be folded into the device region by adding a self-energy to the device region. This can be viewed as an additional self-energy, owing to the transitions between the device and the contacts.

#### C.1. Matrix Truncation

By defining

$$\underline{A} = [\underline{E}\underline{I} - \underline{H} - \underline{\Sigma}_{\text{Scat}}], \tag{175}$$

eqn (87) ( $\underline{A}\underline{G}^r = \underline{I}$ ) can be written as

$$\begin{bmatrix} \underline{A}_{LL} & \underline{A}_{LD} & \\ \underline{A}_{DL} & \underline{A}_{DD} & \underline{A}_{DR} \\ & \underline{A}_{RD} & \underline{A}_{RR} \end{bmatrix} \begin{bmatrix} \underline{G}_{LL}^r & \underline{G}_{LD}^r & \underline{G}_{LR}^r \\ \underline{G}_{DL}^r & \underline{G}_{DD}^r & \underline{G}_{DR}^r \\ \underline{G}_{RL}^r & \underline{G}_{RD}^r & \underline{G}_{RR}^r \end{bmatrix} = \begin{bmatrix} \underline{I} & & \\ & \underline{I} & \\ & & \underline{I} \end{bmatrix}, \tag{176}$$

where

$$\underline{A}_{LL} = \begin{bmatrix} \bullet & \bullet & \bullet & & \\ & \bullet & \bullet & \bullet & \\ & & -t_{L4,3}^\dagger & \underline{A}_{L3} & -t_{L3,2} \\ & & & -t_{L3,2}^\dagger & \underline{A}_{L2} & -t_{L2,1} \\ & & & & -t_{L2,1}^\dagger & \underline{A}_{L1} \end{bmatrix}, \tag{177}$$

corresponds to the left semiinfinite contact,

$$\underline{A}_{RR} = \begin{bmatrix} \underline{A}_{R1} & -t_{R12} & & & \\ -t_{R12}^\dagger & \underline{A}_{R2} & -t_{R23} & & \\ & -t_{R23}^\dagger & \underline{A}_{R3} & -t_{R34} & \\ & & \bullet & \bullet & \bullet \\ & & & \bullet & \bullet & \bullet \end{bmatrix}, \tag{178}$$

corresponds to the right semiinfinite contact, and corresponds to the device region.

The coupling between the left and right contacts and device are, respectively, given by

$$\underline{A}_{DD} = \begin{bmatrix} \underline{A}_1 & -t_{12} & & & \\ -t_{12}^\dagger & \underline{A}_2 & -t_{2,3} & & \\ & \bullet & \bullet & \bullet & \\ & & \bullet & \bullet & \bullet \\ & & & -t_{N-2,N-1}^\dagger & \underline{A}_{N-1} & -t_{N-1,N} \\ & & & & -t_{N-1,N}^\dagger & \underline{A}_N \end{bmatrix}, \tag{179}$$

$$\underline{A}_{LD} = \begin{bmatrix} 0 & 0 & \bullet & \bullet & 0 & 0 \\ 0 & 0 & \bullet & \bullet & 0 & 0 \\ 0 & 0 & \bullet & \bullet & 0 & 0 \\ 0 & 0 & \bullet & \bullet & 0 & 0 \\ -t_{LD} & 0 & \bullet & \bullet & 0 & 0 \end{bmatrix}, \quad (180)$$

and

$$\underline{A}_{RD} = \begin{bmatrix} 0 & 0 & \bullet & \bullet & 0 & -t_{RD} \\ 0 & 0 & \bullet & \bullet & 0 & 0 \\ 0 & 0 & \bullet & \bullet & 0 & 0 \\ 0 & 0 & \bullet & \bullet & 0 & 0 \\ 0 & 0 & \bullet & \bullet & 0 & 0 \end{bmatrix}. \quad (181)$$

It should be noted that  $\underline{A}_{DL} = \underline{A}_{LD}^\dagger$ ,  $\underline{A}_{DR} = \underline{A}_{RD}^\dagger$ , and  $\underline{A}_{LD}$  and  $\underline{A}_{DL}$  ( $\underline{A}_{RD}$ , and  $\underline{A}_{DR}$ ) are sparse matrices. Their only nonzero entry represents the coupling of the left (right) contact and device. From (176), one obtains

$$\underline{G}_{LD}^r = -\underline{A}_{LL}^{-1} \underline{A}_{LD} \underline{G}_{DD}^r, \quad (182)$$

$$\underline{G}_{RD}^r = -\underline{A}_{RR}^{-1} \underline{A}_{RD} \underline{G}_{DD}^r, \quad (183)$$

$$\underline{A}_{DL} \underline{G}_{LD}^r + \underline{A}_{DD} \underline{G}_{DD}^r + \underline{A}_{DR} \underline{G}_{RD}^r = \underline{I}. \quad (184)$$

Substituting (182) and (183) in (184), one obtains a matrix equation with a dimension corresponding to the total number of grid points in device layers,

$$\left[ \underline{A}_{DD} - \underline{A}_{DL} \underline{A}_{LL}^{-1} \underline{A}_{LD} - \underline{A}_{DR} \underline{A}_{RR}^{-1} \underline{A}_{RD} \right] \underline{G}_{DD}^r = \underline{I}. \quad (185)$$

The second and third terms of (185) are self-energies owing to the coupling of the device region to left and right contacts, respectively.

Green's functions of the isolated semiinfinite contacts are defined as

$$\begin{aligned} \underline{A}_{LL} \underline{g}_L^r &= \underline{I}, \\ \underline{A}_{RR} \underline{g}_R^r &= \underline{I}. \end{aligned} \quad (186)$$

The *surface Green's function* of the left and right contacts are Green's function elements corresponding to the first-edge layer of the respective contact,

$$\begin{aligned} \underline{g}_{L,1}^r &= \underline{A}_{LL,1}^{-1}, \\ \underline{g}_{R,1}^r &= \underline{A}_{RR,1}^{-1}. \end{aligned} \quad (187)$$

## C.2. Contact Self-Energies

The surface Green's functions defined in (187) enable us to rewrite (185) in a form very similar to (87),

$$\left[ E\underline{I} - \underline{H} - \underline{\Sigma}_{\text{Scat}}^r - \underline{\Sigma}_C^r \right] \underline{G}_{DD}^r = \underline{I}, \quad (188)$$

where

$$\begin{aligned} \underline{\Sigma}_{C,1}^r &= t_{DL} \underline{g}_{L,1}^r t_{LD} = \underline{\Sigma}_L^r, \\ \underline{\Sigma}_{C,N}^r &= t_{DR} \underline{g}_{R,1}^r t_{RD} = \underline{\Sigma}_R^r. \end{aligned} \quad (189)$$

All the other elements of  $\underline{\Sigma}_C^r$  are zero.  $\underline{\Sigma}_L^r$  and  $\underline{\Sigma}_R^r$  are *self-energies* owing to the left and right contacts, respectively, and  $t_{DL} = t_{LD}^\dagger$  and  $t_{DR} = t_{RD}^\dagger$ . By following the same procedure, one obtains the equation of motion for the lesser and greater Green's functions as [98]

$$\underline{G}_{DD}^> = \underline{G}_{DD}^r \left[ \underline{\Sigma}_{\text{Scat}}^< + \underline{\Sigma}_C^r \right] \underline{G}_{DD}^a, \quad (190)$$

where

$$\underline{\Sigma}_{C,1}^> = t_{DL} \underline{g}_{L,1}^> t_{LD} = \underline{\Sigma}_L^> \quad (191)$$

$$\underline{\Sigma}_{C,N}^> = t_{DR} \underline{g}_{R,1}^> t_{RD} = \underline{\Sigma}_R^>$$

Since the contacts are, by definition, in equilibrium, one obtains (Appendix B.1)

$$\begin{aligned} \underline{g}_{1,1}^< &= i a_{1,1} f_L \\ \underline{g}_{1,1}^< &= i a_{1,1} f_R, \end{aligned} \quad (192)$$

where  $\underline{a} = i(\underline{g}^r - \underline{g}^a) = -2\Im[\underline{g}^r]$  is the spectral function and  $f_{L(R)}$  is the Fermi factor of the left (right) contact. By defining the broadening function as

$$\begin{aligned} \Gamma_{C,1} &= i(\underline{\Sigma}_{1,1}^r - \underline{\Sigma}_{1,1}^a) = t_{DL} \underline{a}_{1,1} t_{LD} = \Gamma_L, \\ \Gamma_{C,1} &= i(\underline{\Sigma}_{1,1}^r - \underline{\Sigma}_{1,1}^a) = t_{DR} \underline{a}_{1,1} t_{RD} = \Gamma_R, \end{aligned} \quad (193)$$

eqn (191) can be rewritten as

$$\begin{aligned} \underline{\Sigma}_L^< &= +i\Gamma_L f_L, \\ \underline{\Sigma}_R^< &= +i\Gamma_R f_R. \end{aligned} \quad (194)$$

In a similar manner, one can show that

$$\begin{aligned} \underline{\Sigma}_L^> &= -i\Gamma_L(1 - f_L), \\ \underline{\Sigma}_R^> &= -i\Gamma_R(1 - f_R). \end{aligned} \quad (195)$$

### C.3. Surface Green's Function

The main information needed to solve (188) are the surface Green's functions of  $\underline{g}_L^r$  and  $\underline{g}_R^r$ . Using the recursive relation [7], eqn (186) can be written as

$$\begin{aligned} \left[ \underline{A}_{L_i} - \underline{t}_{L_{i+1}} \underline{g}_{L_{i+1}}^r \underline{t}_{L_{i+1}}^\dagger \right] \underline{g}_{L_i}^r &= \underline{I}, \\ \left[ \underline{A}_{R_i} - \underline{t}_{R_{i-1}}^\dagger \underline{g}_{R_{i-1}}^r \underline{t}_{R_{i-1}} \right] \underline{g}_{R_i}^r &= \underline{I}. \end{aligned} \quad (196)$$

If the potential does not vary in the left and right contacts, and if the coupling between the different layers are equal, then  $\underline{A}_{LL}$  and  $\underline{A}_{RR}$  become semiinfinite periodic matrices with

$$\begin{aligned} \underline{A}_{L_1} &= \underline{A}_{L_2} = \underline{A}_{L_3} = \dots = \underline{A}_L, \\ \underline{A}_{R_1} &= \underline{A}_{R_2} = \underline{A}_{R_3} = \dots = \underline{A}_R, \\ \underline{t}_{L_{1,2}} &= \underline{t}_{L_{2,3}} = \underline{t}_{L_{3,4}} = \dots = \underline{t}_L, \\ \underline{t}_{R_{1,2}} &= \underline{t}_{R_{2,3}} = \underline{t}_{R_{3,4}} = \dots = \underline{t}_R. \end{aligned} \quad (197)$$

Under this condition, one obtains

$$\begin{aligned} \underline{g}_{L_{1,1}}^r &= \underline{g}_{R_{2,2}}^r = \dots = \underline{g}_L^r, \\ \underline{g}_{R_{1,1}}^r &= \underline{g}_{L_{2,2}}^r = \dots = \underline{g}_R^r. \end{aligned} \quad (198)$$

Therefore, the surface Green's functions can be obtained by solving the quadratic matrix equations,

$$\begin{aligned} \left[ \underline{A}_L - \underline{t}_L \underline{g}_L^r \underline{t}_L^\dagger \right] \underline{g}_L^r &= \underline{I}, \\ \left[ \underline{A}_R - \underline{t}_R^\dagger \underline{g}_R^r \underline{t}_R \right] \underline{g}_R^r &= \underline{I}. \end{aligned} \quad (199)$$

These equations can be solved iteratively by

$$\begin{aligned} \left[ \underline{A}_L - \underline{t}_L \underline{g}_L^{r(m-1)} \underline{t}_L^\dagger \right] \underline{g}_L^{r(m)} &= \underline{I}, \\ \left[ \underline{A}_R - \underline{t}_R^\dagger \underline{g}_R^{r(m-1)} \underline{t}_R \right] \underline{g}_R^{r(m)} &= \underline{I}, \end{aligned} \quad (200)$$

where  $m$  represents the iteration number. It should be noted that the solution to (199) is analytic if the dimension of  $\underline{A}_R$  is one.

## REFERENCES

1. S. Iijima, *Nature* (London) 354, 56 (1991).
2. R. Martel, T. Schmidt, H. Shea, T. Hertel, and P. Avouris, *Appl. Phys. Lett.* 73, 2447 (1998).
3. S. J. Tans, A. R. M. Verschueren, and C. Dekker, *Nature* (London) 393, 49 (1998).
4. J. Guo, S. Datta, and M. Lundstrom, *IEEE Trans. Electron. Devices* 51, 172 (2004).
5. M. Freitag, Y. Martin, J. Misewich, R. Martel, and P. Avouris, *Nano Lett.* 3, 1067 (2003).
6. M. Freitag, J. Chen, J. Tersoff, J. Tsang, Q. Fu, J. Liu, and P. Avouris, *Phys. Rev. Lett.* 93, 076803 (2004).
7. A. Svizhenko, M. P. Anantram, T. R. Govindan, B. Biegel, and R. Venugopal, *J. Appl. Phys.* 91, 2343 (2002).
8. R. Venugopal, Z. Ren, S. Datta, M. Lundstrom, and D. Jovanovic, *J. Appl. Phys.* 92, 3730 (2002).
9. A. Svizhenko and M. Anantram, *Phys. Rev. B* 72, 085430 (2005).
10. Y. Xue, S. Datta, and M. A. Ratner, *Chem. Phys.* 281, 151 (2002).
11. M. S. Dresselhaus, G. Dresselhaus, and R. Saito, *Phys. Rev. B* 45, 6234 (1992).
12. R. Saito, M. Fujita, G. Dresselhaus, and M. S. Dresselhaus, *Appl. Phys. Lett.* 60, 2240 (1992).
13. R. Saito, G. Dresselhaus, and M. Dresselhaus, "Physical Properties of Carbon Nanotubes," London: Imperial College Press, 1998.
14. S. Reich, C. Thomsen, and J. Maultzsch, "Carbon Nanotubes: Basic Concepts and Physical Properties," Weinheim, Germany: Wiley-VCH, 2004.
15. N. Hamada, S. Sawada, and A. Oshiyama, *Phys. Rev. Lett.* 68, 1579 (1992).
16. R. A. Jishi, D. Inomata, K. Nakao, M. S. Dresselhaus, and G. Dresselhaus, *J. Phys. Soc. (Japan)* 63, 2252 (1994).
17. V. Popov, *New J. Phys.* 6, 1 (2004).
18. S. Reich, C. Thomsen, and P. Ordejón, *Phys. Rev. B* 65, 155411 (2002).
19. V. Zólyomi and J. Kürti, *Phys. Rev. B* 70, 085403 (2004).
20. J. W. G. Wildöer, L. C. Venema, A. G. Rinzier, R. E. Smalley, and C. Dekker, *Nature* (London) 391, 59 (1998).
21. S. Datta, "Quantum Transport: From Atoms to Transistors," Cambridge: Cambridge University Press, 2005.
22. J. W. Mintmire and C. White, *Phys. Rev. Lett.* 81, 2506 (1998).
23. R. A. Jishi, L. Venkataraman, M. S. Dresselhaus, and G. Dresselhaus, *Chem. Phys. Lett.* 209, 77 (1993).
24. P. C. Eklunda, J. M. Holden, and R. A. Jishi, *Carbon* 33, 959 (1995).
25. J. Yu, R. K. Kalia, and P. Vashishta, *J. Chem. Phys.* 103, 6697 (1995).
26. M. Menon, E. Richter, and K. R. Subbaswamy, *J. Chem. Phys.* 104, 5875 (1995).
27. J. Kürti, G. Kresse, and H. Kuzmany, *Phys. Rev. B* 58, 8869 (1998).
28. D. Sánchez-Portal, E. Artacho, J. M. Soler, A. Rubio, and P. Ordejón, *Phys. Rev. B* 59, 12678 (1999).
29. O. Dubay, G. Kresse, and H. Kuzmany, *Phys. Rev. Lett.* 88, 235506 (2002).
30. O. Dubay and G. Kresse, *Phys. Rev. B* 67, 035401 (2003).
31. L.-H. Ye, B.-G. Liu, D.-S. Wang, and R. Han, *Phys. Rev. B* 69, 235409 (2004).
32. J. Maultzsch, S. Reich, C. Thomsen, H. Requardt, and P. Ordejón, *Phys. Rev. B* 92, 075501 (2004).
33. S. Piscanec, M. Lazzeri, F. Mauri, A. C. Ferrari, and J. Robertson, *Phys. Rev. B* 93, 185503 (2004).
34. K.-P. Bohnen, R. Heid, H. J. Liu, and C. T. Chan, *Phys. Rev. Lett.* 93, 245501 (2004).
35. D. Connétable, G.-M. Rignanese, J.-C. Charlier, and X. Blase, *Phys. Rev. Lett.* 94, 015503 (2005).
36. V. N. Popov, V. E. V. Doren, and M. Balkanski, *Phys. Rev. B* 59, 8355 (1999).
37. Z. M. Li, V. N. Popov, and Z. K. Tang, *Solid State Commun.* 130, 657 (2004).
38. I. Milošević, E. Dobardžić, and M. Damjanović, *Phys. Rev. B* 72, 085426 (2005).
39. C. Mapelli, C. Castiglioni, G. Zerbi, and K. Müllen, *Phys. Rev. B* 60, 12710 (1999).
40. V. N. Popov and P. Lambin, *Phys. Rev. B* 73, 085407 (2006).
41. S. O. Koswatta, S. Hasan, M. Lundstrom, M. P. Anantram, and D. E. Nikonov, *IEEE Trans. Electron. Devices* 54, 2339 (2007).

42. V. N. Popov and P. Lambin, *Phys. Rev. B* 74, 075415 (2006).
43. G. D. Mahan, *Phys. Rev. B* 68, 125409 (2003).
44. J. Jiang, R. Saito, G. G. Samsonidze, S. G. Chou, A. Jorio, G. Dresselhaus, and M. S. Dresselhaus, *Phys. Rev. B* 72, 235408 (2005).
45. V. N. Popov, V. E. V. Doren, and M. Balkanski, *Phys. Rev. B* 61, 3078 (2000).
46. W. Hoenlein, F. Kreupl, G. Duesberg, A. Graham, M. Liebau, R. Seidel, and E. Unger, *IEEE Trans. Comp. Packag. Technol.* 27, 629 (2004).
47. H. T. Soh, C. F. Quate, A. F. Morpurgo, C. Marcus, J. Kong, and H. Dai, *Appl. Phys. Lett.* 75, 627 (1999).
48. R. Martel, V. Derycke, C. Lavoie, J. Appenzeller, K. K. Chan, J. Tersoff, and P. Avouris, *Phys. Rev. Lett.* 87, 256805 (2001).
49. R. Seidel, M. Liebau, G. Duesberg, F. Kreupl, E. Unger, A. Graham, W. Hoenlein, and W. Pompe, *Nano Lett.* 3, 965 (2003).
50. A. Javey, J. Guo, Q. Wang, M. Lundstrom, and H. Dai, *Nature* (London) 424, 654 (2003).
51. V. Derycke, R. Martel, J. Appenzeller, and P. Avouris, *Nano Lett.* 1, 453 (2001).
52. S. Wind, J. Appenzeller, R. Martel, V. Derycke, and P. Avouris, *Appl. Phys. Lett.* 80, 3817 (2002).
53. J. Appenzeller, J. Knoch, R. Martel, V. Derycke, S. Wind, and P. Avouris, *IEEE Trans. Nanotechnol.* 1, 184 (2002).
54. A. Javey, H. Kim, M. Brink, Q. Wang, A. Ural, J. Guo, P. McIntyre, P. McEuen, M. Lundstrom, and H. Dai, *Nature Mater.* 1, 241 (2002).
55. M. Radosavljevic, J. Appenzeller, P. Avouris, and J. Knoch, *Appl. Phys. Lett.* 84, 3693 (2004).
56. F. Leonard and J. Tersoff, *Phys. Rev. Lett.* 83, 5174 (1999).
57. F. Leonard and J. Tersoff, *Phys. Rev. Lett.* 84, 4693 (2000).
58. J. Appenzeller, M. Radosavljevic, J. Knoch, and P. Avouris, *Phys. Rev. Lett.* 92, 048301 (2004).
59. R. V. Seidel, A. P. Graham, J. Kretz, B. Rajasekharan, G. S. Duesberg, M. Liebau, E. Unger, F. Kreupl, and W. Hoenlein, *Nano Lett.* 5, 147 (2005).
60. M. H. Yang, K. B. K. Teo, L. Gangloff, W. I. Milne, D. G. Hasko, Y. Robert, and P. Legagneux, *Appl. Phys. Lett.* 85, 113507 (2006).
61. J. Appenzeller, J. Knoch, V. Derycke, R. Martel, S. Wind, and P. Avouris, *Phys. Rev. Lett.* 89, 126801 (2002).
62. S. Heinze, J. Tersoff, R. Martel, V. Derycke, J. Appenzeller, and P. Avouris, *Phys. Rev. Lett.* 89, 106801 (2002).
63. M. Freitag, M. Radosavljevic, Y. X. Zhou, A. T. Johnson, and W. F. Smith, *Appl. Phys. Lett.* 79, 3326 (2001).
64. S. Wind, J. Appenzeller, and P. Avouris, *Phys. Rev. Lett.* 91, 058301 (2003).
65. Z. Chen, J. Appenzeller, J. Knoch, Y.-M. Lin, and P. Avouris, *Nano Lett.* 5, 1497 (2005).
66. P. G. Collins, K. Bradley, M. Ishigami, and A. Zettl, *Science* 287, 1801 (2000).
67. V. Derycke, R. Martel, J. Appenzeller, and P. Avouris, *Appl. Phys. Lett.* 80, 2773 (2002).
68. T. Rueckes, K. Kim, E. Joselevich, G. Y. Tseng, C. L. Cheung, and C. M. Lieber, *Science* 289, 94 (2000).
69. J. Kong, N. R. Franklin, C. W. Zhou, M. G. Chapline, S. Peng, K. J. Cho, and H. J. Dai, *Science* 287, 622 (2000).
70. A. Bachtold, P. Hadley, T. Nakanishi, and C. Dekker, *Science* 294, 1317 (2001).
71. Z. Chen, J. Appenzeller, Y.-M. Lin, J. Sippel-Oakley, A. G. Rinzier, J. Tang, S. J. Wind, P. M. Solomon, and P. Avouris, *Science* 311, 1735 (2006).
72. D. Frank and J. Appenzeller, *IEEE Electron. Device Lett.* 25, 34 (2004).
73. X. Huo, M. Zhang, P. C. H. Chan, Q. Liang, and Z. K. Tang, in "IEDM Tech. Dig., IEEE, San Francisco, 2004, pp. 691–694.
74. S. Rosenblatt, H. Lin, V. Sazonova, S. Tiwari, and P. L. McEuen, *Appl. Phys. Lett.* 87, 153111 (2005).
75. J. Schwinger, *J. Math. Phys.* 2, 407 (1961).
76. L. P. Kadanoff and G. Baym, "Quantum Statistical Mechanics: Green's Function Methods in Equilibrium and NonEquilibrium Problems," New York, W.A.: Benjamin, 1962.
77. C. B. Duke, "Tunneling in Solids," New York: Academic Press, 1969.
78. G. Kim and G. B. Arnold, *Phys. Rev. B* 38, 3252 (1988).
79. C. Caroli, R. Combescot, P. Nozieres, and D. Saint-James, *J. Phys. C: Solid State Phys.* 4, 916 (1971).
80. C. Caroli, R. Combescot, D. Lederer, P. Nozieres, and D. Saint-James, *J. Phys. C: Solid State Phys.* 4, 2598 (1971).
81. R. Combescot, *J. Phys. C: Solid State Phys.* 4, 2611 (1971).
82. C. Caroli, R. Combescot, P. Nozieres, and D. Saint-James, *J. Phys. C: Solid State Phys.* 5, 21 (1972).
83. W. R. Bandy and A. J. Glick, *Phys. Rev. B* 13, 3368 (1976).
84. W. R. Bandy and A. J. Glick, *Phys. Rev. B* 16, 2346 (1977).
85. M. Cini, *Phys. Rev. B* 22, 5887 (1980).
86. H. Haug, Ed., "Optical Nonlinearities and Instabilities in Semiconductors," Boston: Academic Press, 1988.
87. A. P. Jauho and J. W. Wilkins, *Phys. Rev. B* 29, 1919 (1984).
88. G. D. Mahan, *Phys. Rep.* 145, 251 (1987).
89. R. Bertoni and A.-P. Jauho, *Phys. Rev. Lett.* 68, 2826 (1992).
90. E. V. Anda and F. Flores, *J. Phys. Cond. Matt.* 3, 9087 (1991).
91. L. Y. Chen and C. S. Ting, *Phys. Rev. Lett.* 64, 3159 (1990).
92. J. Zang and J. L. Birman, *Phys. Rev. B* 46, 5020 (1992).
93. R. K. Lake and S. Datta, *Superlatt. Microstruct.* 11, 83 (1992).
94. C. H. Grein, E. Runge, and H. Ehrenreich, *Phys. Rev. B* 47, 12590 (1993).
95. R. Lake and S. Datta, *Phys. Rev. B* 45, 6670 (1992).
96. A. L. Yeyati, F. Flores, and E. V. Anda, *Phys. Rev. B* 47, 10543 (1993).
97. R. Lake, G. Klimeck, R. C. Bowen, C. Fernando, M. Leng, T. Moise, and Y. C. Kao, *Superlatt. Microstruct.* 20, 279 (1996).
98. R. Lake, G. Klimeck, R. C. Bowen, and D. Jovanovic, *J. Appl. Phys.* 81, 7845 (1997).
99. G. Kim, H. Suh, and E. Lee, *Phys. Rev. B* 52, 2632 (1995).
100. L. Y. Chen and C. S. Ting, *Phys. Rev. B* 43, 4534 (1991).
101. R. Lake and S. Datta, *Phys. Rev. B* 46, 4757 (1992).
102. G. Klimeck, R. Lake, R. C. Bowen, and W. R. Frensley, *Appl. Phys. Lett.* 67, 2539 (1995).
103. R. C. Bowen, G. Klimeck, R. K. Lake, W. R. Frensley, and T. Moise, *J. Appl. Phys.* 81, 3207 (1997).
104. M. J. McLennan, Y. Lee, and S. Datta, *Phys. Rev. B* 43, 13846 (1991).
105. S.-C. Lee and A. Wacker, *Phys. Rev. B* 66, 245314 (2002).
106. C. Rivas, R. Lake, G. Klimeck, W. R. Frensley, M. V. Fischetti, P. E. Thompson, S. L. Rommel, and P. R. Berger, *Appl. Phys. Lett.* 78, 814 (2001).
107. C. Rivas, R. Lake, W. R. Frensley, G. Klimeck, P. E. Thompson, S. L. Rommel, and P. R. Berger, *J. Appl. Phys.* 94, 5005 (2003).
108. R. Lake, D. Jovanovic, and C. Rivas, in "Progress in Nonequilibrium Green Functions," pp. 143–158, New Jersey: World Scientific, 2003.
109. A. Svizhenko and M. P. Anantram, *IEEE Trans. Electron. Devices* 50, 1459 (2003).
110. R. Venugopal, M. Paulsson, S. Goasguen, S. Datta, and M. S. Lundstrom, *J. Appl. Phys.* 93, 5613 (2003).
111. Y.-J. Ko, M. Shin, S. Lee, and K. W. Park, *J. Appl. Phys.* 89, 374 (2001).
112. C. Rivas and R. Lake, *Phys. Status Solidi B* 239, 94 (2003).
113. M. B. Nardelli, *Phys. Rev. B* 60, 7828 (1999).
114. J. Taylor, H. Guo, and J. Wang, *Phys. Rev. B* 63, 245407 (2001).
115. H. Mehrez, J. Taylor, H. Guo, J. Wang, and C. Roland, *Phys. Rev. Lett.* 84, 2682 (2000).
116. D. Orlikowski, H. Mehrez, J. Taylor, H. Guo, J. Wang, and C. Roland, *Phys. Rev. B* 63, 155412 (2001).
117. J. Taylor, H. Guo, and J. Wang, *Phys. Rev. B* 63, 121104 (2001).
118. M. P. Anantram, *Appl. Phys. Lett.* 78, 2055 (2001).
119. A. Maiti, A. Svizhenko, and M. P. Anantram, *Phys. Rev. Lett.* 88, 126805 (2002).

120. C.-C. Kaun, B. Larade, H. Mehrez, J. Taylor, and H. Guo, *Phys. Rev. B* 65, 205416 (2002).
121. G. Cuniberti, R. Gutiérrez, G. Fagas, F. Grossmann, K. Richter, and R. Schmidt, *Physica E* 12, 749 (2002).
122. J. J. Palacios, A. J. P. Jimenez, E. Louis, E. SanFabioa, and J. Verges, *Phys. Rev. Lett.* 90, 106801 (2003).
123. Y. Xue and M. A. Ratner, *Appl. Phys. Lett.* 83, 2429 (2003).
124. T.-S. Xia, L. F. Register, and S. K. Banerjee, *J. Appl. Phys.* 95, 1597 (2004).
125. M. Brandbyge, J.-L. Mozos, P. Ordejón, J. Taylor, and K. Stokbro, *Phys. Rev. B* 65, 165401 (2002).
126. E. Louis, J. A. Vergés, J. J. Palacios, A. J. Pérez-Jiménez, and E. SanFabian, *Phys. Rev. B* 67, 155321 (2003).
127. W. Tian, S. Datta, S. Hong, R. Reifenberger, J. I. Henderson, and C. P. Kubiak, *J. Chem. Phys.* 109, 2874 (1998).
128. P. A. Derosa and J. M. Seminario, *J. Phys. Chem. B* 105, 471 (2001).
129. Y. Xue, S. Datta, and M. A. Ratner, *J. Chem. Phys.* 115, 4292 (2001).
130. Y. Xue, Dissertation, Purdue University, 2000.
131. P. S. Damle, A. W. Ghosh, and S. Datta, *Phys. Rev. B* 64, 201403 (2001).
132. J. M. Seminario, A. G. Zacarias, and P. A. Derosa, *J. Phys. Chem. A* 105, 791 (2001).
133. J. M. Seminario and P. A. Derosa, *J. Am. Chem. Soc.* 123, 12418 (2001).
134. J. Taylor, M. Brandbyge, and K. Stokbro, *Phys. Rev. Lett.* 89, 138301 (2002).
135. J. Palacios, E. Louis, A. J. Pérez-Jiménez, E. S. Fabián, and J. Vergés, *Nanotechnology* 13, 378 (2002).
136. J. Heurich, J. C. Cuevas, W. Wenzel, and G. Schön, *Phys. Rev. Lett.* 88, 256803 (2002).
137. Y. Xue and M. A. Ratner, *Phys. Rev. B* 68, 115406 (2003).
138. J. M. Seminario, L. E. Cordova, and P. A. Derosa, *Proc. IEEE* 91, 1958 (2003).
139. M. Galperin, A. Nitzan, S. Sek, and M. Majda, *J. Electroanal. Chem.* 550–551, 337 (2003).
140. P. A. Derosa, S. Guda, and J. M. Seminario, *J. Am. Chem. Soc.* 125, 14240 (2003).
141. Y. Xue and M. A. Ratner, *Phys. Rev. B* 69, 085403 (2004).
142. A. W. Ghosh, T. Rakshit, and S. Datta, *Nano Lett.* 4, 565 (2004).
143. F. Evers, F. Weigend, and M. Koentopp, *Phys. Rev. B* 69, 235411 (2004).
144. U. Gunsenheimer and A. D. Zaikin, *Phys. Rev. B* 50, 6317 (1994).
145. W. I. Babiczayk and B. R. Bulka, *J. Phys. Cond. Matt.* 16, 4001 (2004).
146. S. Krompiewski, *J. Phys. Cond. Matt.* 16, 2981 (2004).
147. M. A. Davidovich, E. V. Anda, C. Tejedor, and G. Platero, *Phys. Rev. B* 47, 4475 (1993).
148. G. Klimeck, R. Lake, and D. K. Blanks, *Phys. Rev. B* 58, 7279 (1998).
149. A. Groshev, T. Ivanov, and V. Valtchinov, *Phys. Rev. Lett.* 66, 1082 (1991).
150. L. Y. Chen and C. S. Ting, *Phys. Rev. B* 44, 5916 (1991).
151. S. Hershfield, J. Davies, and J. Wilkins, *Phys. Rev. Lett.* 67, 3720 (1991).
152. S. Hershfield, J. Davies, and J. Wilkins, *Phys. Rev. B* 46, 7046 (1992).
153. Y. Meir, N. S. Wingreen, and P. A. Lee, *Phys. Rev. Lett.* 70, 2601 (1993).
154. N. S. Wingreen and Y. Meir, *Phys. Rev. B* 49, 11040 (1994).
155. Y. Xue and M. A. Ratner, *Phys. Rev. B* 68, 235410 (2003).
156. S. Hershfield, *Phys. Rev. B* 46, 7061 (1992).
157. S. Datta and M. P. Anantram, *Phys. Rev. B* 45, 13761 (1991).
158. M. P. Anantram and S. Datta, *Phys. Rev. B* 51, 7632 (1995).
159. N. S. Wingreen, A. Jauho, and Y. Meir, *Phys. Rev. B* 48, 8487 (1993).
160. A. Jauho, N. S. Wingreen, and Y. Meir, *Phys. Rev. B* 60, 5528 (1994).
161. C. A. Stafford and N. S. Wingreen, *Phys. Rev. Lett.* 76, 1916 (1996).
162. P. Kral, *Phys. Rev. B* 53, 11034 (1996).
163. P. Kral and A.-P. Jauho, *Phys. Rev. B* 59, 7656 (1999).
164. H. Haug and A.-P. Jauho, "Quantum Kinetics in Transport and Optics of Semiconductors," Springer Series in Solid-State Sciences, Vol. 123, Springer: Berlin, 1996.
165. S. Datta, "Electronic Transport in Mesoscopic Systems," New York: Cambridge University Press, 1995.
166. S. Datta, *Superlatt. Microstruct.* 28, 253 (2000).
167. R. K. Lake and R. R. Pandey, in "Handbook of Semiconductor Nanostructures and Devices" (A. A. Balandin and K. L. Wang, Eds.), Vol. 3, Chap. Nonequilibrium Green Functions in Electronic Device Modeling, pp. 409–443, Los Angeles: American Scientific Publishers, 2006.
168. W. Fichtner, Technical Report, Integrated Systems Laboratory, ETH Zürich [unpublished].
169. A. L. Fetter and J. D. Walecka, "Quantum Theory of Many-Particle Systems," San Francisco: McGraw-Hill, 1971.
170. G. D. Mahan, "Many-Particle Physics, Physics of Solids and Liquids," 2nd ed., New York: Plenum Press, 1990.
171. H. J. Kreuzer, "Nonequilibrium Thermodynamics and Its Statistical Foundations," New York: Oxford University Press, 1981.
172. V. Korenman, *Ann. Phys.* 39, 72 (1966).
173. A. Schmid, *J. Low. Temp. Phys.* 49, 609 (1982).
174. L. V. Keldysh, *Soviet Phys. JETP* 20, 1018 (1965).
175. R. P. Feynman, *Phys. Rev.* 76, 769 (1949).
176. P. Danielewicz, *Ann. Phys.* 152, 239 (1984).
177. W. Schäfer and M. Wegener, "Semiconductor Optics and Transport Phenomena," Berlin: Springer, 2002.
178. D. C. Langreth, in "Linear and Nonlinear Electron Transport in Solids" (J. T. Devreese and E. van Doren, Eds.), NATO ASI Ser., Ser. B, Vol. 17, pp. 3–18, New York: Plenum Press, 1976.
179. M. Karner, A. Gehring, S. Holzer, M. Pourfath, M. Wagner, W. Goes, M. Vasicek, et al., *J. Comput. Electron.* 6, 179 (2007).
180. M. Krüger, M. R. Buitelaar, T. Nussbaumer, C. Schönenbergera, and L. Forró, *Appl. Phys. Lett.* 78, 1291 (2001).
181. S. Rosenblatt, Y. Yaish, J. Park, J. Gore, V. Sazonova, and P. L. McEuen, *Nano Lett.* 2, 869 (2002).
182. W. B. Choi, J. U. Chu, K. S. Jeong, E. J. Bae, J. W. Lee, J. J. Kim, and J. O. Lee, *Appl. Phys. Lett.* 79, 3696 (2001).
183. A. V. Melechko, V. I. Merkulov, T. E. McKnight, M. A. Guillorn, K. L. Klein, D. H. Lowndes, and M. L. Simpson, *J. Appl. Phys.* 97, 041301 (2005).
184. F. Zahid, A. Ghosh, M. Paulsson, E. Polizzi, and S. Datta, *Phys. Rev. B* 70, 245317 (2004).
185. D. John, L. Castro, P. Pereira, and D. Pulfrey, in "Proc. NSTI Nanotech," Vol. 3, pp. 65–68, 2004.
186. A. Okabe, B. Boots, and K. Sugihara, "Spatial Tessellations: Concepts and Applications of Voronoi Diagrams," New York: Wiley, 1992.
187. S. Selberherr, "Analysis and Simulation of Semiconductor Devices," Vienna: Springer-Verlag, 1984.
188. F. Leonard and J. Tersoff, *Appl. Phys. Lett.* 81, 4835 (2002).
189. D. Z. Y. Ting and Y.-C. Change, *Phys. Rev. B* 36, 4359 (1987).
190. J. Guo, S. Datta, M. Lundstrom, and M. Anantram, *Int. J. Multiscale Comput. Eng.* 2, 257 (2004).
191. N. Nemecek, D. Tománek, and G. Cuniberti, *Phys. Rev. Lett.* 96, 076802 (2006).
192. D. Kienle and A. W. Ghosh, *J. Comput. Electron.* 4, 97 (2005).
193. A. B. Migdal, *Soviet Phys. JETP* 7, 996 (1958).
194. P. Davis and P. Rabinowitz, "Methods of Numerical Integration," 2nd ed., Orlando: Academic Press, 1984.
195. J. Lyness, *J. ACM* 16, 483 (1969).
196. T. Espelid, *BIT* 43, 319 (2003).
197. T. Espelid, "Numerical Integration, Recent Developments, Software and Applications, NATO ASI Ser., Ser. C," Vol. 357, p. 367, 1992.



198. M. Malcolm and R. Simpson, *ACM Trans. Math. Soft.* 1, 129 (1975).
199. O. Pinaud, *J. Appl. Phys.* 92, 1987 (2002).
200. S. Laux, A. Kumar, and M. Fischetti, *J. Appl. Phys.* 95, 5545 (2004).
201. F. Stern, *J. Comput. Phys.* 6, 56 (1970).
202. T. Kerkhoven, A. Galick, U. Ravaioli, J. Arends, and Y. Saad, *J. Appl. Phys.* 68, 3461 (1990).
203. A. Trellakis, A. T. Galick, A. Pacelli, and U. Ravaioli, *J. Appl. Phys.* 81, 7880 (1997).
204. F. Venturi, R. Smith, E. Sangiorgi, M. Pinto, and B. Ricco, *IEEE Trans. Comput. Aid. Des.* 8, 360 (1989).
205. A. Pacelli, *IEEE Trans. Electron. Dev.* 44, 1169 (1997).
206. M. Pourfath and H. Kosina, in "Large Scale Scientific Computing, Lecture Notes in Computer Science" (I. Lirkov, S. Margenov, and J. Wasniewski, Eds.), Vol. 3743, pp. 578–585, Springer, 2006, Chap. Fast Convergent Schrödinger–Poisson Solver for the Static and Dynamic Analysis of Carbon Nanotube Field Effect Transistors.
207. S. O. Koswatta, S. Hasan, M. Lundstrom, M. P. Anantram, and D. E. Nikonov, *Appl. Phys. Lett.* 89, 023125 (2006).
208. M. Pourfath and H. Kosina, *IOP J. Nanotechnol.* 18, 424036 (2007).
209. M. Lundstrom, *IEEE Electron. Dev. Lett.* 18, 361 (1997).
210. J. Park, S. Rosenblatt, Y. Yaish, V. Sazonova, H. Ustunel, S. Braig, T. Arias, P. Brouwer, and P. McEuen, *Nano Lett.* 4, 517 (2004).
211. A. Javey, J. Guo, D. Farmer, Q. Wang, E. Yenilmez, R. Gordon, M. Lundstrom, and H. Dai, *Nano Lett.* 4, 1319 (2004).
212. D. Singh, K. Jenkins, J. Appenzeller, D. Neumayer, A. Grill, and H.-S. P. Wong, *IEEE Trans. Nanotechnol.* 3, 383 (2004).
213. Y. Yoon, Y. Ouyang, and J. Guo, *IEEE Trans. Electron. Dev.* 53, 2467 (2006).
214. A. L. Louarn, F. Kapche, J.-M. Bethoux, H. Happy, G. Dambrine, V. Derycke, P. Chenevier, N. Izard, M. F. Goffman, and J. Bourgoin, *Appl. Phys. Lett.* 90, 233108 (2007).
215. M. Pourfath, H. Kosina, and S. Selberherr, *IOP J. Phys.: Conf. Ser.* 38, 29 (2006).
216. A. Wacker, *Phys. Rep.* 357, 1 (2002).
217. M. P. Anantram, M. S. Lundstrom, and D. E. Nikonov, cond-mat /0610247, 2006.

